

**Document Version**

Final published version

**Licence**

CC BY

**Citation (APA)**

Huber, S. T., Terwiel, D., Evers, W. H., Maresca, D., & Jakobi, A. J. (2023). Cryo-EM structure of gas vesicles for buoyancy-controlled motility. *Cell*, 186(5), 975-986.e13. <https://doi.org/10.1016/j.cell.2023.01.041>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.  
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

**Sharing and reuse**

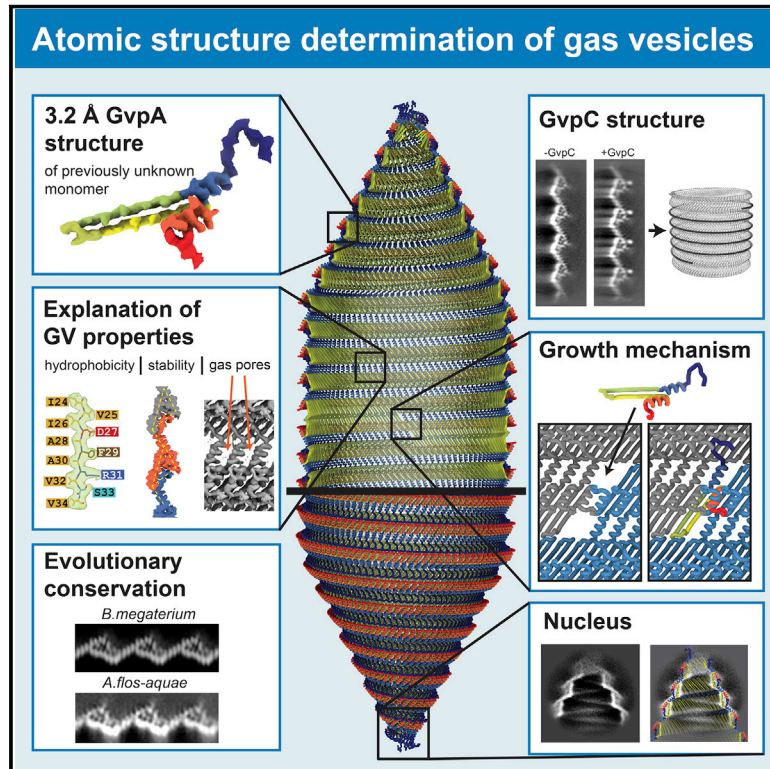
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Cryo-EM structure of gas vesicles for buoyancy-controlled motility

## Graphical abstract



## Authors

Stefan T. Huber, Dion Terwiel,  
Wiel H. Evers, David Maresca,  
Arjen J. Jakobi

## Correspondence

d.maresca@tudelft.nl (D.M.),  
a.jakobi@tudelft.nl (A.J.J.)

## In brief

Cryo-EM structure of the gas vesicle shell provides insights into the evolutionarily conserved mechanistic basis of gas vesicle biogenesis and selective permeability.

## Highlights

- 3.2 Å cryo-EM structure of microbial gas vesicles reveals atomic model of vesicle shell
- The structure explains the molecular basis of selective permeability and vesicle growth
- Fold and assembly are highly conserved between evolutionarily distant species
- Binding interface and geometry of secondary structural protein GvpC are identified



## Article

# Cryo-EM structure of gas vesicles for buoyancy-controlled motility

Stefan T. Huber,<sup>1</sup> Dion Terwiel,<sup>2</sup> Wiel H. Evers,<sup>1</sup> David Maresca,<sup>2,\*</sup> and Arjen J. Jakobi<sup>1,3,\*</sup><sup>1</sup>Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft 2628CD, the Netherlands<sup>2</sup>Department of Imaging Physics, Delft University of Technology, Delft 2628CD, the Netherlands<sup>3</sup>Lead contact\*Correspondence: [d.maresca@tudelft.nl](mailto:d.maresca@tudelft.nl) (D.M.), [a.jakobi@tudelft.nl](mailto:a.jakobi@tudelft.nl) (A.J.J.)<https://doi.org/10.1016/j.cell.2023.01.041>

## SUMMARY

Gas vesicles are gas-filled nanocompartments that allow a diverse group of bacteria and archaea to control their buoyancy. The molecular basis of their properties and assembly remains unclear. Here, we report the 3.2 Å cryo-EM structure of the gas vesicle shell made from the structural protein GvpA that self-assembles into hollow helical cylinders closed off by cone-shaped tips. Two helical half shells connect through a characteristic arrangement of GvpA monomers, suggesting a mechanism of gas vesicle biogenesis. The fold of GvpA features a corrugated wall structure typical for force-bearing thin-walled cylinders. Small pores enable gas molecules to diffuse across the shell, while the exceptionally hydrophobic interior surface effectively repels water. Comparative structural analysis confirms the evolutionary conservation of gas vesicle assemblies and demonstrates molecular features of shell reinforcement by GvpC. Our findings will further research into gas vesicle biology and facilitate molecular engineering of gas vesicles for ultrasound imaging.

## INTRODUCTION

Microorganisms utilize active motility systems to move toward or away from a variety of environmental stimuli such as chemicals and light.<sup>1</sup> These include swimming by rotation of rigid flagella and movement over solid surfaces with filamentous appendages.<sup>2</sup> Other forms of motility do not rely on active propulsion. Aquatic bacteria and archaea have evolved mechanisms to regulate buoyancy and can—similar to ballast tanks in submarines—create and eliminate gas-filled compartments to allow vertical migration in the water column. The cellular compartments providing positive buoyancy are formed by gas-filled protein shells called gas vesicles (GVs).<sup>3</sup>

There are very specific requirements for such structures: to achieve net buoyancy, GV must occupy a substantial proportion of the cell, which involves forming compartments that extend over hundreds of nanometers in size. To maximize buoyancy the shell must be constructed from minimal material. At the same time, the shell needs to provide resistance to hydrostatic pressure to maintain buoyancy with changes in water depth.<sup>4</sup>

GVs have therefore evolved as rigid, thin-walled structures composed of a single protein unit that typically polymerizes into large cylindrical shells closed off by conical tips.<sup>5,6</sup> The shell allows gas to diffuse passively between the GV lumen and the surrounding liquid, while effectively repelling water.<sup>7</sup> All GV identified to date appear to be constructed from the same components.<sup>8</sup> The ~7 kDa primary GV protein GvpA forms the core of the GV shell and the cone-shaped tips. A second protein,

GvpC, binds the exterior of the GV and provides additional structural reinforcement.<sup>9,10</sup>

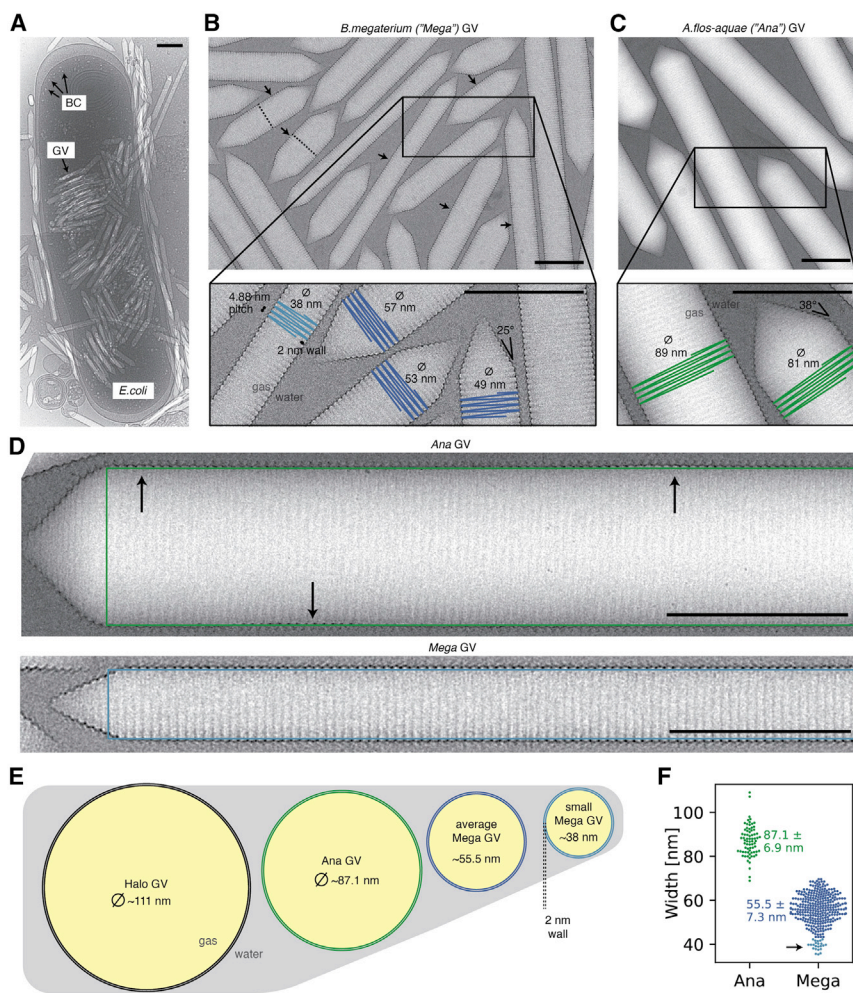
With molar masses exceeding hundreds of MDa, GV range among the largest protein-based macromolecular assemblies reported to date. Despite intensive efforts,<sup>5,11–17</sup> the molecular structure of GV and therefore a molecular-level understanding of their distinctive properties have remained elusive. Here, we present the cryogenic electron microscopy (cryo-EM) structure of the canonical GV shell, providing detailed insight into the biogenesis of GV and the unique evolutionary adaptations that enable buoyancy-controlled motility.

## RESULTS

### Cryo-EM structure of the gas vesicle wall

We expressed and purified *Bacillus megaterium* GV that form narrow tubes most suitable for cryo-EM analysis (Figure 1). The small diameter enables preparation of samples with thin ice to maximize contrast and reduces the propensity of GV to deform from ideal cylindrical shape. The native *B. megaterium* GV gene cluster contains two almost identical GvpA homologs, named GvpA and GvpB; for consistency in naming convention, we will refer to them as GvpA1 and GvpA2. A minimal gene cluster including GvpA2 but lacking GvpA1 is sufficient for GV assembly in *E. coli*<sup>19,20</sup> (Figure 1A). Cryo-EM images showed GV forming 0.1–1 μm long cylinders with varying diameters (55 ± 7 nm), consistent with previous data.<sup>18</sup> A subset of images (~16%) contained GV with diameters as small as 34–42 nm





**Figure 1. Gas vesicles**

(A) Cryo-EM micrograph of an *E. coli* cell heterologously producing *B. megaterium* gas vesicles (GVs). Mature GV and small bicone (BC) nuclei are visible inside bacteria and in the surrounding medium.

(B) Cryo-EM micrograph of purified *B. megaterium* GVs. GV appears brighter than the surrounding solvent due to the lower density of the gas inside the GV lumen. Dashed lines on a subset of GVs represent midpoints of the helical GV segment; arrows indicate the location of the seam. The inset shows close-up examples of average-sized GVs (blue) and a small diameter GV (light blue) used for structure determination. The 4.88 Å helical pitch, the 2 nm-thick wall, and the 25° cone angle are annotated.

(C) Representative cryo-EM micrograph of Ana GVs with inset showing a close-up indicating diameters and the ~38° cone angle.

(D) *A. flos-aquae* (Ana) GVs with mean diameter of 87 nm deform in cryo-EM sample preparation and do not stay perfectly cylindrical. The green rectangle depicts the ideal, non-deformed shape of an ideal cylinder in projection. Arrows point at deviations of the GV from the ideal shape. Helical reconstruction in cryo-EM assumes perfect helical crystals of the repeating unit. This assumption is violated when GVs “squish” during sample preparation. Very thin *B. megaterium* (Mega) GVs maintain a cylindrical shape in the thin ice layer of the cryo-EM sample.

(E) Schematic cross-sections of *H. salinarum*, *A. flos-aquae*, and *B. megaterium* GVs drawn to scale. Average diameter of *H. salinarum* GV from Dutka et al.<sup>18</sup>

(F) Width measurements in 20 micrographs from both *A. flos-aquae* and *B. megaterium* GV datasets (this study). A long tail of small-diameter outliers with 34–42 nm diameter GV (black arrow) was observed in *B. megaterium* GVs.

All scale bars 100 nm.

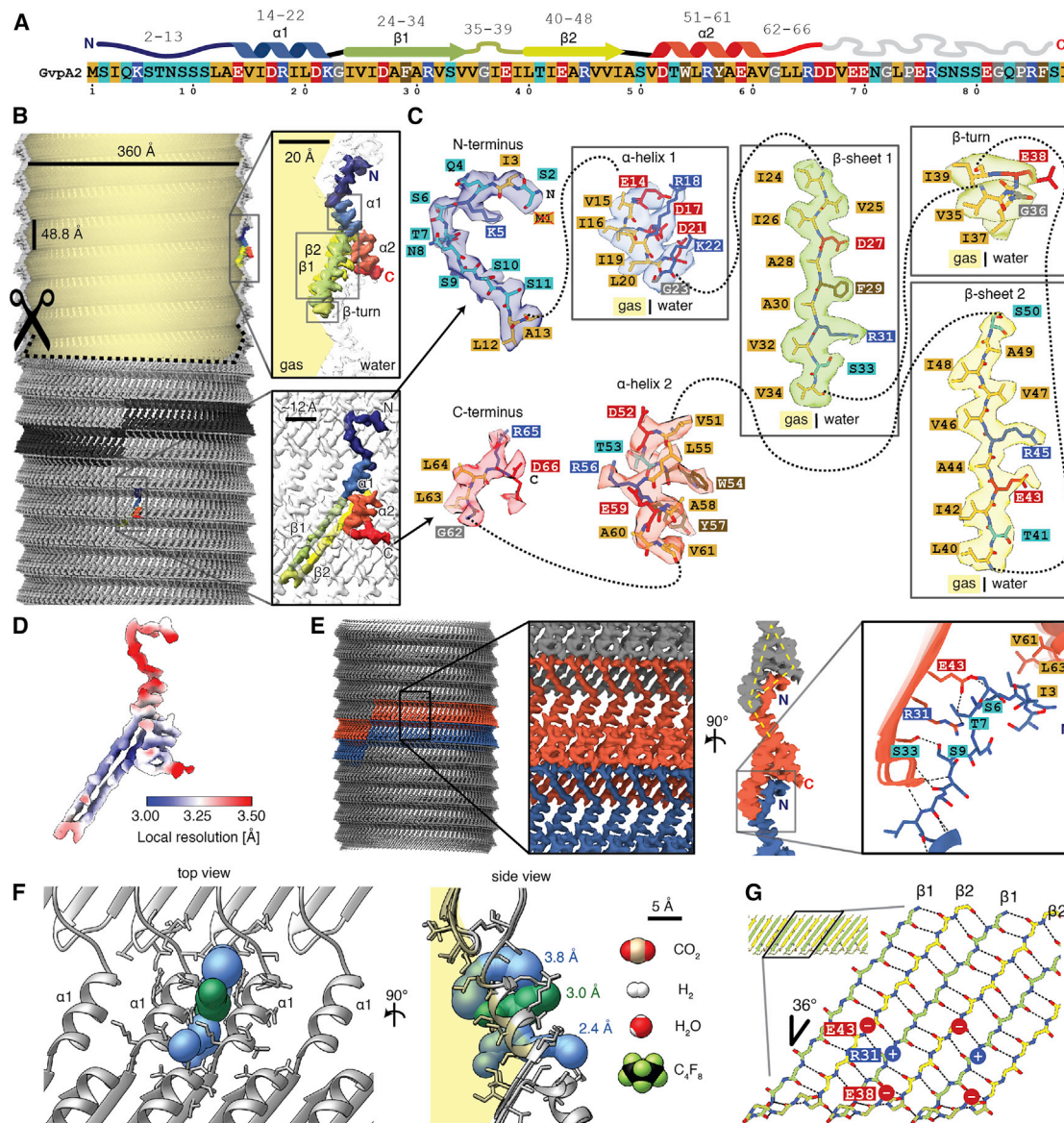
(Figure 1B), which had their cylinder shape best preserved in the thin liquid film of cryo-EM samples (Figures 1B–1F).

We used a combination of two- and three-dimensional (2D and 3D) classification techniques to select 4% of the small GV subset corresponding to 35.6 nm diameter (Figure S1) and used helical reconstruction to obtain a cryo-EM density of 3.2 Å resolution (Figure 2). The final reconstruction yielded a cylindrical GV shell assembly with ~93 GvpA2 monomers per helical turn, which represents one member of a range of helical polymorphs with diameters ranging from 34 to 70 nm and with 90–183 monomers per helical turn.

The reconstructed density allowed *de novo* atomic model building of GvpA2 in the structural context of its native assembly (Figure 2; Table S1). The cylindrical shell is constructed from thousands of GvpA2 monomers polymerized side by side into ribs spiraling into a left-handed helix with a helical pitch of 48.8 Å and  $-3.87^\circ$  helical twist, resulting in 92.93 GvpA2 units per helical turn (Figures 2A and 2B; Video S1). Contrary to postulated models,<sup>14,16</sup> GvpA monomers and not antiparallel dimers form the repeating units of the helical assembly. GvpA2 adopts

a coil- $\alpha$ - $\beta$ - $\beta$ - $\alpha$ -coil fold (Figures 2A and 2B). The carboxyl-terminal residues Asp<sub>67</sub>-Ile<sub>88</sub> are flexible and not resolved in our structure. The helical lattice forms an array of ribs consisting of densely packed GvpA2 subunits with a lateral center-to-center distance of ~12 Å. The central  $\beta$ -hairpin, tilted at  $-36^\circ$  relative to the long axis of the cylinder, forms the core of the GV ribs. Helix  $\alpha$ 2 folds back onto the hairpin, and helix  $\alpha$ 1 forms a bridge across the ~16 Å gap separating adjacent ribs. The GV wall is therefore only one or two peptide layers thick (Figure 2B). The inner wall of the GV shell forms a continuous hydrophobic surface consisting of a dense pattern of hydrophobic residues located on the luminal side of the  $\beta$ -hairpin and helix  $\alpha$ 1 (Figures 2C and 2D). Connections between the ribs of the GV shell are formed by the predominantly polar N terminus, which extends perpendicular to helix  $\alpha$ 1 and folds across the  $\beta$ -hairpin of the adjacent rib, stabilized by interactions with several residues in the  $\beta$ -hairpin and helix  $\alpha$ 2 (Figure 2E).

The extreme hydrophobicity of the luminal GV surface constitutes an energetic barrier for diffusion of liquid water or condensation of gaseous H<sub>2</sub>O. Consistently, GVs have been



**Figure 2. Cryo-EM structure of the gas vesicle wall**

(A) Primary and secondary structure of *B. megaterium* GvpA2. Residues in the primary structure are colored based on physicochemical properties.

(B) GvpA2 monomers form thin-walled gas-filled protein shells assembled into a left-handed helix. One individual rib formed by 93 monomers is highlighted in dark gray. The top part of the GV is cut open to visualize the gas space. The 3.2 Å resolution cryo-EM density is shown with a single monomer colored according to sequence (N terminus, blue; C terminus, red) from a side view (top inset) and front view (bottom inset).

(C) Entire cryo-EM density of a GvpA2 monomer annotated with respective amino acid one-letter-codes and the atomic model. Aliphatic residues (Ala, Val, Leu, Ile) line the gas-facing side of the GV wall.

(D) Estimate of local resolution mapped onto the monomer density.

(E) Cryo-EM density of the GV wall with two ribs highlighted in orange and blue. Side view illustrating corrugated zigzag structure and triangular cross-sections of the GV wall (yellow lines). Close up of inter-rib interactions mediated by the N terminus (blue), which binds across the  $\beta$ -hairpin and the C terminus (orange) of adjacent ribs, stabilized by hydrogen bonds from backbone, polar side chains (Ser<sub>6</sub>, Thr<sub>7</sub>, Ser<sub>9</sub>), and hydrophobic contacts (Ile<sub>3</sub>).

(F) A slit between  $\alpha$ 1 helices allows diffusion of gas through the wall. Three computed tunnels approximate the slit and have bottleneck sizes ranging from 2.4 to 3.8 Å. The van der Waals surfaces of several gas molecules known to diffuse through the wall are shown to scale, with perfluorocyclobutane being the largest with 6.3 Å collision diameter.<sup>7</sup>

(G) Schematic of the  $\beta$  strand rib providing most lateral connections for the assembly through backbone hydrogen bonding (dotted lines) and electrostatic interactions (Glu<sub>43</sub>-Arg<sub>31</sub>-Glu<sub>38</sub>).

See also [Figure S1](#), [Table S1](#), and [Video S1](#).

shown to be impermeable to water but to be highly permeable to gas molecules.<sup>7</sup> How gas molecules pass through the GV wall has so far been unclear. We located pores in the GV shell formed by slit-like openings between  $\alpha$ 1 helices of adjacent GvpA2 monomers (Figure 2F). We quantified the resulting pore size in the GV assembly computationally using Voronoi diagrams<sup>22</sup> and retrieved three different access routes with minimal constrictions ranging from 2.4 to 3.8 Å, compatible with the collisional cross-sections of gases dissolved in the cytosol (Figure 2F).<sup>23</sup>

Despite its limited thickness, the GV shell can resist several atmospheres of pressure without collapse.<sup>4</sup> GvpA2 monomers are held together tightly by lateral connections along the GV ribs formed by an extensive hydrogen-bonding network between the  $\beta$  strand backbones (Figure 2G). The hydrogen bonds are oriented at an angle of 54° relative to the cylinder axis, which is close to the "magic angle" 54.7° at which transverse and longitudinal stresses are equal in the wall of a cylinder.<sup>3,24</sup> Additional reinforcements are made by a continuous network of salt bridges formed by Glu<sub>43</sub>-Arg<sub>31</sub> between two monomers and Arg<sub>31</sub>-Glu<sub>38</sub> within a monomer. The GvpA2 shell consists of alternating line segments and triangular cross-sections providing force-bearing elements (Figure 2E). The corrugations increase stiffness along the rib direction, while the linear segments provide compliant hinge elements that increase elasticity of GVs, increasing their capacity to accommodate deformations orthogonal to the rib without collapse.<sup>25</sup>

### Gas vesicles consist of two half shells with inverted orientation

The thin film of the cryo-EM sample orients the large GVs into a sideways orientation, providing a consistent viewing direction onto the GV edges in projection. Detailed inspection of GV edges revealed that GvpA2 monomers are always oriented with their  $\beta$ -turns pointing toward the center of the GV cylinder (Figure 3A), which contains a structural irregularity that has previously been referred to as a seam.<sup>13</sup> 2D class averages of GV edges around this seam showed two oppositely oriented GvpA2 monomers that make contact via their  $\beta$ -turns (Figures 3A and S2A) implying that this is the contact site of two GV half shells with inverted orientation.

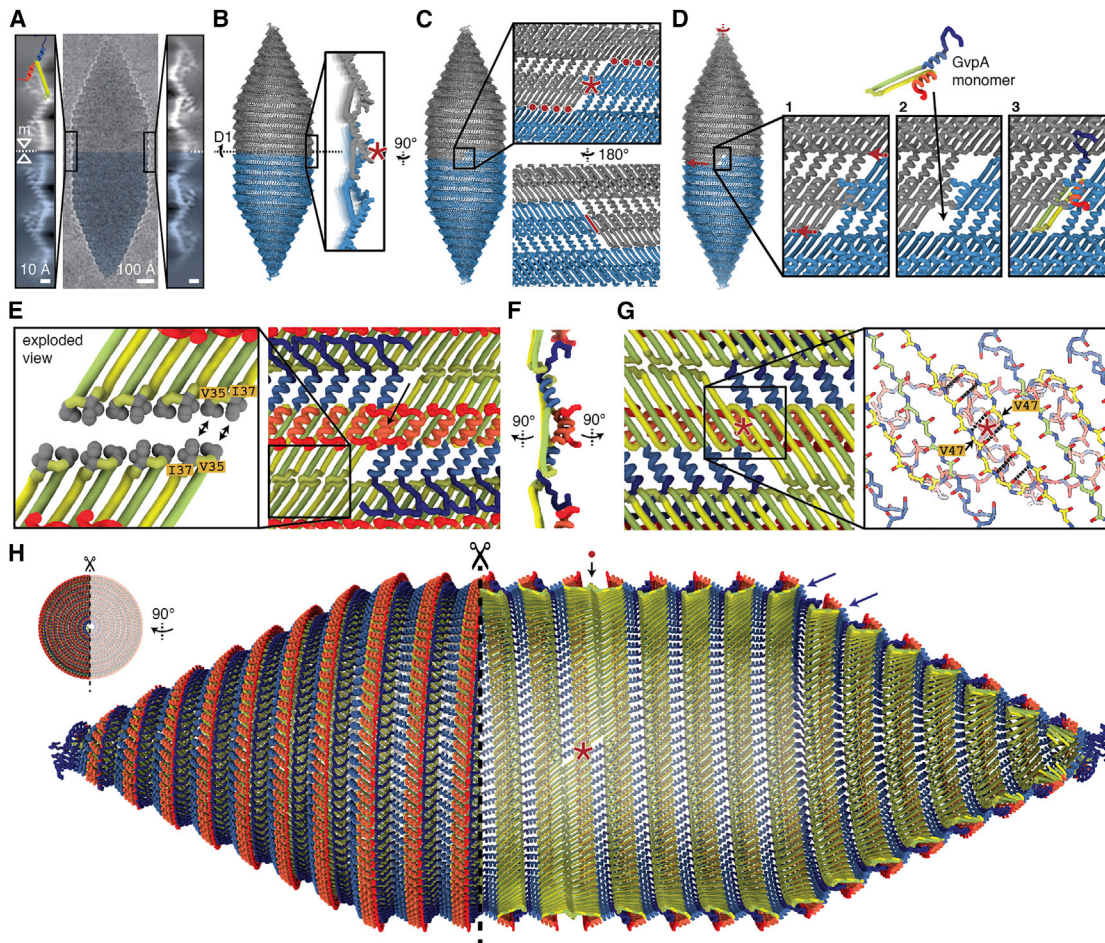
While two contacting cylinders or cones have the same contact geometry along the entire circumference, GVs are assembled from two contacting helices. Continuity of the helicity implies there must be a unique polarity reversal point (PRP) at a point along the circumference, where an upwards- and a downwards-oriented GvpA2 monomer meet side by side. Apart from classes showing contacting  $\beta$ -turns, there is one less frequent set of 2D classes of the seam that displays two overlapping GvpA2 monomers in inverted orientation (Figures 3A and S2B). We posit that this 2D class is a projection view of the PRP located at the GV edge. As GVs can freely rotate around the cylinder axis in the thin ice layer, the PRP will be located exactly at the edge only at special rotation angles, explaining why this class is observed less frequently. The mirror symmetry in the 2D class implies a viewing direction perpendicular to a 180° (C2/D1) rotation axis (Figure S2C) that points through the PRP.

Using restraints from our 3D reconstruction and 2D class averages, we constructed a pseudo-atomic model of a GV half shell. Starting at the PRP, we extended the model from the D1 symmetry axis using the known helical symmetry for the cylindrical part of the GV shell and allowed transitioning into the conical tips by linearly decreasing the radius set by the 25° semi-angle of the cone and while refining structural adaptation of GvpA monomers at defined hinge-points to match the experimental data (details in STAR Methods and Figure S3). We then duplicated the half shell by rotating around the D1 axis, leading to a complete GV model consisting of 1,730 monomers and a total molecular mass of 12.2 MDa. Simulated density projections from this model closely match the experimental 2D classes (Figures S2D and S2E).

### Molecular mechanism of gas vesicle biogenesis

Our pseudo-atomic model of the GV assembly shows that half shells interact through contacts at the GvpA  $\beta$ -turns around the circumference of the seam, as well as at the PRP where the GvpA2 rib reverses its polarity (Figure 3C). The pattern of side-by-side contacts between  $\beta$  sheets of GvpA2 monomers ( $d=d=d$ ) along a rib is identical everywhere except at the PRP, where  $\beta$ -hairpins align inversely ( $d=d-p=p$ ). The PRP is therefore a unique point in the GV assembly that may be recognized by the molecular machinery that facilitates GV growth and is likely the point at which new GvpA molecules are inserted during GV growth. Expansion of the GV assembly could occur either by inserting two oppositely oriented monomers, leading to symmetric expansion of both GV half shells, or by stochastic insertion of monomers into either half shell. If GV growth occurs by insertion of dimers, the PRP should always be located exactly at the midpoint of the cylindrical segment. Instead, we observe that the PRP can be located away from the midpoint of the helical GV segment (Figure 1B), consistent with previous observations.<sup>26</sup> We therefore propose that insertion of new monomers on either side of the PRP occurs stochastically through ratcheting of the two GV half shells by rotation relative to each other, generating a single monomer gap at the PRP (Figure 3D). This would involve breaking the lateral hydrogen bonds between the two monomers around the PRP, along with breaking and re-establishing hydrophobic contacts (Val<sub>35</sub>, Ile<sub>37</sub>) of the  $\beta$ -turns at the seam with no net loss in energy (Figure 3E). In our model, two factors suggest that the PRP represents the weakest point in the assembly. First, the monomer orientation at the PRP leads to steric hindrance by the  $\alpha$ 2 helices of the two symmetry-related monomers (Figure 3E). Second, the two oppositely oriented monomers at the PRP are connected by only 6 hydrogen bonds between segments of strand  $\beta$ 2 around Val<sub>47</sub> as compared to 11 hydrogen bonds formed between monomers along the rib (Figure 3G). We propose that the energetic disadvantage of this conformation facilitates opening of the seam for addition of new monomers during growth.

If the GV grows by adding monomers at the PRP, backward extrapolation starting from a mature GV leads to a state at which the seam forms between two conical half shells (Figure 4A). This is the point where conical growth transitions into cylindrical growth. This transition requires adaptation of the relative orientation of adjacent ribs, which we propose is mediated through a



**Figure 3. Pseudo-atomic model of a gas vesicle with two half shells in inverted orientation**

(A) Raw cryo-EM image of a single GV (with inverted contrast).  $\beta$ -hairpins of GvpA2 (cartoon) always point toward the seam at the center of the GV cylinder. Two different types of 2D class averages of the seam (left and right) are observed. The mirror symmetry (mirror axis: m) suggests a  $180^\circ$  symmetry axis (D1) at the point where two inversely oriented GV half shells meet.

(B) Pseudo-atomic model of a GV constructed from two identical halves (gray, blue) with close up showing a side view of the polarity reversal point (PRP, red asterisk).

(C) Close-up view onto the GvpA2 lattice around the PRP (red asterisk). Red dots indicate molecular contacts along the GV circumference where  $\beta$ -turns contact. The red line indicates contact between parts of hairpin strands  $\beta_2$  at the PRP.

(D) Model of monomer insertion at the PRP. The two GV halves rotate against each other, with the  $\beta$ -hairpin contacts sliding over each other (red arrow) in a ratcheting fashion to allow monomer insertion in the resulting gap. Insertion of the monomer in the orientation opposite to that depicted is geometrically equivalent and would enlarge the other GV half shell.

(E) View onto the seam between the two GV half shells with color scheme as in Figure 2B. The seam is sealed by contacting  $\beta$ -turns and the hydrophobic side chains Val35 and Ile37 (exploded view in inset). The lateral orientation of the hairpins is not uniquely determined by the data. An unresolved steric clash between  $\alpha_2$  helices from GvpA monomers around the PRP is marked by an arrow.

(F) Side view of the PRP with color scheme as in Figure 2B.

(G) View of the PRP from inside the GV. Oppositely oriented monomers next to the PRP are connected by six hydrogen bonds formed between the backbones of the  $\beta_2$  strands, with amino acid Val47 being located around the D1 symmetry axis (red asterisk).

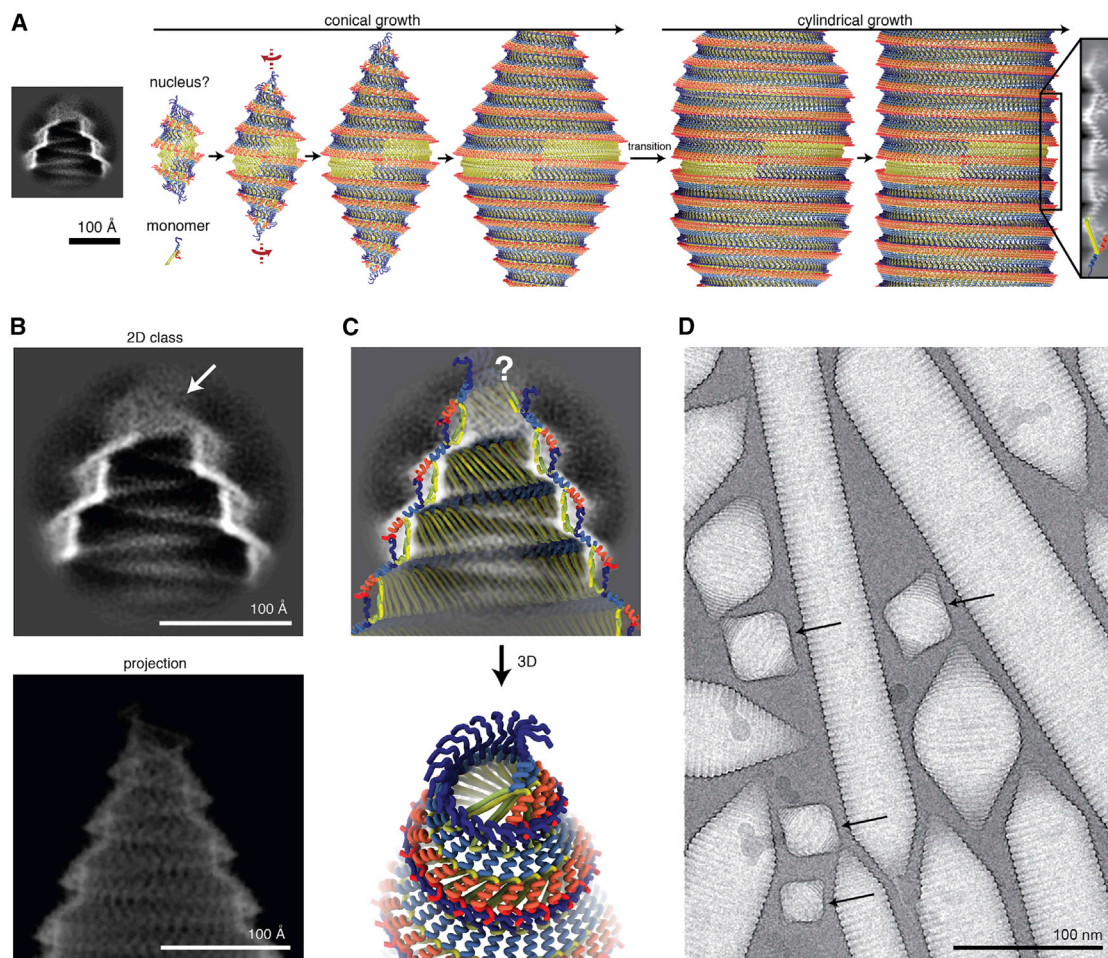
(H) Enlarged, sideways-rotated, and cut-open GV model with PRP (asterisk) and seam (dot) annotated. The proposed hinging motion of the N terminus required for adaption to the cylinder-to-cone transition and the reducing diameter of the tip is shown (blue arrows).

See also Figures S2 and S3.

hinging motion of the N terminus relative to helix  $\alpha_1$  to accommodate for the reduction in diameter at the transition point and along the cone toward the tip (Figure 3H).

Continuing this extrapolation leads to a biconical nucleus that must initially form to start GV biogenesis. According to our model, the original nucleus would remain at the cone tips

on both half shells after maturation (Video S2). The 2D classes and the fitted pseudo-atomic model suggest that N-terminal residues of GvpA2 monomers crowd together at the tip (Figures 4B and 4C). At diameters lower than  $50 \text{ \AA}$ , the 2D class averages of these cone tips display weak density (Figure S2F) sealing off the opening. The fuzzy appearance of the density



**Figure 4. Model of gas vesicle growth from an initial nucleus**

(A) GV growth from a hypothetical nucleus to mature GVs. The transition from conical to cylindrical growth is annotated. A 2D class of GV tips with proposed nucleus remnants is shown to scale.

(B) Class average of the GV tip shows no molecular order at diameters lower than  $50 \text{ \AA}$  toward the cone end (top). 2D projection of the pseudo-atomic model is shown to scale (bottom).

(C) 2D class overlaid with cut-through of the pseudo-atomic model (top) and 3D view onto the tip (bottom). The simplified model of a helix with linearly decreasing radius breaks down at the very tip, leading to steric clashes between the main chains. N termini from GvpA monomers at the tip come into proximity and might close off the gas space.

(D) Cryo-EM micrograph with immature GV bicones (arrows).

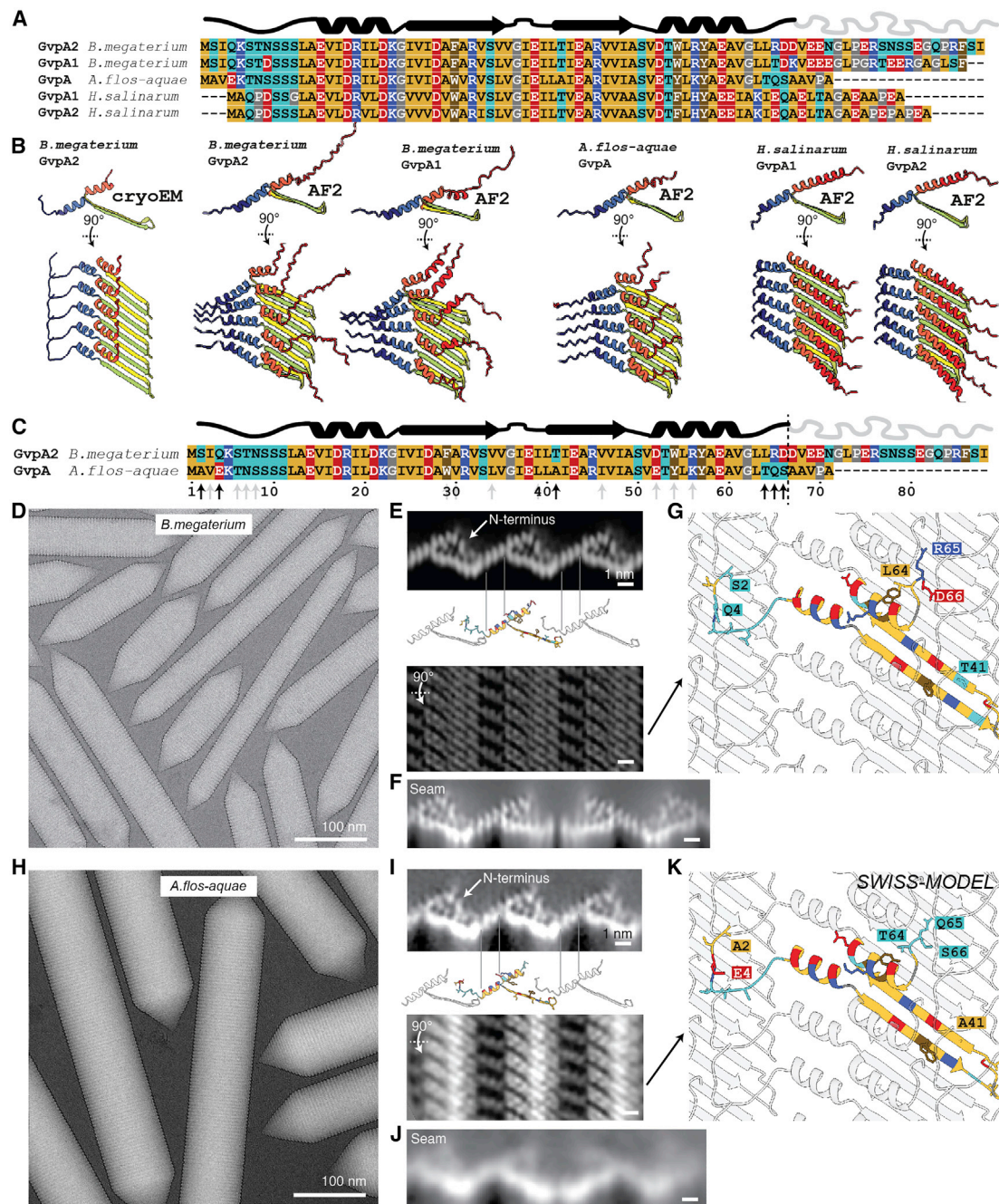
See also [Video S2](#).

at the tip is indicative of structural heterogeneity in the nucleating monomers.

### Conservation of gas vesicle shell architecture

Sequence alignments between GvpAs of three evolutionary diverse bacterial and archaeal species producing GVs reveal a high degree of sequence conservation in the structural parts of GvpA, suggesting that the overall mode of GvpA shell assembly must be similar. This is supported by computational predictions using AlphaFold2, revealing highly similar assemblies (ribs) of GvpA oligomers for the evolutionarily distant GvpAs in firmicutes (*B. megaterium*), cyanobacteria (*A. flos-aquae*) and the archaeon *H. salinarum*, that all resemble our experimental structure (Figures 5A and 5B).

To verify these predictions, we employed cryo-EM to image GVs of *A. flos-aquae*. Differences in the GvpA sequences between *A. flos-aquae* and *B. megaterium* GvpA2 locate mainly to the N-terminal and C-terminal regions of the folded core (Figure 5C). *A. flos-aquae* GVs formed cone-ended cylinders with mean diameter of  $87 \pm 7 \text{ nm}$ , consistent with previous observations.<sup>18</sup> Despite significant effort, 3D refinement using a similar helical reconstruction strategy as for *B. megaterium* did not converge. Closer inspection of full-length GVs revealed them to flatten in the thin ice of the cryo-EM sample, breaking symmetry assumptions of helical reconstruction. We therefore resorted to obtain high-resolution structural information from 2D classification. Class averages of GV edges showed a corrugated zigzag pattern of the  $\sim 2 \text{ nm}$ -thick wall formed by GvpA (Figure S4A).



**Figure 5. Highly conserved GvpA from *B. megaterium* and *A. flos-aquae* adopt the same fold and assembly**

(A) Sequence alignment of *B. megaterium* (Mega) GvpA1 (GvpA) and GvpA2 (GvpB), *A. flos-aquae* (Ana) GvpA, and *H. salinarum* (Halo) GvpA1 and GvpA2 show high degree of conservation despite forming GVs of different diameters.

(B) AF2 prediction of GvpA 5-mers for selected genes in (A) compared to the cryo-EM structure. Only the middle monomer is shown in side view. AF2 predicts the general arrangement of a GvpA rib and the angle between the  $\beta$ -hairpin and  $\alpha$  helix 1 accurately. The conformation of the N-terminal coil appears different, and the distance between  $\alpha$  helix 1 (where gas pores are located) is not accurately modeled by AF2.

(C) Protein primary sequences of wall-forming protein GvpA from both *B. megaterium* and *A. flos-aquae* are very similar. Black arrows show 6 property-changing mutations; gray arrows show 12 property-conserving mutations in the ordered part of the structure.

(D) Representative cryo-EM micrograph of *B. megaterium* GVs.

(E) 2D-projected side view and top view of the 3.2 Å cryo-EM density of *B. megaterium* GVs.

(F) 2D class average of seam between two GV half shells.

(legend continued on next page)

Another set of 2D class averages, obtained from collapsed GVs also present in the data, revealed the *A. flos-aquae* GV wall to consist of a periodic array of ribs consisting of dense  $5.0 \times 1.25$  nm GvpA subunits tilted at  $-36^\circ$  relative to the long axis of the cylinder (Figure S4B). The class averages with a resolution of better than 4.8 Å allow discerning the  $\alpha$ -helical repeat of helix  $\alpha_1$  bridging the gap of adjacent ribs and the individual  $\beta$  strands of the polymerizing  $\beta$ -hairpin. A cumulative Fourier spectrum computed from in-plane rotated GV segments showed a typical Fourier transform of a helix (Figure S4C), consistent with our 3D reconstruction from *B. megaterium* GVs (Figure S4D). We compared the *A. flos-aquae* 2D classes of GV wall side and top views with equivalent projections of the *B. megaterium* GV structure. On the level of the main-chain fold, the two GV shell assemblies are indistinguishable, hence confirming the conserved architecture of GvpA assembly of the GV shell (Figure 5). Our model of two contacting half shells also applies to *A. flos-aquae* GVs as we observe similar contacting  $\beta$ -turns in 2D class averages of the seam (Figures 5F and 5J).

More generally, our results establish that 2D classification of GV edges in cryo-EM data can give valuable structural insight into GV architecture for cases when 3D reconstruction is out of reach. This is due to the small unit cell of the GV wall, where not many structural features overlap in side views at the GV edges. 2D class averages can be computed to sufficient resolution to see secondary structure elements and even large side chains. We suggest that this approach can be used for comparative studies of GVs from different species with larger sequence divergence to reveal different assembly modes, such as resulting from differences in binding modes of the evolutionarily less conserved GvpA N terminus (Figure S5A).

### Reinforcement of the gas vesicle shell by GvpC

Many GV gene clusters contain a second structural GV protein GvpC, which is absent from our model GV from *B. megaterium*.<sup>28</sup> GvpC binds on the outside of GVs and increases the critical collapse pressure of GVs.<sup>9,29</sup> While the essential role of GvpC has been firmly established, how it binds to and reinforces GVs has remained elusive. We acquired cryo-EM datasets of *A. flos-aquae* GVs, and to investigate the structural role of GvpC, we imaged the GVs in the presence and absence of GvpC and compared 2D class averages of the GV edges (Figure 6A).

The class average containing GvpC shows an additional dot-like density ( $\sim 10$  Å diameter) located in vicinity to helix  $\alpha_2$  (Figure 6A). This feature and its dimensions are consistent with the projection of an  $\alpha$  helix viewed along its helical axis. Indeed, GvpC is predicted to be all  $\alpha$ -helical<sup>10</sup> and consists of five 33-residue repeats (Figures 6B and 6C) that are highly similar in sequence. The class averages suggest that GvpC binds along the ribs of GVs and contacts helix  $\alpha_2$  of GvpA (Figure 6B), which

is supported by a molecular envelope of the *A. flos-aquae* GV wall determined by electron tomography.<sup>30</sup> We analyzed the evolutionary conservation of the repeat sequence. In a set of 91 GvpC sequences (Figure S5B) from different organisms, we find nine residues to be conserved in more than 90% of those sequences, including a strongly conserved set of leucine (Leu<sub>4,12,30</sub>), phenylalanine (Phe<sub>11,33</sub>), and arginine (Arg<sub>19</sub>) residues (Figure 6D). A helical wheel plot of the repeat shows that all conserved residues cluster on the same face (Figure 6E) that likely forms the binding interface. We further confirmed the importance of these residues by designing alanine mutants for occurrences in all five repeats (Figure S6A). All GvpC mutants abrogated stabilization of GVs (Figure 6F).

Our 2D class averages provide strong spatial restraints on the positioning of GvpC relative to the GvpA ribs, while the conservation pattern and mutants pinpoint residues essential for binding. We used these data as restraints to predict a model for GvpC binding by computational docking (Figure 6G and S6E–S6H).

While our data do not allow us to decisively distinguish whether GvpC follows the left-handed spiral of the rib upwards toward the tips, or downwards, we obtained the highest-scoring docking solution with the downward orientation (Figure 6H). A single GvpC repeat spans four monomers of GvpA. In this GvpA tetrad, glutamate residues Glu<sub>52</sub> and Glu<sub>59</sub> in GvpA monomers 1, 2, and 4 form hydrogen bonds with GvpC. In our model, Glu<sub>52</sub> in monomer 3 binds to the conserved Arg<sub>19</sub> of GvpC, while Glu<sub>59</sub> in monomer 4 binds to the conserved Gln<sub>26</sub>. The (Leu<sub>4</sub>, Leu<sub>12</sub>, Leu<sub>30</sub>) triplet inserts between the  $\alpha_2$  helices of the GvpA tetrad, while Phe<sub>11</sub> and Phe<sub>33</sub> are sandwiched in between  $\alpha_2$  in GvpA monomers 2 and 3 and face-on on helix  $\alpha_2$  of monomer 4.

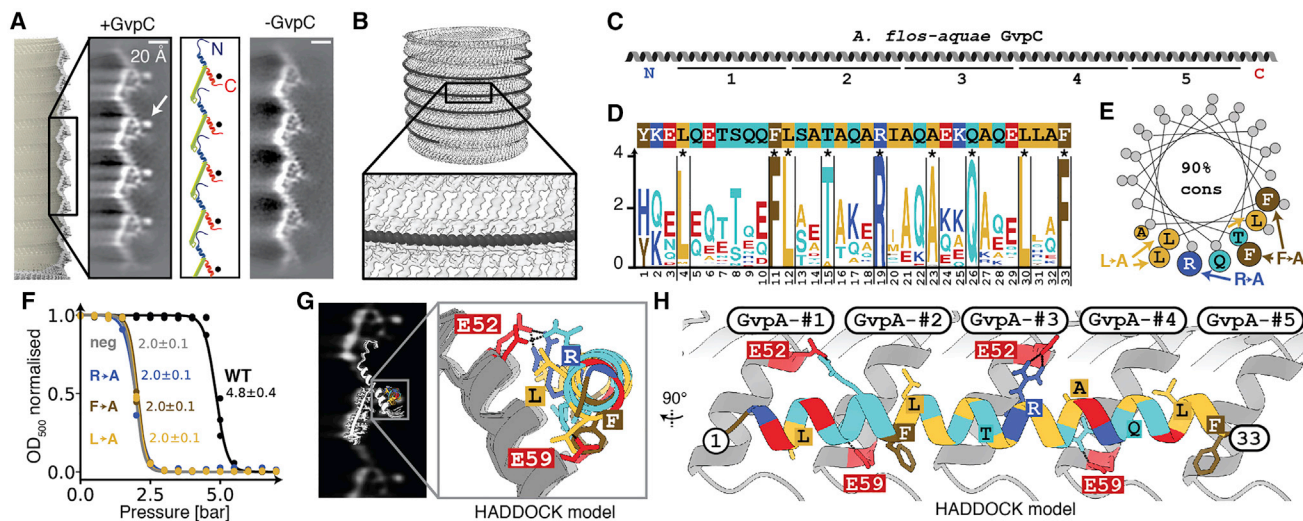
### DISCUSSION

GVs represent a remarkable example of biomolecular self-assembly. Our results provide a canonical structural framework for the unique molecular properties of GVs, including their selective permeability to gases,<sup>7,31</sup> their mechanical stability,<sup>25</sup> and their distinctive ability to grow without compromising the integrity of its shell.<sup>3</sup> Our work establishes an atomic resolution model of the mature GV shell formed exclusively by GvpA. A key question is how GvpA nucleates to form an elementary bicone from which the shell extends during GV growth. Besides GvpA, many GV gene clusters contain genes encoding the proteins GvpJ, GvpM, or GvpS<sup>28</sup> that exhibit high sequence homology and predicted folds similar to that of GvpA. A dominant structural role of these homologs in mature GVs is unlikely, as none of them were found in intact GVs,<sup>29</sup> suggesting a putative role as nucleation factors. Additional support for this role comes from the observation that, in some species, GvpJ and GvpM are expressed as part of a separate transcript exclusively during early exponential growth,<sup>8</sup> whereas a transcript of gvpACNO is

(G) Atomic model of *B. megaterium* GV wall with different side chains between both species displayed and property-changing mutations highlighted by one-letter-code. Residues are colored according to side chain chemistry as indicated in (A).

(H–K) Equivalent data for *A. flos-aquae* GVs shows the high degree of conservation of GvpA fold and assembly. The 2D views in (I) are computed by 2D classification of GV edges and collapsed GVs and have a resolution better than 5.4 Å or 4.8 Å as the  $\alpha$ -helical pitch or the  $\beta$  strands are resolved. The homology model of the *A. flos-aquae* GV wall was computed using SWISS-MODEL.<sup>27</sup>

See also Figure S4.



**Figure 6. The secondary wall protein GvpC binds along the GvpA ribs of *A. flos-aquae* GVs**

(A) Comparison of 2D class averages of GV edges with (left) and without (right) GvpC reveals an additional circular density (arrow). A cartoon model of GvpA helps locating the GvpC density to helix  $\alpha 2$  (red).  
 (B) Artist impression of GvpC molecules wrapping around GVs.  
 (C) Predicted secondary structure and 33-amino-acid repeats 1–5 of *A. flos-aquae* GvpC.  
 (D) Consensus sequence of the GvpC repeats with logo representation of evolutionary conservation reveals nine highly conserved residues. The height of the characters depicts the degree of conservation (information content in bits).  
 (E) Helical wheel plot of highly conserved (>90% conserved in 91 GvpC sequences from other species) amino acids reveal that there is one highly conserved face of the  $\alpha$ -helical repeat. Experimentally tested GvpC mutants are indicated.  
 (F) Critical collapse pressure measurements of *A. flos-aquae* GVs supplemented with wild-type GvpC or GvpC mutants or stripped of GvpC.  
 (G) Comparison of 2D class average of GvpC-bound GVs and predicted binding mode between a GvpC consensus repeat and seven *A. flos-aquae* GvpA monomers.  
 (H) Rotated view of binding model with predicted interactions of residues.  
 See also Figures S5 and S6.

expressed at later stages, possibly to enlarge the formed nuclei. How the initial nucleus is structured and by which mechanism other Gvps assist in growth remain open questions.<sup>32</sup>

In most organisms, biconical GVs transition their growth mode when reaching a certain diameter and continue extending cylindrically. This transition occurs over a range of diameters, different for each individual GV. We observed mature cylindrical *B. megaterium* GVs with a diameter of  $55.5 \pm 7.3$  nm corresponding to  $145 \pm 19$  monomers per helical turn and mature *A. flos-aquae* GVs with diameter of  $87.1 \pm 6.9$  nm corresponding to  $227 \pm 18$  monomers per helical turn. The mechanism for the bi-cone-to-cylinder transition is unknown. The GvpA ribs in the cones are highly curved and may be energetically disfavored. Insertion of new monomers in a bicone results in enlargement of the cone and reduces rib curvature. However, the expanding cone does not provide defined interactions between monomers of adjacent ribs. In contrast, cylindrical GV segments have crystalline order where the N terminus always has identical contacts to the adjacent rib. The interplay of curvature preference and the energetic advantage of crystallinity could favor a cone-to-cylinder transition at a certain critical diameter. This suggests that the GvpA N terminus could play a role in determining the size distribution of GVs. Consistently, the sequence of the N terminus is most divergent. Further evidence for the decisive role of GvpA sequence in defining the final diameter comes from a hybrid

GV gene cluster where *A. flos-aquae* GvpA integrated into the native gene cluster of *B. megaterium*, resulting in GVs with diameters consistent with native *A. flos-aquae* GVs.<sup>33</sup>

Our structure reveals that the gas permeability of the GV wall can be ascribed to a large number of molecular pores formed between  $\alpha 1$  helices of the GvpA shell, the size of which is compatible with the collision diameters (2.65–3.64 Å) of typical atmospheric gases dissolved in the cytosol.<sup>23</sup> Surprisingly, the GV wall has been shown to be permeable to perfluorocyclobutane ( $C_4F_8$ ) with a collision diameter (6.3 Å) exceeding the estimated pore size in our structure. While atmospheric gases appear to diffuse freely through the GV wall,<sup>31</sup> the diffusion coefficient of  $C_4F_8$  is consistent with a very small number (~11) of such pores.<sup>7</sup> Based on our pseudo-atomic model, such pores would need to be located at the seam, PRP or at the conical tips. Alternatively, flexibility in the GV shell may modulate size of the regular gas pores to allow passage of  $C_4F_8$ .

We inferred a binding mode of a single 33-amino-acid repeat of GvpC to four consecutive GvpA monomers assembled in the cylindrical *A. flos-aquae* GV shell. The requirement for a binding site comprised of consecutive GvpA monomers in a helical assembly is consistent with previous observations in split-GFP assays showing that monomeric GvpA and GvpC do not interact.<sup>32</sup> What remains open is how the five consecutive GvpC repeats are structured. It is likely that all repeats would

form identical molecular contacts with the GvpA ribs. The side-by-side distance of GvpA monomers measured between identical points of the inner  $\beta$ -hairpins is 12.1 Å. Due to the curvature of GVs, this distance increases radially outwards across the GV wall toward the binding site for GvpC. For an average 87 nm *A. flos-aquae* GV, this distance is  $\sim$ 12.7 Å. A single perfectly  $\alpha$ -helical repeat of GvpC would span 49.5 Å ( $33 \times 1.5$  Å/residue), only slightly shorter than the 50.8 Å spanned by a stretch of four GvpA monomers. The five tandem repeats of native GvpC would bind to 20 GvpA monomers. Taking into account additional space for the N and C termini of GvpC, this corresponds well to a previously determined 1:25 M ratio of GvpA to GvpC.<sup>10</sup> The modest distance mismatch between one GvpC repeat and a GvpA tetrad could be accommodated by slight deviation of GvpC from perfect helicity. It has been previously suggested that this would also be required to maintain the relative orientation of consecutive repeats, as with perfect helicity they would be  $(100^\circ \times 33) \% 360 = 60^\circ$  rotated toward each other.<sup>10</sup> A small unstructured stretch would allow GvpC to adapt to different curvatures of the GvpA ribs in cylindrical and conical parts and in GVs of different diameters.

We show that GvpC binds to helix  $\alpha 2$  of the GV wall, thereby increasing the critical collapse pressure. The wall thickness measured between the extremes of the corrugation pattern increases from  $\sim$ 3 nm to  $\sim$ 4 nm when GvpC is bound. Because the buckling pressure of thin-walled cylinders is proportional to the third power of the wall thickness,<sup>25</sup> the most straightforward explanation for enhanced stability is that GvpC bound on the outside of GVs increases the wall thickness. The real picture is likely more complicated. Our structure reveals that GVs are assembled from two helical half shells that are connected by a seam. Because of its structure, it is conceivable that the seam is a weak point of the GV where failure will occur first. In this case, GvpC would be most critical to prevent failure at the seam, which questions its requirement elsewhere on the GV shell.

Recently, GVs have been repurposed as genetically encoded acoustic reporters.<sup>34,35</sup> The high contrast in density between gas-filled GVs and surrounding cellular structures makes them amenable to ultrasound imaging.<sup>36</sup> While native GVs display little shell deformation under ultrasound exposure, GVs that are stripped of GvpC become less stiff and scatter non-linearly above a certain pressure threshold.<sup>37</sup> This behavior enables amplitude modulation imaging and multiplexing of stripped and unstripped GVs in an *in vivo* context.<sup>37,38</sup> Engineering conditional binding strength of GvpC can transform GVs into biosensors with switchable acoustic properties.<sup>39,40</sup> Our insights into GvpA-GvpC interaction provide a rational basis for designing such sensors. Moreover, the high-resolution structure of the GvpA shell may enable development of designer GVs with custom mechanical shell properties through direct engineering of the GvpA sequence.

Together, our results establish the molecular basis of a widely conserved buoyancy-controlled motility apparatus in aquatic bacteria and archaea. Our study will form the foundation for addressing a multitude of open questions on GV biogenesis such as nucleation, growth, width regulation, function of other GV gene products in GV assembly, and species-to-species variability of GV gene clusters.

### Limitations of the study

While our structure provides an atomic model of the main structural protein GvpA as it forms the GV wall assembly and our data allowed inferring the binding geometry of GvpC to the GV wall, the function and/or structural role of other accessory proteins in GV gene clusters in GV biogenesis is yet to be determined. Some of those proteins presumably work together to create a nucleus whose precise structure remains unknown.

We also report a pseudo-atomic model of an entire GV. The precise structure of the GvpA monomer in this model is supported by cryo-EM data at 3.2 Å resolution; however, the special structural features of the assembly such as the seam, polarity reversal point, and tip are inferred from 2D projection views alone. The exact molecular details at these special positions can slightly deviate from the pseudo-atomic model we present.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - *Escherichia coli* BL21(DE3)pLysS
  - *Anabena flos-aquae* CCAP 1403/13F
- METHOD DETAILS
  - *B. megaterium* gas vesicle expression and purification
  - *A. flos-aquae* gas vesicle purification
  - GvpC mutant purification
  - Preparation of gas vesicles with recombinant GvpC and collapse pressure measurements
  - Cryo-EM of *B. megaterium* gas vesicles
  - Cryo-EM of *A. flos-aquae* gas vesicles
  - Data processing and structure determination of *B. megaterium* gas vesicles
  - Atomic model building and refinement
  - 2D classification of edges, seams and tips
  - Pore analysis
  - Pseudo-atomic modeling of a complete gas vesicle
  - AlphaFold2 structure predictions
  - Bioinformatics analysis
  - HADDOCK modeling of GvpC binding to *A. flos-aquae* gas vesicles
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Gas vesicle width measurements
  - Collapse pressure measurements

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2023.01.041>.

## ACKNOWLEDGMENTS

We thank Cecilia de Agrela Pinto and Tarannum Ara for help with cloning and reconstitution of GvpC constructs and Willem Noteborn for data collection support at NeCEN. We thank Sander Tans and Cees Dekker for critically reading the manuscript and helpful suggestions. We acknowledge support by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. ERC-StG-852880 to A.J.J.), the Dutch Research Council (NWO.STU.018-2.007 to A.J.J.; STU.019.021 to D.M.), the 4TU Precision Medicine Program, the Medical Delta Ultra HB program, and the Kavli Institute of Nanoscience Delft. The data on *B. megaterium* GVs were collected at the Netherlands Centre for Electron Nanoscopy (NeCEN) made possible through financial support from the Dutch Roadmap Grant NEMI (NWO.GWI.184.034.014 to A.J.J.).

## AUTHOR CONTRIBUTIONS

S.T.H., D.T., D.M., and A.J.J. conceived the project. D.T. and D.M. provided initial samples. S.T.H. and D.T. purified gas vesicles. S.T.H. prepared cryo-EM samples. S.T.H. and W.H.E. optimized cryo-EM data collection for GVs. S.T.H. processed cryo-EM data. S.T.H. and A.J.J. analyzed cryo-EM data, interpreted the structure, and built the atomic model and the mature GV assembly. S.T.H. performed bioinformatic analyses and docking experiments. S.T.H. and A.J.J. designed GvpC mutants. S.T.H. expressed and purified GvpC mutants. D.T. performed collapse experiments. All authors discussed the results. S.T.H. and A.J.J. wrote the manuscript. All authors commented on the final version. A.J.J. supervised the overall project and contributions by S.T.H. and W.H.E. D.M. supervised contributions by D.T.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: July 25, 2022

Revised: November 14, 2022

Accepted: January 30, 2023

Published: March 2, 2023

## REFERENCES

1. Szurmant, H., and Ordal, G.W. (2004). Diversity in Chemotaxis Mechanisms among the Bacteria and Archaea. *Microbiol. Mol. Biol. Rev.* **68**, 301–319. <https://doi.org/10.1128/mmb.68.2.301-319.2004>.
2. Jarrell, K.F., and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. *Nat. Rev. Microbiol.* **6**, 466–476. <https://doi.org/10.1038/nrmicro1900>.
3. Walsby, A.E. (1994). Gas vesicles. *Microbiol. Rev.* **58**, 94–144. <https://doi.org/10.1128/mmb.58.1.94-144.1994>.
4. Hayes, P.K., and Walsby, A.E. (1986). The inverse correlation between width and strength of gas vesicles in cyanobacteria. *Br. Phycol. J.* **21**, 191–197. <https://doi.org/10.1080/00071618600650221>.
5. Blaurock, A.E., and Walsby, A.E. (1976). Crystalline structure of the gas vesicle wall from *Anabaena flos-aquae*. *J. Mol. Biol.* **105**, 183–199. [https://doi.org/10.1016/0022-2836\(76\)90106-6](https://doi.org/10.1016/0022-2836(76)90106-6).
6. Offner, S., Ziese, U., Wanner, G., Typke, D., and Pfeifer, F. (1998). Structural characteristics of halobacterial gas vesicles. *Microbiology* **144**, 1331–1342. <https://doi.org/10.1099/00221287-144-5-1331>.
7. Walsby, A.E. (1982). Permeability of gas vesicles to perfluorocyclobutane. *J. Gen. Microbiol.* **128**, 1679–1684. <https://doi.org/10.1099/00221287-128-8-1679>.
8. Pfeifer, F. (2012). Distribution, formation and regulation of gas vesicles. *Nat. Rev. Microbiol.* **10**, 705–715. <https://doi.org/10.1038/nrmicro2834>.
9. Walsby, A.E., and Hayes, P.K. (1988). The Minor Cyanobacterial Gas Vesicle Protein, GVPc, Is Attached to the Outer Surface of the Gas Vesicle. *Microbiology* **134**, 2647–2657. <https://doi.org/10.1099/00221287-134-10-2647>.
10. Buchholz, B.E.E., Hayes, P.K., and Walsby, A.E. (1993). The distribution of the outer gas vesicle protein, GvpC, on the *Anabaena* gas vesicle, and its ratio to GvpA. *J. Gen. Microbiol.* **139**, 2353–2363. <https://doi.org/10.1099/00221287-139-10-2353>.
11. McMaster, T.J., Miles, M.J., and Walsby, A.E. (1996). Direct observation of protein secondary structure in gas vesicles by atomic force microscopy. *Biophys. J.* **70**, 2432–2436. [https://doi.org/10.1016/S0006-3495\(96\)79813-2](https://doi.org/10.1016/S0006-3495(96)79813-2).
12. Sivertsen, A.C., Bayro, M.J., Belenky, M., Griffin, R.G., and Herzfeld, J. (2009). Solid-state NMR evidence for inequivalent GvpA subunits in gas vesicles. *J. Mol. Biol.* **387**, 1032–1039. <https://doi.org/10.1016/j.jmb.2009.02.015>.
13. Sivertsen, A.C., Bayro, M.J., Belenky, M., Griffin, R.G., and Herzfeld, J. (2010). Solid-state NMR characterization of gas vesicle structure. *Biophys. J.* **99**, 1932–1939. <https://doi.org/10.1016/j.bpj.2010.06.041>.
14. Strunk, T., Hamacher, K., Hoffgaard, F., Engelhardt, H., Zillig, M.D., Faist, K., Wenzel, W., and Pfeifer, F. (2011). Structural model of the gas vesicle protein GvpA and analysis of GvpA mutants in vivo. *Mol. Microbiol.* **81**, 56–68. <https://doi.org/10.1111/j.1365-2958.2011.07669.x>.
15. Bayro, M.J., Daviso, E., Belenky, M., Griffin, R.G., and Herzfeld, J. (2012). An amyloid organelle, solid-state NMR evidence for cross- $\beta$  assembly of gas vesicles. *J. Biol. Chem.* **287**, 3479–3484. <https://doi.org/10.1074/jbc.M111.313049>.
16. Ezzeldin, H.M., Klauda, J.B., and Soares, S.D. (2012). Modeling of the major gas vesicle protein, GvpA: from protein sequence to vesicle wall structure. *J. Struct. Biol.* **179**, 18–28. <https://doi.org/10.1016/j.jsb.2012.04.015>.
17. Knitsch, R., Schneefeld, M., Weitzel, K., and Pfeifer, F. (2017). Mutations in the major gas vesicle protein GvpA and impacts on gas vesicle formation in *Haloflex volcanii*. *Mol. Microbiol.* **106**, 530–542. <https://doi.org/10.1111/mmi.13833>.
18. Dutka, P., Malounda, D., Metskas, L.A., Chen, S., Hurt, R.C., Lu, G.J., Jensen, G.J., and Shapiro, M.G. (2021). Measuring gas vesicle dimensions by electron microscopy. *Protein Sci.* **30**, 1081–1086. <https://doi.org/10.1002/pro.4056>.
19. Li, N., and Cannon, M.C. (1998). Gas vesicle genes identified in *Bacillus megaterium* and functional expression in *Escherichia coli*. *J. Bacteriol.* **180**, 2450–2458. <https://doi.org/10.1128/jb.180.9.2450-2458.1998>.
20. Lakshmanan, A., Lu, G.J., Farhadi, A., Nety, S.P., Kunth, M., Lee-Gosselin, A., Maresca, D., Bourdeau, R.W., Yin, M., Yan, J., et al. (2017). Preparation of biogenic gas vesicle nanostructures for use as contrast agents for ultrasound and MRI. *Nat. Protoc.* **12**, 2050–2080. <https://doi.org/10.1038/nprot.2017.081>.
21. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296. <https://doi.org/10.1038/nmeth.4169>.
22. Sehnal, D., Svobodová Vařeková, R., Berka, K., Pravda, L., Navrátilová, V., Banáš, P., Ionescu, C.-M., Otyepka, M., and Koča, J. (2013). MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J. Cheminformatics* **5**, 39. <https://doi.org/10.1186/1758-2946-5-39>.
23. Ismail, A.F., Chandra Khulbe, K., and Matsuura, T. (2015). Gas Separation Membranes: Polymeric and Inorganic, First Edition (Springer International Publishing) <https://doi.org/10.1007/978-3-319-01095-3>.
24. Horgán, C.O., Murphy, J.G., and Saccomandi, G. (2018). The complex mechanical response of anisotropic materials in simple experiments. *Int. J. Non-Linear Mech.* **106**, 274–279. <https://doi.org/10.1016/j.ijnonlinmec.2018.05.025>.

25. Walsby, A.E. (1982). The elastic compressibility of gas vesicles. *Proc. R. Soc. Lond. B Biol. Sci.* *216*, 355–368. <https://doi.org/10.1098/rspb.1982.0079>.
26. Walsby, A.E. (1972). Structure and function of gas vacuoles. *Bacteriol. Rev.* *36*, 1–32. <https://doi.org/10.1128/membr.36.1.1-32.1972>.
27. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., Beer, T.A.P.D., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* *46*, 296–303. <https://doi.org/10.1093/nar/gky427>.
28. Tashiro, Y., Monson, R.E., Ramsay, J.P., and Salmond, G.P.C. (2016). Molecular genetic and physical analysis of gas vesicles in buoyant enterobacteria. *Environ. Microbiol.* *18*, 1264–1276. <https://doi.org/10.1111/1462-2920.13203>.
29. Dunton, P.G., Mawby, W.J., Shaw, V.A., and Walsby, A.E. (2006). Analysis of tryptic digests indicates regions of GvpC that bind to gas vesicles of *Anabaena flos-aquae*. *Microbiology* *152*, 1661–1669. <https://doi.org/10.1099/mic.0.28755-0>.
30. Dutka, P., Metskas, L.A., Hurt, R.C., Salahshoor, H., Wang, T.-Y., Malounda, D., Lu, G., Chou, T.-F., Shapiro, M.G., and Jensen, G.J. (2022). Structure of *Anabaena flos-aquae* gas vesicles revealed by cryo-ET. *bioRxiv*. Preprint at. <https://doi.org/10.1101/2022.06.21.496981>.
31. Walsby, A.E. (1969). The permeability of blue-green algal gas-vacuole membranes to gas. *Proc. R. Soc. Lond. B Biol. Sci.* *173*, 235–255. <https://doi.org/10.1098/rspb.1969.0049>.
32. Völkner, K., Jost, A., and Pfeifer, F. (2020). Accessory Gvp Proteins Form a Complex During Gas Vesicle Formation of Haloarchaea. *Front. Microbiol.* *11*, 610179. <https://doi.org/10.3389/fmicb.2020.610179>.
33. Bourdeau, R.W., Lee-Gosselin, A., Lakshmanan, A., Farhadi, A., Kumar, S.R., Nety, S.P., and Shapiro, M.G. (2018). Acoustic reporter genes for noninvasive imaging of microorganisms in mammalian hosts. *Nature* *553*, 86–90. <https://doi.org/10.1038/nature25021>.
34. Shapiro, M.G., Goodwill, P.W., Neogy, A., Yin, M., Foster, F.S., Schaffer, D.V., and Conolly, S.M. (2014). Biogenic gas nanostructures as ultrasonic molecular reporters. *Nat. Nanotechnol.* *9*, 311–316. <https://doi.org/10.1038/nnano.2014.32>.
35. Maresca, D., Lakshmanan, A., Abedi, M., Bar-Zion, A., Farhadi, A., Lu, G.J., Szablowski, J.O., Wu, D., Yoo, S., and Shapiro, M.G. (2018). Biomolecular ultrasound and sonogenetics. *Annu. Rev. Chem. Biomol. Eng.* *9*, 229–252. <https://doi.org/10.1146/annurev-chembioeng-060817-084034>.
36. Heiles, B., Terwiel, D., and Maresca, D. (2021). The Advent of Biomolecular Ultrasound Imaging. *Neuroscience* *474*, 122–133. <https://doi.org/10.1016/j.neuroscience.2021.03.011>.
37. Maresca, D., Lakshmanan, A., Lee-Gosselin, A., Melis, J.M., Ni, Y.-L., Bourdeau, R.W., Kochmann, D.M., and Shapiro, M.G. (2017). Nonlinear ultrasound imaging of nanoscale acoustic biomolecules. *Appl. Phys. Lett.* *110*, 073704. <https://doi.org/10.1063/1.4976105>.
38. Maresca, D., Sawyer, D.P., Renaud, G., Lee-Gosselin, A., and Shapiro, M.G. (2018). Nonlinear X-wave ultrasound imaging of acoustic biomolecules. *Phys. Rev. X* *8*, 041002. <https://doi.org/10.1103/PhysRevX.8.041002>.
39. Lakshmanan, A., Farhadi, A., Nety, S.P., Lee-Gosselin, A., Bourdeau, R.W., Maresca, D., and Shapiro, M.G. (2016). Molecular Engineering of Acoustic Protein Nanostructures. *ACS Nano* *10*, 7314–7322. <https://doi.org/10.1021/acsnano.6b03364>.
40. Lakshmanan, A., Jin, Z., Nety, S.P., Sawyer, D.P., Lee-Gosselin, A., Malounda, D., Swift, M.B., Maresca, D., and Shapiro, M.G. (2020). Acoustic biosensors for ultrasound imaging of enzyme activity. *Nat. Chem. Biol.* *16*, 988–996. <https://doi.org/10.1038/s41589-020-0591-0>.
41. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* *7*, e42166. <https://doi.org/10.7554/eLife.42166>.
42. Zhang, K. (2016). Gctf : Real-time CTF determination and correction. *J. Struct. Biol.* *193*, 1–12. <https://doi.org/10.1016/j.jsb.2015.11.003>.
43. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* *27*, 14–25. <https://doi.org/10.1002/pro.3235>.
44. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* *66*, 486–501. <https://doi.org/10.1107/S0907444910007493>.
45. Croll, T.I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D. Struct. Biol.* *74*, 519–530. <https://doi.org/10.1107/S2059798318002425>.
46. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D. Struct. Biol.* *75*, 861–877. <https://doi.org/10.1107/S2059798319011471>.
47. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* *16*, 1153–1160. <https://doi.org/10.1038/s41592-019-0575-8>.
48. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera — A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* *25*, 1605–1612. <https://doi.org/10.1002/jcc.20084>.
49. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* *596*, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
50. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* *19*, 679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
51. Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* *16*, 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2).
52. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* *44*, W344–W350. <https://doi.org/10.1093/nar/gkw408>.
53. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* *39*, W29–W37. <https://doi.org/10.1093/nar/gkr367>.
54. Brown, N.P., Leroy, C., and Sander, C. (1998). MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* *14*, 380–381. <https://doi.org/10.1093/bioinformatics/14.4.380>.
55. Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res.* *14*, 1188–1190. <https://doi.org/10.1101/gr.849004>.
56. van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kaspritts, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* *428*, 720–725. <https://doi.org/10.1016/j.jmb.2015.09.014>.
57. Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* *41*, 207–234. <https://doi.org/10.1016/j.pep.2005.01.016>.
58. Gautier, R., Douguet, D., Antonny, B., and Drin, G. (2008). HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics* *24*, 2101–2102. <https://doi.org/10.1093/bioinformatics/btn392>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
BL21 (DE3) pLysS	EMBL protein expression core facility	AJLC0003
BL21 (DE3)	EMBL protein expression core facility	AJLC0002
<i>A.flos-aquae</i>	CCAP (Culture Collection of Algae and Protozoa)	CCAP 1403/13F
<b>Chemicals, peptides, and recombinant proteins</b>		
SoluLyse	genlantis	L100125
(Isopropyl β-D-1-thiogalactopyranoside) IPTG	VWR	CAS 367-93-1
Lysozyme	Roth	CAS 12650-88-3
DNAseI	ITW reagents	CAS 9003-98-9
BG11	Sigma	73,816
<b>Deposited data</b>		
Atomic coordinates of <i>Bacillus megaterium</i> GvpA2	This study	PDB: 7r1c
Cryo-EM structure of <i>Bacillus megaterium</i> gas vesicles (helically symmetric)	This study	EMDB: 14340
Cryo-EM structure of <i>Bacillus megaterium</i> gas vesicles (after local refinement)	This study	EMDB: 14238
Pseudo-atomic model of a complete <i>Bacillus megaterium</i> gas vesicle	This study	<a href="https://doi.org/10.5281/zenodo.6458345">https://doi.org/10.5281/zenodo.6458345</a>
2D classes of gas vesicle seams from cryo-EM micrographs of <i>Bacillus megaterium</i> and <i>Anabaena flos-aquae</i> gas vesicles	This study	<a href="https://doi.org/10.5281/zenodo.6867443">https://doi.org/10.5281/zenodo.6867443</a>
AlphaFold2 models of <i>Anabaena flos-aquae</i> and <i>Halobacterium salinarum</i> GvpA	This study	<a href="https://doi.org/10.5281/zenodo.6867443">https://doi.org/10.5281/zenodo.6867443</a>
Atomic models from molecular docking solutions for <i>Anabaena flos-aquae</i> GvpC bound to GvpA	This study	<a href="https://doi.org/10.5281/zenodo.6867443">https://doi.org/10.5281/zenodo.6867443</a>
Surface models from gas pore analysis in <i>Bacillus megaterium</i> gas vesicle	This study	<a href="https://doi.org/10.5281/zenodo.6867443">https://doi.org/10.5281/zenodo.6867443</a>
<b>Recombinant DNA</b>		
pET28a-GvpC_His6	Genscript, this study	AJLD0162
pET28a-GvpC_His6 (R19A)	Genscript, this study	AJLD0163
pET28a-GvpC_His6 (F11A,F33A)	Genscript, this study	AJLD0164
pET28a-GvpC_His6 (L4A,L12A,L30A)	Genscript, this study	AJLD0166
pST39-pNL29 (Lakshmanan et al. <sup>20</sup> )	Addgene	#91696
<b>Software and algorithms</b>		
MOLE 2.5	Sehna et al. <sup>22</sup>	<a href="https://mole.upol.cz/">https://mole.upol.cz/</a>
SWISS-MODEL	Waterhouse et al. <sup>27</sup>	<a href="https://swissmodel.expasy.org/">https://swissmodel.expasy.org/</a>
RELION 3.1	Zivanov et al. <sup>41</sup>	<a href="https://github.com/3dem/relion">https://github.com/3dem/relion</a>
Gctf 1.06	Zhang <sup>42</sup>	<a href="https://www.mrc-lmb.cam.ac.uk/kzhang/">https://www.mrc-lmb.cam.ac.uk/kzhang/</a>
cryoSPARC 3.1 and 3.3	Punjani et al. <sup>21</sup>	<a href="https://cryosparc.com/download">https://cryosparc.com/download</a>
ChimeraX 1.4	Goddard et al. <sup>43</sup>	<a href="https://www.cgl.ucsf.edu/chimerax/download.html">https://www.cgl.ucsf.edu/chimerax/download.html</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
WinCoot 0.9.8.1	Emsley et al. <sup>44</sup>	<a href="https://bernhardcl.github.io/cool/wincoot-download.html">https://bernhardcl.github.io/cool/wincoot-download.html</a>
Isolde 1.4	Croll <sup>45</sup>	<a href="https://isolde.cimr.cam.ac.uk/download/">https://isolde.cimr.cam.ac.uk/download/</a>
Phenix 1.13	Liebschner et al. <sup>46</sup>	<a href="https://phenix-online.org/download">https://phenix-online.org/download</a>
TOPAZ v.0.23	Bepler et al. <sup>47</sup>	<a href="https://github.com/tbepler/topaz">https://github.com/tbepler/topaz</a>
Chimera 1.13.1	Pettersen et al. <sup>48</sup>	<a href="https://www.cgl.ucsf.edu/chimera/download.html">https://www.cgl.ucsf.edu/chimera/download.html</a>
AlphaFold2	Jumper et al. <sup>49</sup>	<a href="https://github.com/deepmind/alphafold">https://github.com/deepmind/alphafold</a>
ColabFold	Mirdita et al. <sup>50</sup>	<a href="https://github.com/sokrypton/ColabFold">https://github.com/sokrypton/ColabFold</a>
EMBOSS	Rice et al. <sup>51</sup>	<a href="https://emboss.sourceforge.net/download/">https://emboss.sourceforge.net/download/</a>
ConSurf web server	Ashkenazy et al. <sup>52</sup>	<a href="https://consurf.tau.ac.il/">https://consurf.tau.ac.il/</a>
HMMer v.3.3.2 web server	Finn et al. <sup>53</sup>	<a href="https://www.ebi.ac.uk/Tools/hmmer/">https://www.ebi.ac.uk/Tools/hmmer/</a>
MView	Brown et al. <sup>54</sup>	<a href="https://desmid.github.io/mview/">https://desmid.github.io/mview/</a>
WebLogo	Crooks et al. <sup>55</sup>	<a href="https://weblogo.berkeley.edu/logo.cgi">https://weblogo.berkeley.edu/logo.cgi</a>
HADDOCK 2.4	van Zundert et al. <sup>56</sup>	<a href="https://wenmr.science.uu.nl/haddock2.4/">https://wenmr.science.uu.nl/haddock2.4/</a>
<b>Other</b>		
R2/1 300 mesh holey carbon grids	Quantifoil	N/A

**RESOURCE AVAILABILITY**

**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Arjen J. Jakobi ([a.jakobi@tudelft.nl](mailto:a.jakobi@tudelft.nl)).

**Materials availability**

Plasmids for GvpC mutants are freely available on request.

**Data and code availability**

- The refined atomic model of *B. megaterium* GvpA2 (GvpB) has been deposited in the PDB under accession code PDB: 7R1C. The cryo-EM density of a subsection of the *B. megaterium* wall (after local refinement) is available in the Electron Microscopy DataBank under accession code EMD: EMD-14238. A symmetrized density of an entire GV cylinder is available under accession code EMD-14340. The presented pseudo-atomic model of a complete *B. megaterium* GV is available on Zenodo (<https://doi.org/10.5281/zenodo.6458345>). 2D class averages of *A. flos-aquae* and *B. megaterium* GVs, predicted gas pores, AlphaFold2 structure predictions and computational docking results of *A. flos-aquae* GvpA with GvpC are also available on Zenodo (<https://doi.org/10.5281/zenodo.6867443>).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

***Escherichia coli* BL21(DE3)pLysS**

The *E. coli* BL21 host strain expressing T7 RNA polymerase and T7 lysozyme was used for heterologous production of *B. megaterium* GVs. *E. coli* BL21(DE3)pLysS cells were grown in lysogeny broth (LB) containing 0.2% (w/v) glucose at 37°C, or at 30°C after induction of heterologous expression by IPTG.

***Anabena flos-aquae* CCAP 1403/13F**

*A. flos-aquae* (CCAP 1403/13F), also known as *Dolichospermum flos-aquae*, were grown in G625 medium complemented by BG11 medium (Sigma C3061) for approximately 2 weeks until confluence.

## METHOD DETAILS

### *B. megaterium* gas vesicle expression and purification

The purification protocol for Mega GVs was derived from Lakshmanan et al.<sup>20</sup> In brief, BL21(DE3)pLysS *E. coli* cells were transformed with the pST39-pNL29 plasmid (a gift from Mikhail Shapiro; Addgene #91696)<sup>20</sup> and 1 L of LB containing 0.2% (w/v) glucose was inoculated with 10 mL of overnight culture. The culture was grown at 37°C until OD = 0.5 and GV expression was induced with 20 μM IPTG. Following induction, cells were grown at 30°C for 20 h.

The culture was centrifuged at 500 rcf for 2 h in 50 mL Falcon tubes. The floating fraction was collected with a peristaltic pump. This process was repeated once more. The resulting 25 mL of liquid were lysed chemically by adding 5 mL of SoluLyse reagent per 50 mL of liquid, 0.25 mg/mL lysozyme and 10 μg/mL DNaseI, and slowly rotated for 1 h at room temperature.

GVs were purified in three overnight rounds of floatation separation by centrifugation at 800 rcf in 50 mL Falcon tubes. After each centrifugation step, the GV-containing top layer was gently removed with a pipette after which the GVs were resuspended in PBS containing 6M urea (first round), and subsequently in PBS alone. Final concentration was determined as OD<sub>500</sub> = 3.1 by optical density measurement at 500 nm against a sonicated blank. The sample was dialyzed into imaging buffer (20 mM Tris, 50 mM NaCl, pH = 8) prior to cryo-EM sample preparation.

### *A. flos-aquae* gas vesicle purification

The purification protocol for GVs from *A. flos-aquae* was derived from Lakshmanan et al.<sup>20</sup> Briefly, *A. flos-aquae* (CCAP 1403/13F), also known as *D. flos-aquae*, were grown in 250 mL G625 medium complemented by BG11 medium (Sigma C3061) for approximately 2 weeks until confluence. The culture was centrifuged at 350 rcf for 4 h or until a floating layer of cells was observed. Subnatant was removed using a syringe before lysing the cells in 500 mM sorbitol and 10% v/v lysis buffer (SoluLyse) at 4°C for 6–8 h while gently rotating. GVs were purified by three rounds of floatation separation with 4–6 h centrifugation at 350 rcf. After each centrifugation, subnatant was removed by syringe after which GVs were resuspended in PBS at pH 7.4.

### GvpC mutant purification

Codon-optimized genes for *A. flos-aquae* wild-type and mutant (R19A; F-11,33-A; L-4,12,30-A) GvpC (Uniprot: P09413) including a C-terminal 'GSGSGS' linker and a C-terminal 6xHis-tag in a pET-28a(+) vector were obtained from GenScript (New Jersey, United States). The mutations were engineered into all five repeats.

Proteins were expressed in *E. coli* BL21-DE3 cells grown in 750 mL autoinduction medium<sup>57</sup> for 3 h at 37°C before the temperature was lowered to 20°C for additional 20 h. The bacteria were harvested by centrifugation and lysed by freeze-thaw cycles in lysis buffer (20 mM Tris, 500 mM NaCl, 0.1% Triton X-, 20 mM imidazole; 5 mL per gram of pellet). Lysozyme (0.15 mg/mL) and DNaseI (10 μg/mL) were added and the lysate rotated for 2 h at room temperature. Isolation of inclusion body and IMAC purification was performed as described previously.<sup>20</sup> Protein purity was assessed by SDS-PAGE and concentration was determined according to Bradford.

### Preparation of gas vesicles with recombinant GvpC and collapse pressure measurements

*A. flos-aquae* GVs were stripped of GvpC by resuspending in 6 M urea and 60 mM Tris buffer, using 3 rounds of floatation separation as previously described.<sup>20</sup> Recombinant GvpC was added to stripped GVs according to the formulation:  $2 \times \text{OD}_{500} \times 198 \text{ nM} \times \text{GV volume (L)} = \text{nmol GvpC}$ . GvpC will be present in a 2-fold excess under the assumption of a 1:25 M ratio of GvpC/GvPA.<sup>20</sup> The urea solution was then slowly replaced with PBS at pH 7.4 by 2 rounds of ~12 h of dialysis over a 7–10 kDa MWCO dialysis membrane. Finally, 3 rounds of floatation separation at 350 rcf removed traces of urea. GVs were diluted to OD<sub>500</sub> = 0.1–0.4 for collapse pressure measurements. Samples were loaded in a pressure vessel and hydrostatic pressure was increased in increments of 0.5 bar using pressurized nitrogen gas. Samples were allowed to equilibrate for 5 s after pressure changes before absorption was measured using a spectrophotometer (Ocean optics) at 500 nm OD<sub>500</sub> values were normalized between the minimum and maximum for each measurement. Three independent re-additions per GvpC mutant were performed and measured.

A sigmoid function with  $p_0$  as the inflection point and  $k$  as the width was fitted to the curves using the means and standard deviations of measured triplicates ( $n = 3$ ) as input for the scipy 'curve\_fit' function.

$$\text{OD}_{500,\text{norm}} = \frac{1}{1 + e^{k(p - p_0)}}$$

The error of  $p_0$  was determined using a bootstrapping approach, performing the fit 50 times with a random set of  $n$  out of  $n$  measurement points with replacement. The parameters and their uncertainty were estimated as their mean and SD over the 50 fits.

### Cryo-EM of *B. megaterium* gas vesicles

*B. megaterium* GVs at OD<sub>500</sub> = 3.1 were applied to a freshly glow-discharged Quantifoil R2/1 grid and frozen using a Leica plunger set to 95% humidity, front-side blotting and 20°C with blot times ranging from 5 to 11 s. Micrographs were collected on a Titan Krios

(Thermo Fisher Scientific) microscope at the Netherlands Center for Electron Nanoscopy (NeCEN) operated at 300 kV. Dose-fractionated movies were acquired on a Gatan K3 direct electron detector at a pixel size of 1.37 Å with 60 frames over an exposure of 30 e<sup>-</sup>/Å<sup>2</sup> and a defocus range from -0.25 to -1.25 μm.

### Cryo-EM of *A.flos-aquae* gas vesicles

Native *A.flos-aquae* GVs (containing GvpC) at OD<sub>500</sub> = 13 in imaging buffer (20 mM Tris, 50 mM NaCl, 0.5 mM EDTA, pH = 8) were applied to a freshly glow-discharged Quantifoil R2/1 grid and frozen using a Leica plunger set to backside-blotting, 95% humidity and 20°C with 10 s blot time. 1273 cryo-EM micrographs at 1.288 Å pixel size were acquired on a JEOL 3200 microscope with a K2 detector using 62 e<sup>-</sup> total exposure over 60 frames. *A.flos-aquae* GVs stripped from GvpC (OD<sub>500</sub> = 1, in PBS buffer) were prepared and imaged in a similar same way, using 17.6 e<sup>-</sup> over 50 frames, 1.288 Å pixel size, acquiring 58 micrographs.

### Data processing and structure determination of *B.megaterium* gas vesicles

4351 collected super-resolution movies were 2x binned and motion-corrected in RELION 3.1.<sup>41</sup> CTF determination was performed using Gctf 1.06.<sup>42</sup> 709 micrographs containing thin GVs with a diameter of 42 nm or less were identified manually. 1021 tubes were manually picked in RELION by selecting start and end coordinates. 36,295 overlapping segments with 512 pixels were extracted along the cylindrical sections with a step size of 49 Å (2x binned). 2D classification was done in RELION3.1 with the 'ignore CTFs until first peak' option turned on. The resulting 2D class averages were grouped by projecting them along the helical axis and calculating the rim-to-rim distance (Figure S2B) between the two density maxima. A class with 35.6 nm edge-to-edge distance was selected containing 2911 segments.

Analysis of in-plane rotated power spectra of the segments using Helixplorer (<http://rico.ibs.fr/helixplorer/>) revealed a likely helical symmetry between 90.92 and 95.92 units per helical rung, with ~49 Å helical pitch. 2D classification was not sufficient to separate all symmetries and the final set of segments originated from several assemblies with different symmetries.

3D classification starting from a featureless cylinder and imposing candidate symmetry parameters was used to further select for segments adhering to a particular symmetry, leading to 1460 segments with symmetry of 92.93 units per turn. 3D refinement of those particles with CTF refinement and Bayesian polishing led to a resolution of 3.6 Å at FSC = 0.143. Convergence of 3D refinements was only achieved when using a 'tau\_fudge' parameter of 5.

A final round of particle polishing was used to create a new particle stack extracted only from frames 1–20 (0–10 e<sup>-</sup>/Å<sup>2</sup>) of the movies. Particles were exported to cryoSPARC 3.1.0<sup>21</sup> and high-pass filtered to 100 Å. 3D refinement with the helical reconstruction algorithm implemented in cryoSPARC also led to a reconstruction at 3.6 Å. To account for small deviations from helical symmetry, e.g. by flattening of the tube in ice, a mask encompassing ~3x9 GvpA2 monomers was created in ChimeraX.<sup>43</sup> The particle stack expanded by helical symmetry was subjected to focused refinement in cryoSPARC using the mask, which increased the final resolution to 3.2 Å at FSC = 0.143. The final maps were cropped from a box size of 512<sup>3</sup> voxels to a box size of 128<sup>3</sup> voxels centered on the refined region. Local resolution was calculated in cryoSPARC 3.1.0<sup>21</sup> and mapped onto the monomer structure in ChimeraX.<sup>43</sup>

### Atomic model building and refinement

To build an atomic model of a GvpA2 monomer the final map density was traced *de novo* using COOT.<sup>44</sup> The monomer was expanded using helical symmetry (rise: 0.525 Å, twist: -3.87°) into a 15 subunit segment (three ribs with 5 monomers each) to account for connections between monomers and manually adjusted in ISOLDE<sup>45</sup> before automatic real-space refinement in PHENIX 1.13.<sup>46</sup> We used phenix\_real\_space\_refine with automatic restraint weighting to refine coordinates and atomic displacement factors against the globally sharpened experimental density map using secondary structure and non-crystallographic symmetry restraints. The central monomer of the 15 subunit segment was used as the asymmetric unit for PDB deposition. Renderings of the cryo-EM density and atomic models were made in ChimeraX 1.4.<sup>43</sup>

### 2D classification of edges, seams and tips

From the *B.megaterium* GVs cryo-EM dataset, several hundred particles of either seams or tips were picked manually and used to generate a template for automated picking in cryoSPARC 3.3.<sup>21</sup> Particles were high-pass filtered to 100 Å to eliminate the large negative contrast of the gas space in the GV interior. The picked particles were cleaned up by several rounds of 2D classification to give a clean set of particles of either the seam, the PRP or the tips. These particle sets were used to train a neural network for particle picking (TOPAZ v0.23<sup>47</sup>), which was then applied to the micrographs to pick seams, PRPs, and tips. Those particles were cleaned by several rounds of 2D classification and led to the final presented 2D classes. For display, the 2D classes were treated in ImageJ using an 'unsharp mask' filter.

The two cryo-EM datasets of *A.flos-aquae* GVs with and without GvpC were used to obtain 2D class images of the edges leading to side-views of the wall. Movies were imported into cryoSPARC v3.3,<sup>21</sup> motion-corrected and CTF-estimated. For both dataset, frames were used only until an exposure of ~15 e<sup>-</sup>/Å<sup>2</sup> because shrinking of GVs was observed for high exposures leading to GV edges blurring out. A few hundred edges were manually selected for 2D classification to generate picking references for the cryoSPARC filament tracer. Particles were extracted with 192 pixel box size and high-pass filtered to 100 Å. Several rounds of 2D classification and selection of sharp classes led to the final 2D classes of the GV edges with and without GvpC. For display, the 2D classes were treated in ImageJ using an 'unsharp mask' filter.

Similarly, 2D classes of the GvpA lattice from collapsed GVs were calculated from the *A.flos-aquae* GV dataset above including GvpC. For this, the dataset was preprocessed in RELION 3.1,<sup>41</sup> points on the lattice manually picked to create a 2D class, which was then used for automated particle picking. Particles were extracted with a box size of 128 pixels and aligned with 2D classification to generate a view onto the collapsed GV wall. The biggest class containing ~34,000 particles was selected and was treated in ImageJ using an 'unsharp mask' filter.

### Pore analysis

Gas pores in the GV wall were analyzed using MOLE 2.5.<sup>22</sup> The gap between  $\alpha 1$  helices has a slit-type shape, enabling multiple possible routes for gas diffusion. Several start and endpoints of tunnels around both sides of the slit were selected and tunnels were computed with MOLE. The slit was modeled by three tunnels and displayed in Chimera.<sup>48</sup> The minimal constriction of the tunnel was calculated as the diameter of the smallest sphere in the respective tunnel model.

### Pseudo-atomic modeling of a complete gas vesicle

The model was generated from the solved atomic model of *B.megaterium* GvpA2. The GvpA2 monomer was placed next to the x axis (D1 axis) manually such that a 180°, symmetry copy operation around the x axis would reproduce a side view of the GV edge compatible with the determined 2D class average.

A left-handed parametric helix in 3D space was defined with the parameter  $t$  corresponding to turns of the helix:

$$t_{cap} = \frac{r_{max}}{P \cdot \sin(\alpha)}$$

$$r(t) = \begin{cases} r_{max}, & \text{if } t < t_{cyl} \\ r_{max} \cdot \left(1 - \frac{t - t_{cyl}}{t_{cap}}\right), & \text{if } t \geq t_{cyl} \end{cases}$$

$$x(t) = r(t) \cdot \cos(2\pi t)$$

$$y(t) = -r(t) \cdot \sin(2\pi t)$$

$$z(t) = \begin{cases} P \cdot t, & \text{if } t < t_{cyl} \\ P \cdot t_{cyl} + P \cdot \cos(\alpha) \cdot (t - t_{cyl}), & \text{if } t \geq t_{cyl} \end{cases}$$

where  $t_{caps}$  is the number of turns in the cap,  $t_{cyl}$  is a user-parameter of how many cylindrical turns the model should have (5),  $P$  is the pitch of the helix of 48.8 Å,  $r_{max}$  the radius of the assembly of 178.4 Å, and  $\alpha$  is the cone angle of the tip (25°).

The starting point of the curve was adjusted to go through the pivot point (in the center of the two  $\beta$ -sheets between amino acid (AA) 28 and AA 42 carbonyl oxygen atom) of the placed monomer by applying a shift along the z axis and rotation around the z axis. Points were placed along the curve at a distance of 12.07 Å, calculated as  $\sqrt{[(2\pi r)^2 + P^2]}/ut$ , where  $ut$  is the number of monomers per turn defined by the solved helical symmetry (92.93). 835 points were placed with the last 4 points toward the tip being omitted.

A model only with placement of monomers does not reflect the experimental 2D classes well. Four additional parameters were introduced rotating and modifying the monomer.

- (1) a correction for the change of helix angle toward the ends of the tip as the helix becomes steeper with narrower radius. The rotation center is between the AA28 and AA42 carbonyl oxygen atoms, and the rotation axis normal to a fitted plane through C- $\alpha$  atoms of all amino acids of the  $\beta$ -hairpin (AA23-49).
- (2) a rotation of the entire monomer to account for the tilt of monomers following the cone angle. The rotation center is carbonyl oxygen 36 and the axis normal to the plane of AA24-33 C- $\beta$  atoms.
- (3) a hinging motion of the two  $\beta$ -strands to account for the deformation visible in 2D classes of the PRP where GvpA ribs from both GVs halves meet. The rotation axis is defined by a line between the AA23 C- $\alpha$  atom and the AA49 C- $\alpha$  atom, and moves the atoms of AA23-49.
- (4) a hinging motion of the  $\alpha$ -helix 1 and the N-terminus, to account for gaps in the assembly formed when monomers tilt toward the cone angle. The axis is defined as in (3), and the rotation moves the atoms of AA2-23.

For the purpose of illustrating GV growth from nuclei to mature cylinders, the model was replicated in Blender 3.0 in a simplified form (without modifications of the monomer) and animated using varying parameters for the diameter and the  $t_{cyl}$  parameter.

### AlphaFold2 structure predictions

Complexes of five monomers of GvpA corresponding to a rib section of the GV wall were predicted using AlphaFold2<sup>49</sup> run in a Google Colab environment<sup>50</sup> with each five generated models per run and six recycles. The highest-scoring model was selected and displayed in ChimeraX.<sup>43</sup>

### Bioinformatics analysis

The sequences of the five 33 AA repeats of *A.flos-aquae* GvpC (Uniprot P09413) was converted into a consensus sequence 'YKELQETSQQFLSATAQARIAQAEKQAQELLAF' using the software 'cons' from the EMBOSS package.<sup>51</sup> The sequence was used as input for the ConSurf Server<sup>52</sup> to search the Uniref. 90 database with the HMMER web server<sup>53</sup> using one iteration, resulting in 91 sequences. The resulting sequence alignment was displayed and consensus sequences calculated in MView.<sup>54</sup> A sequence logo was calculated from the sequence alignment using WebLogo.<sup>55</sup> Helical wheel plots were generated in Heliquest.<sup>58</sup>

### HADDOCK modeling of GvpC binding to *A.flos-aquae* gas vesicles

A  $\alpha$ -helical model of the 33 residue consensus repeat of *A.flos-aquae* GvpC ('YKELQETSQQFLSATAQARIAQAEKQAQELLAF') was generated in ChimeraX<sup>43</sup> using ideal  $\alpha$ -helical backbone dihedral angles  $\phi = -57^\circ$  and  $\psi = -47^\circ$ . A homology model of the *A.flos-aquae* GvpA monomer was generated using SWISS-MODEL<sup>27</sup> based on the structure of *B.megaterium* structure GvpA2 and the *A.flos-aquae* sequence for GvpA (UniProt: P10397). The homology model was extended with the symmetry parameters from the *B.megaterium* assembly in ChimeraX<sup>43</sup> into a rib of 7 adjacent monomers.

Both models were used as input for HADDOCK 2.4.<sup>56</sup> For GvpC, residues 4,11,12,15,19,23,26,30,33, which all are >90% conserved in the bioinformatics analysis were chosen as active residues. For GvpA, residues 51–61 of GvpA, part of helix  $\alpha$ 2 and adjacent to the GvpC density in the 2D classes, were chosen as active residues. All remaining settings were used at default values.

The highest scoring cluster of the docking solutions (HADDOCK score:  $-108.4 \pm 5.0$ ) showed GvpC binding across several GvpA monomers along the helical spiral of the rib, with the GvpC sequence oriented inversely to the direction of helical propagation (with the direction from seam to tip). A very similar cluster was found shifted by one GvpA monomer laterally with the same molecular contacts and was discarded. A second, lower scoring type of cluster (HADDOCK score:  $-90.7 \pm 4.6$ ) showed GvpC binding mode with the GvpC sequence oriented to the direction of helical propagation.

Docking clusters were displayed in ChimeraX<sup>43</sup> and hydrogen bonds between GvpA and GvpC highlighted with the 'hbonds' command. The highest-scoring solution was manually fitted into a 2D class of the *A. flos-aquae* GV wall with GvpC.

## QUANTIFICATION AND STATISTICAL ANALYSIS

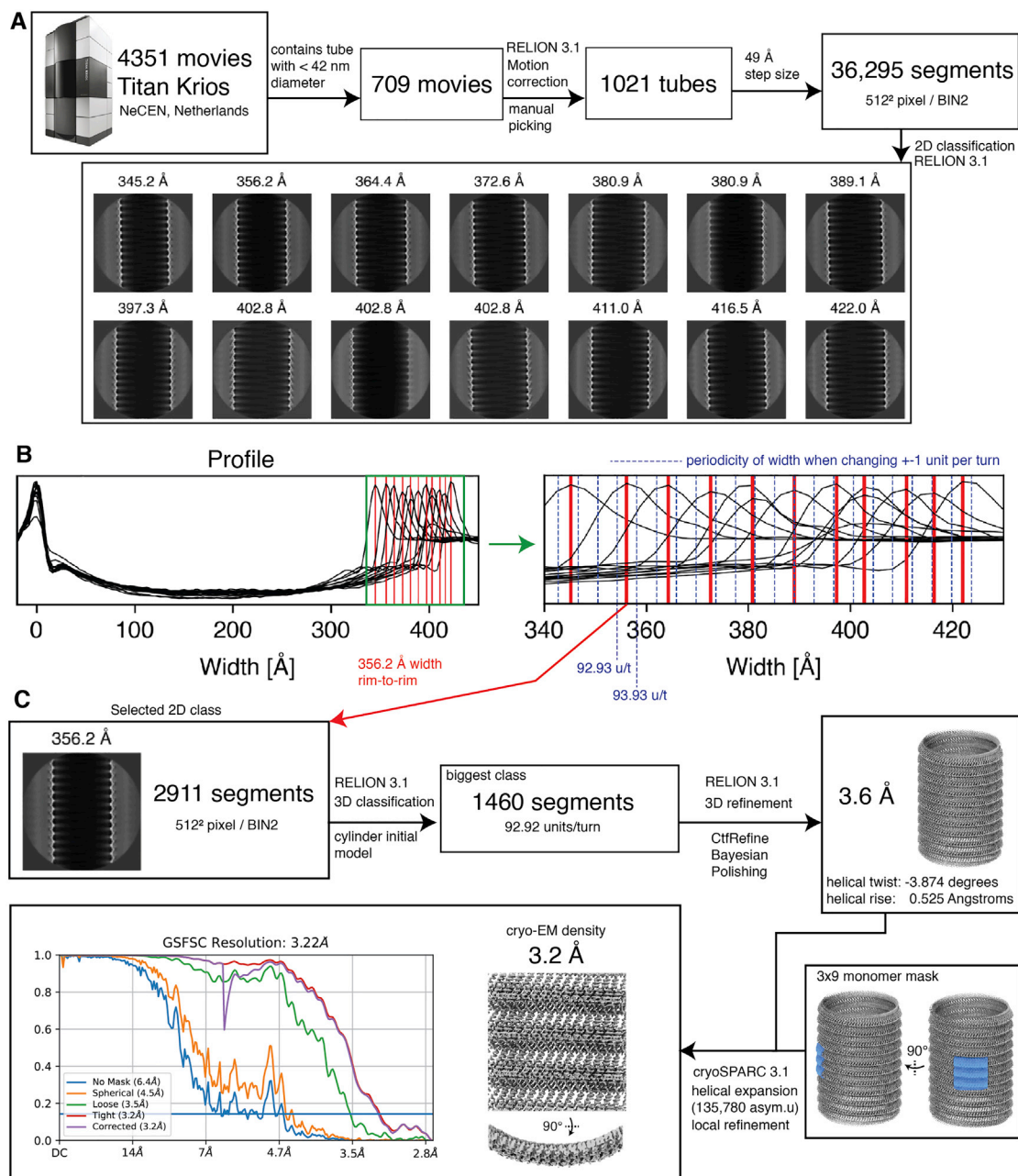
### Gas vesicle width measurements

Diameters of *A.flos-aquae* and *B.megaterium* GVs were determined from measurements of the widths of GV projection images from 20 cryo-EM micrographs each for both species. Figure 1E displays all individual measurement data points. Stated are mean and SD of the distributions.

### Collapse pressure measurements

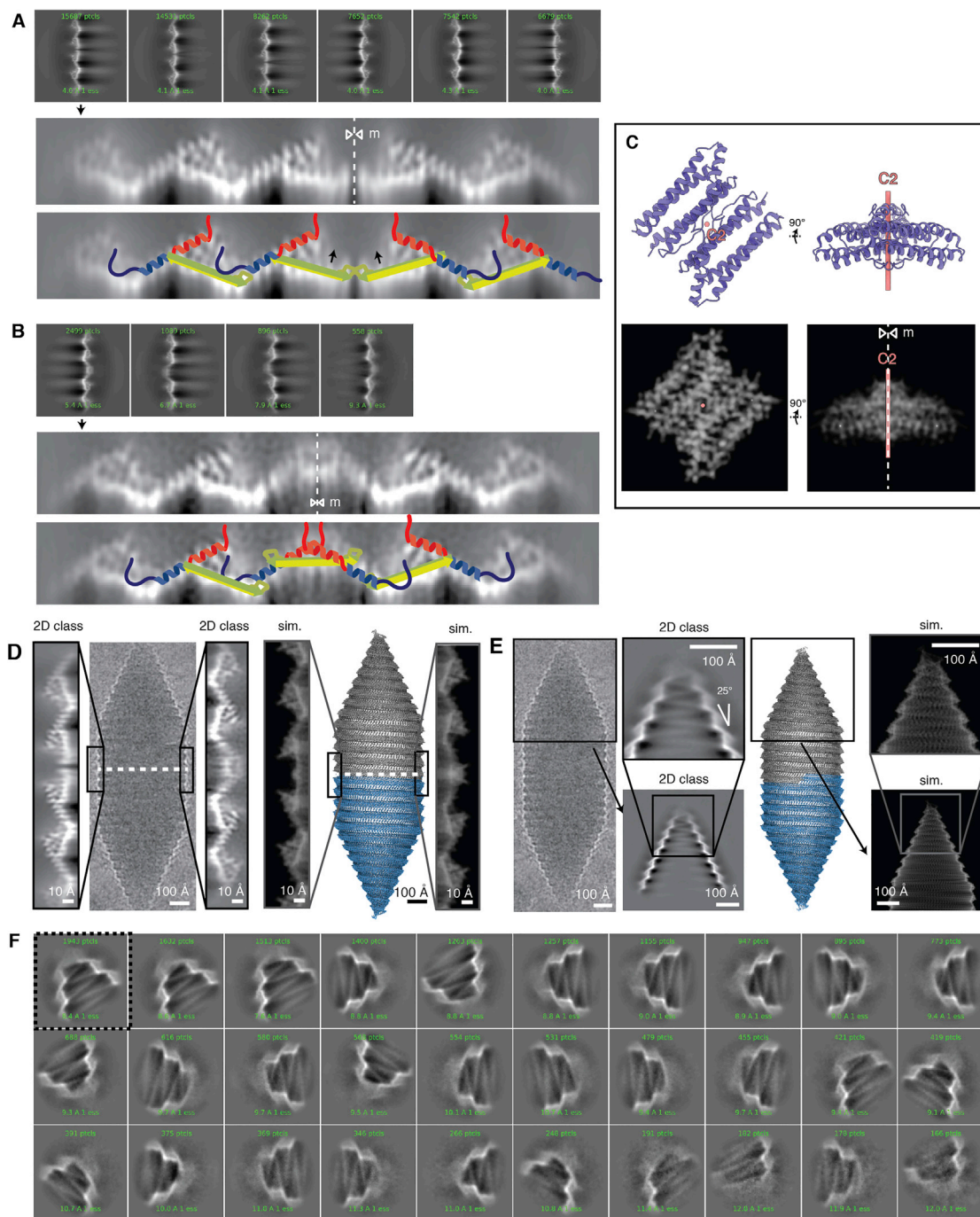
The collapse pressure of *A.flos-aquae* GVs supplemented with different GvpC mutants (Figure 6F and S6D) was determined as described in the STAR Methods details (section "Preparation of GVs with recombinant GvpC and collapse pressure measurements"). Displayed in Figure 6F and S6D are the collapse pressures and their standard deviations determined in a bootstrapping approach.

# Supplemental figures



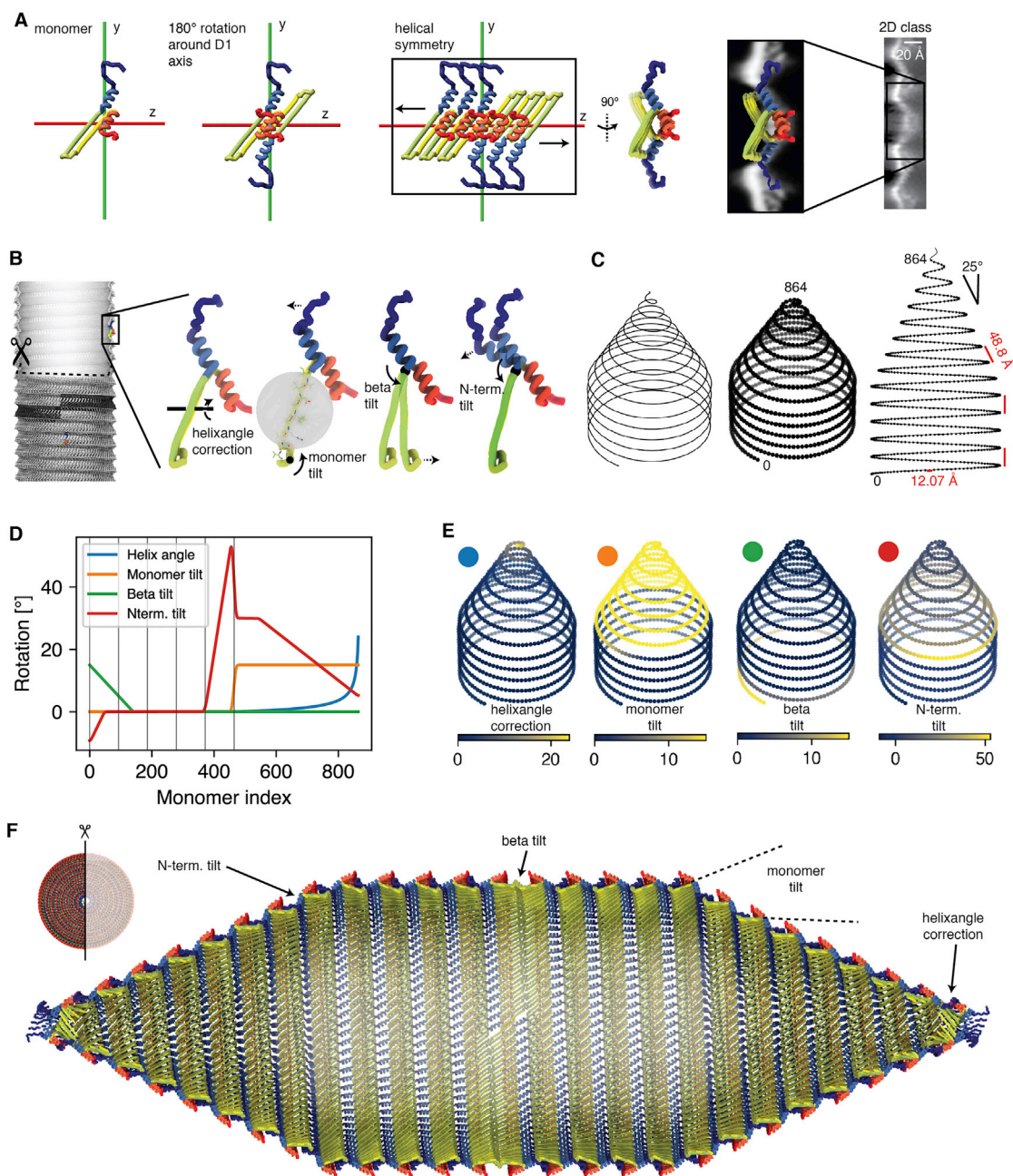
**Figure S1. Data processing of *B. megaterium* gas vesicle dataset, related to Figure 2**

(A) Preprocessing, manual picking, segment extraction and 2D classification leads to 2D class averages of GVs with different diameters. (B) The 2D classes were projected along the helical axis to generate profiles. The profiles were aligned with respect to the left peak. Zooming into the right peaks shows the distribution of GV widths in the 2D classes. Peaks are marked with a vertical red line. Blue lines indicate the periodicity of widths when an increment of one monomer per helical turn is assumed, based on a side-to-side distance of monomers of 12 Å, leading to a diameter increment of  $12/\pi = 3.8$  Å. Two to three different helical polymorphs are part of the particle subset belonging to a single 2D class average. (C) Processing steps starting from 2911 selected segments of a particular 2D class. The particle subset from the 2D class was further selected by 3D classification, imposing possible symmetry candidates between 90 and 95 units per wrung to select 1460 segments. Focused refinement on a 3x9 monomer segment of the wall in cryoSPARC 3.1<sup>21</sup> after symmetry expansion to 135,780 asymmetric units leads to the final result of a 3.2 Å resolution cryo-EM density of the GV wall.



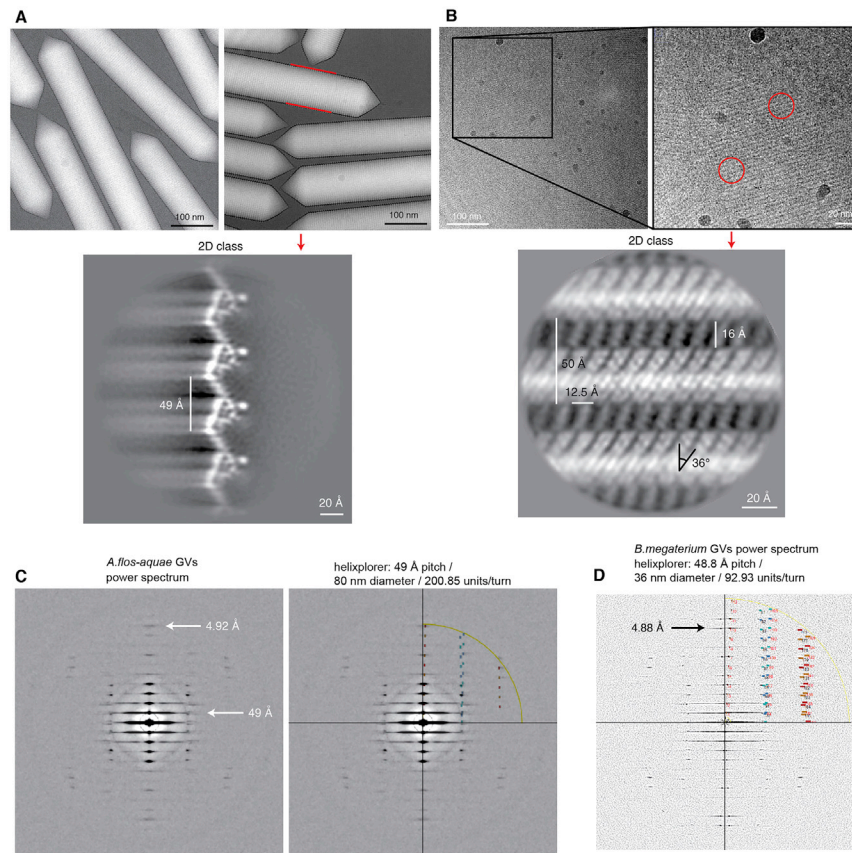
**Figure S2. 2D classification of *B. megaterium* seams and tips, related to Figure 3**

(A) 2D classes from the seam show perfect or near-perfect mirror symmetry.  $\beta$ -hairpins seem to hinge upwards at the seam (black arrows). (B) 2D classes from the putative PRP. Selected classes were magnified and sharpened for easier depiction. The mirror line is shown (m, dotted white line). A cartoon is drawn on the 2D classes to visualize GvpA molecules with the N-terminus in blue and the C-terminus in red. (C) Demonstration of the fact that views orthogonal to a 180° rotation axis show mirror symmetry in projection. Apoferritin dimer (PDB: 7ohf) is shown along the C2 axis and orthogonal to the C2 axis. 2D projection images (right) orthogonal to the C2 axis show mirror symmetry. (D) 2D class from the seam and PRP and simulated EM density from the pseudo-atomic model are in close agreement. (E) Pseudo-atomic model of a GV with simulated 2D projections of the tip, closely matching the experimental data. The 2D class average of GV tips with large box size reveals a linear decrease in radius at the tips with a cone angle of 25°. (F) 2D classes of GV tips with smaller box size reveal more detail, but all end in a blurry density at the tip. Alignment of secondary structure features is not possible and indicated strong structural heterogeneity at the tips.



**Figure S3. Construction of pseudo-atomic model of a whole *B. megaterium* gas vesicle, related to Figure 3**

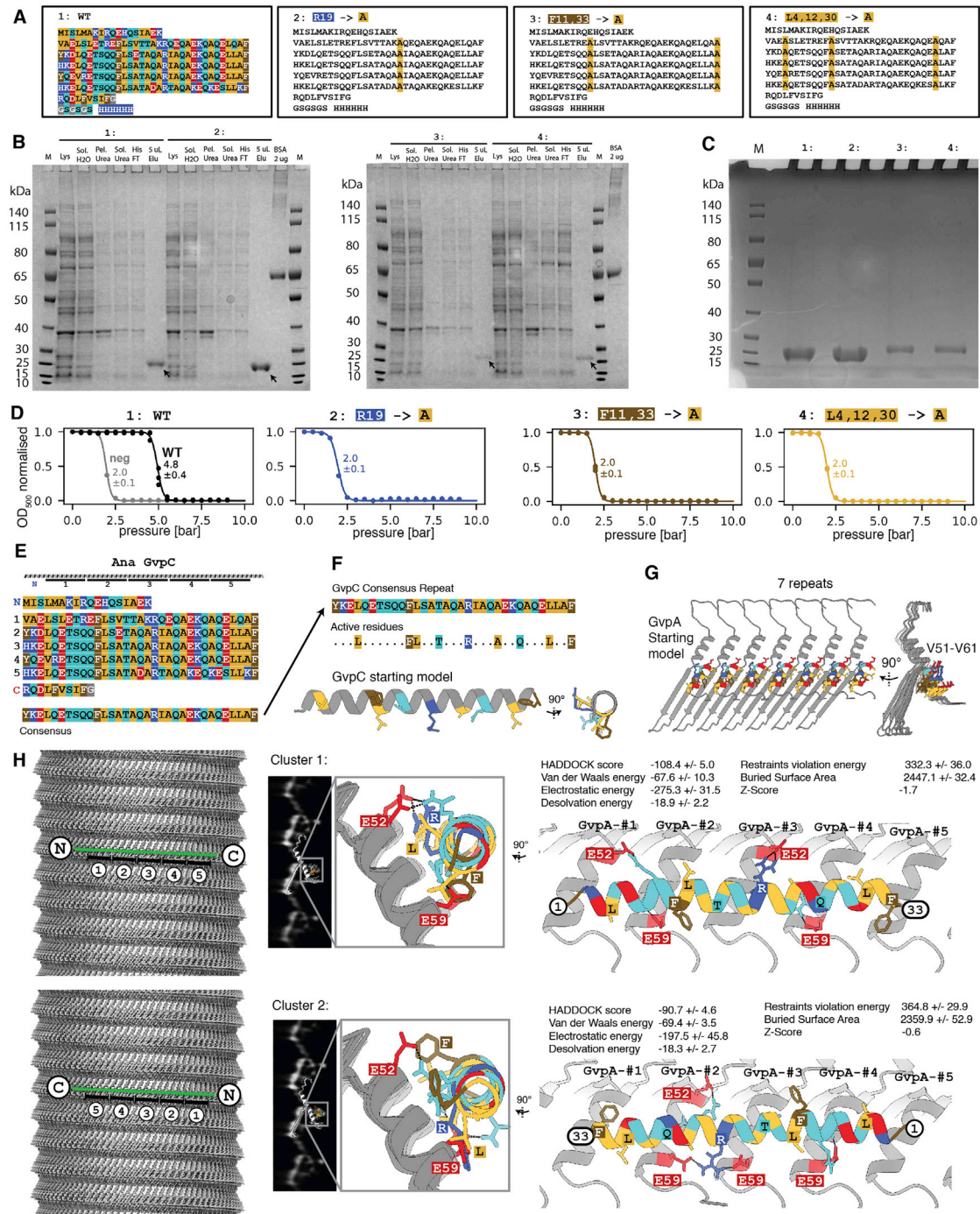
(A) A GvpA2 monomer was placed next to the symmetry axis (along  $x$ ) such that a  $180^\circ$  rotation would reproduce a view corresponding to the experimental 2D class average. The  $\beta$ -sheets meet in an angle at this stage, which is later corrected by tilting the sheets. (B) Four rotation parameters (helix angle correction, monomer tilt,  $\beta$ -hairpin tilt and N-term. tilt) used in the model are visualized. (C) The model is based on a helical curve in space with a linearly decreasing radius in the cones. The pitch in both the cylinder and cone is  $48.8 \text{ \AA}$ . 835 monomers are placed equidistantly on the curve with an intermonomer distance of  $12.07 \text{ \AA}$ . (D) The helix angle parameter was extracted from the curve, while the other three parameters were manually tuned to fit the 2D class average data, make the  $\beta$ -sheets line up in the PRP and avoid gaps by adjusting the N-termini. Vertical lines every 93 units depict the monomers on each wrung of the cylindrical part. (E) Parameters mapped onto the monomers (F) Cross-section of the final model highlighting the impact of the four model parameters. The N-terminal tilt is required to maintain the binding site of the N-terminus to the adjacent rib despite the constricting cone radius. The  $\beta$  tilt is required to make the orientation of the  $\beta$ -hairpin compatible with the 2D classes at the seam. The monomer tilt is required to make the monomer orientation compatible with the 2D classes of GV tips. The helix angle correction is required to accommodate the increasing helical angle in the constricting cone toward the tips.



**Figure S4. Cryo-EM of *A. flos-aquae* gas vesicles, related to Figure 5**

(A) Representative micrographs of *A. flos-aquae* GVs. GV edges were analyzed by 2D class averaging to give a low-noise high-resolution 2D view of the edges, to reveal a repetitive zigzag pattern. The 2D view shows details of at least  $5.4 \text{ \AA}$  resolution corresponding to the  $\alpha$ -helical pitch. (B) The same dataset contained collapsed GV wall segments. Those can be averaged as well by 2D class averaging to reveal a high-resolution top-view of the GV wall with better than  $4.8 \text{ \AA}$  resolution as the  $\beta$ -strands are resolved. (C) Computing the sum of in-plane rotated power spectra of segments of all GVs in the dataset gives rise to a layer-line pattern typical for helical assemblies. This approach can be seen as a form of fiber diffraction where fibers are aligned computationally. Overlay with computed layer line patterns (Helixplorer, <http://rico.ibs.fr/helixplorer/>) shows good agreement to a helix with  $49 \text{ \AA}$  pitch and  $200.85$  units per helical turn. (D) Power spectrum of *B. megaterium* power spectrum with first layer line at  $48.8 \text{ \AA}$ . The upper right quadrant is overlaid by a screenshot from Helixplorer (<http://rico.ibs.fr/helixplorer/>) used for interactive exploration of helical symmetry.





**Figure S6. Binding analysis of GvpC to *A.flos-aquae* gas vesicles, related to Figure 6**

(A) A GvpC wild-type construct and three mutants were designed with F, L and R residues in all five repeats mutated to alanine. All constructs have a C-terminal GSGSGS linker and 6x His-tag. (B) SDS-PAGE following purification steps of wild-type GvpC and point mutants. Arrow indicates final product. Legend as follows: P: Pellet/SolH2O: Soluble fraction in aqueous solution/Pel.Urea: Pelleting fraction in urea-containing buffer/Sol.Urea: Soluble fraction in urea-containing buffer/HisFT: Flowthrough from Ni-NTA column/5uLElu: Eluted fraction from Ni-NTA column (C) Overloading an SDS-PAGE gel with GvpC mutant sample indicates high degree of purity. (D) Collapse pressure measurements of the four constructs confirm binding to *A.flos-aquae* GV's of the wild-type construct and loss of binding for the mutants. Stripped GV's before re-addition of GvpC were measured as a negative control. Circles are measurement points from three independent re-addition experiments. Solid curves are fits of a sigmoid function, with the stated number being the pressure when the normalized OD500 drops to 0.5. (E) Primary sequence of Ana GvpC with 5 repeats and the consensus sequence of the repeats. (F) Identified highly conserved residues in the repeat are highlighted and chosen as 'active residues' in HADDOCK. A perfect  $\alpha$ -helical peptide is used as the starting model. (G) Amino acids V51 to V61 on  $\alpha$ -helix 2 are chosen as the

(legend continued on next page)

---

active residues (on a homology model of GvpA for *A.flos-aquae*) based on observed close proximity in 2D class averages. GvpA was repeated 7 times according to the helical symmetry of the solved *B.megaterium* assembly. (H) Two possible binding geometries of GvpC are shown on a *B.megaterium* GV density oriented with the seam on the bottom and the tip on the top: direction from C to N-terminus is following the left-handed loop of the helix, or reverse. HADDOCK protein docking results between GvpC and GvpA fall into two main clusters, one highest-scoring solution with a score of  $-108.4$  and a solution with reverse binding polarity of GvpC with a score of  $-90.7$ . The best solution of each cluster is shown. The solution was fitted into 2D class averages of *A.flos-aquae* edges.