Solid-State NMR Characterization of Gas Vesicle Structure

Astrid C. Sivertsen,[†‡] Marvin J. Bayro,[‡§] Marina Belenky,[†] Robert G. Griffin,[‡§] and Judith Herzfeld[†*]
[†]Department of Chemistry, Brandeis University, Waltham, Massachusetts; and [‡]Francis Bitter Magnet Laboratory and [§]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

ABSTRACT Gas vesicles are gas-filled buoyancy organelles with walls that consist almost exclusively of gas vesicle protein A (GvpA). Intact, collapsed gas vesicles from the cyanobacterium *Anabaena flos-aquae* were studied by solid-state NMR spectroscopy, and most of the GvpA sequence was assigned. Chemical shift analysis indicates a coil-α-β-β-α-coil peptide backbone, consistent with secondary-structure-prediction algorithms, and complementary information about mobility and solvent exposure yields a picture of the overall topology of the vesicle subunit that is consistent with its role in stabilizing an air-water interface.

INTRODUCTION

Gas vesicles are buoyancy organelles that are found in a wide range of microorganisms (1). The hollow protein shells (2) are assembled and disassembled in vivo, allowing the cells to change their overall density and thereby adjust their position in the water column in response to light intensity (1) and aeration (3). Of the 14 genes in the gas vesicle operon of *Halobacterium halobium*, eight are necessary and sufficient to code gas vesicle expression, and six of these have homologs in *Anabaena flos-aquae* (4). The gas vesicle operon has most likely spread by lateral gene transfer (5), such that the ability to form gas vesicles occurs more widely even than photosynthesis (1).

Electron microscopy of intact vesicles (6,7) shows shapes ranging from regular cylinders with conical end caps to irregular acornlike bodies. Vesicles are typically 5000 Å long and 750 Å wide, but proportions vary considerably across species, and even within a cell. The vesicles are bipolar, with an apparent insertion seam. To either side of the insertion seam, monomers of the ~7-kDa gas vesicle protein A (GvpA) are assembled into a low-pitch helix (8). GvpA is highly hydrophobic and negatively charged at neutral pH (9), with a highly conserved sequence that is predicted by PSIPRED (10) and other algorithms to contain both α-helix and β-strand secondary-structure segments. Substantial β-sheet is indicated by Fourier transform infrared spectroscopy (Mark Braiman, Syracuse University, personal communication, 1994) and x-ray diffraction (11). Matrix-assisted laser desorption/ionization time of flight studies of GvpA in intact *Anabaena flos-aquae* vesicles have shown that 1), only five residues can be cleaved from the C-terminus by using carboxypeptidase Y; 2), none of the nine D-X and E-X bonds in GvpA, situated throughout the sequence from E3-K4 to E58-A59 are accessible to endoproteinase GluC; and 3), only the N-terminal K4-T5 bond is accessible to trypsin (12) —except for a rarely observed cleavage of the R17-I18 bond (13), the other R-X and K-X bonds situated from R17-I18 to K55-Y56 are not cleaved.

Among the other gas vesicle proteins, GvpC is the best characterized. It is about three times larger than GvpA (14) and adheres to the outside of the shell formed by GvpA, strengthening the overall structure (15). In addition to GvpA and GvpC, the proteins GvpF, GvpG, GvpJ, GvpL, and GvpM have been found in *Halobacterium salinarum* gas vesicles in amounts detectable only by immunoblotting (i.e., well under 1% of overall gas vesicle protein). GvpJ and GvpM have partial sequence similarity with GvpA and are hypothesized to be involved in cap formation, whereas GvpF, GvpL, and GvpG are hypothesized to be involved in nucleation of the gas vesicle assembly process at the tips of the caps (16). In addition, the *gvpo* and *gvpk* genes are required for synthesis of gas vesicles, although GvpO and GvpK proteins have not been detected in gas vesicles, indicating that these proteins may act as chaperones in the gas vesicle assembly process.

The low-pitch helix of GvpA subunits gives rise to regularly spaced ribs running nearly perpendicular to the vesicle long axis, and it has been shown by x-ray diffraction from intact, partially aligned vesicles (11) and by atomic force microscopy (17) that the distance between neighboring ribs is 45.7 Å and that the ribs incorporate β-strands that are tilted at an angle of 36° from the vesicle long axis. This orients the H-bonds at a 54° angle relative to the vesicle axis, which is the ideal angle for mechanical stability (1), as the strength is the same in the length and width directions of the vesicle.

The hollow gas vesicles are filled with air of atmospheric composition, in equilibrium with gas molecules dissolved in the aqueous phase. Although the gas vesicle wall is permeable to H_2, N_2, O_2, Ar, CO, and CO_2, and even to a molecule as large as perfluorocyclobutane, C_4F_8 (6.3 Å in diameter), water vapor does not condense inside. This is believed to be due to a highly hydrophobic and tightly curved inner face without suitable nucleation sites (1,18). However, the details of the formation of the air-water interface are not known.

Submitted March 25, 2010, and accepted for publication June 7, 2010.

*Correspondence: herzfeld@brandeis.edu

Editor: Kathleen B. Hall.

0006-3495/10/09/1932/8 $2.00

doi: 10.1016/j.bpj.2010.06.041

A recent solid-state NMR study of intact, collapsed vesicles from *Anabaena flos-aquae* (19) has shown duplication of several of the most easily assigned resonances, suggesting that GvpA is a member of the growing group of metamorphic proteins (20). Using secondary-structure-prediction algorithms, the duplicated resonances were interpreted in terms of asymmetric GvpA dimers that can explain how the strands of an antiparallel β-sheet can be tilted at an angle of 36° relative to the gas vesicle axis while pairing oppositely charged side chains in salt bridges on one side of the β-sheet and leaving the other side fully hydrophobic. However, the accuracy of the predicted secondary structure remained unknown.

In this article, we describe magic-angle-spinning NMR studies that extend resonance assignments to 81% of the peptide backbone and thereby allow characterization of the secondary structure of GvpA. These extensive resonance assignments also allow site-specific characterization of mobility and solvent accessibility.

MATERIALS AND METHODS

Sample preparation

Anabaena flos-aquae cells were grown under a $^{13}CO_2$ and $^{15}N_2$ atmosphere, and the gas vesicles were isolated, stripped of GvpC, collapsed, and prepared for magic-angle-spinning experiments as described previously (19). Approximately 10 and 24 mg of sample were packed into 2.5- and 3.2-mm rotors, respectively. A ^{2}H-exchanged sample was prepared by multiple rounds of centrifugally accelerated flotation of inflated vesicles in 99% $^{2}H_2O$ (Cambridge Isotope Laboratories, Andover, MA). Vesicles were then collapsed and ~28 mg of sample was packed into a 3.2-mm rotor. The solvent included 15% $^{2}H_8$-glycerol (wt/wt) to prevent dehydration of the samples.

NMR spectroscopy

Solid-state NMR experiments were carried out using custom-designed spectrometers (D. J. Ruben, Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, MA) operating at 700 MHz and 750 MHz ^{1}H Larmor frequencies, and a Bruker spectrometer (Billerica, MA), operating at 900 MHz ^{1}H Larmor frequency. The 750 and 900 MHz spectrometers were equipped with triple-resonance $^{1}H/^{13}C/^{15}N$ Bruker probes with 2.5-mm stators, and the 700 MHz spectrometer was equipped with a triple-resonance $^{1}H/^{13}C/^{15}N$ Varian-Chemagnetics (Palo Alto, CA) probe with a 3.2-mm stator. The magic angle spinning (MAS) frequency was controlled to ± 5 Hz using Bruker MAS controllers, and the samples were cooled with a stream of cold N_2 gas during experiments, maintaining exit gas temperatures of -40°C to 5°C. Sample temperatures are estimated to be ~10°C higher than the exit gas temperature, depending on the MAS frequency and the intensity of radio frequency irradiation. All spectra were referenced to external 2,2-dimethyl-2-silapentane-5-sulfonic acid (21,22).

One-dimensional spectra were acquired at 700 MHz ^{1}H Larmor frequency, 12.5 MHz MAS, and approximate sample temperatures of $+3$°C and -30°C. The cross-polarization (23) (CP) and direct-polarization spectra are the averages of 64 scans each, whereas the spectra from insensitive nuclei enhanced by polarization transfer (INEPT) (24) are the average of 256 scans each. Homonuclear ^{13}C-^{13}C and heteronuclear ^{15}N-^{13}C-^{13}C correlation spectra were acquired using various two-dimensional dipolar recoupling experiments at 5°C. A ^{13}C-^{13}C radio-frequency-driven dipolar recoupling (RFDR) (25–27) spectrum was recorded at 750 MHz ^{1}H Larmor frequency and 18 kHz MAS. A 2D NCOCX spectrum was obtained at 900 MHz ^{1}H Larmor frequency and 20 kHz MAS, using specific CP (28,29) for ^{15}N-^{13}CO polarization transfer and 40 ms of proton-driven spin diffusion mixing (30,31) for subsequent ^{13}CO-^{13}CX transfer. A NCACX spectrum was obtained at 900 MHz ^{1}H Larmor frequency and 20 kHz MAS, using specific CP for ^{15}N-^{13}CA polarization transfer followed by 6.4 ms of homonuclear mixing via RFDR. Both heteronuclear experiments were recorded with 128 real and 128 imaginary points in the indirect dimension and 2048 points in the direct dimension, with dwell times of 100 μs and 12 μs, respectively, and 83 kHz two-pulse phase modulation decoupling (32) during the chemical shift evolution and detection periods. The spectra are the average of 384 transients per t_1 point. A 2D INEPT-through-bond-correlation-spectroscopy (TOBSY) spectrum was acquired at 750 MHz ^{1}H Larmor frequency and 8.333 kHz MAS. In this experiment, ^{1}H-^{13}C magnetization transfer via INEPT (24) was used to excite mobile sites, after which 5.76 ms of TOBSY (33,34) ^{13}C-^{13}C mixing was applied using the $P9_3^1$ condition (35) with a 50 kHz ^{13}C radio frequency field. This spectrum was acquired with 256 real and 256 imaginary points in the indirect dimension and 2048 points in the direct dimension, with dwell times of 60 μs and 12 μs, respectively, and 50 kHz two-pulse phase modulation decoupling during the chemical shift evolution and detection periods. The spectra are the average of 224 scans for each t_1 point. All data sets were processed with the NMRPipe package (36) and analyzed using Sparky (37).

RESULTS

Spectral quality

Intact, collapsed gas vesicles are highly ordered molecular assemblies and present highly resolved MAS NMR spectra with average line widths of 0.5 and 0.8 ppm for ^{13}C and ^{15}N resonances, respectively (equivalent to full width at half-height of 80 and 70 Hz, respectively, at 700 MHz ^{1}H Larmor frequency). These line widths are comparable to those observed for such microcrystalline proteins as BPTI (38) and ubiquitin (39), and such amyloid fibrils as those formed by α-synuclein (40,41), PI3-SH3 (42), and HET-s(218-289) (43), but they do not allow one-bond ^{13}C-^{13}C *J*-couplings to be resolved. Fig. 1 shows CP and direct-polarization ^{13}C MAS spectra acquired at 3°C and -30°C. A certain degree of line broadening is observed at the lower temperature, indicating some freezing of dynamic disorder, in particular for side-chain ^{13}C resonances.

The CP signal enhancement observed is well below the ~2.5 factor observed for rigid peptides and proteins and reveals the highly mobile nature of GvpA. This dynamic feature also manifests itself in the low intensity of many of the crosspeaks in dipolar correlation spectra, despite their narrow line widths, as can be observed in the ^{13}C-^{13}C spectrum of Fig. 2. However, as shown below, only a small number of residues exhibited solution-like mobility, and we therefore relied mainly on dipolar recoupling experiments for establishing intraresidue and sequential correlations.

Sequential resonance assignments

Narrow ^{15}N and ^{13}C line widths allowed us to obtain site-specific sequential resonance assignments via 2D dipolar

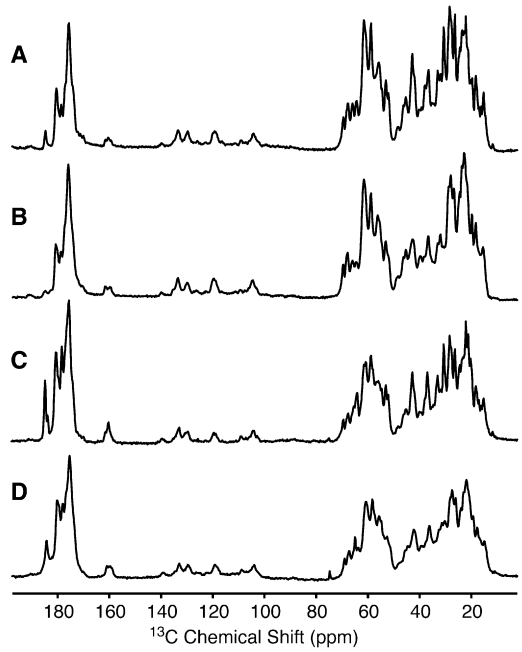

FIGURE 1 ^{13}C MAS spectra of gas vesicles acquired at 700 MHz ^{1}H Larmor frequency and 12.5 kHz MAS frequency, using cross-polarization (*A* and *B*) and direct polarization (*C* and *D*) at 3°C (*A* and *C*) and −30°C (*B* and *D*). The relatively low cross-polarization enhancement over direct excitation indicates considerable mobility in the protein. Each spectrum is an average of 64 scans recorded with recycle delays of 3 s and 30 s in CP and DP experiments, respectively. A glycerol resonance is visible at ~75 ppm in *D*.

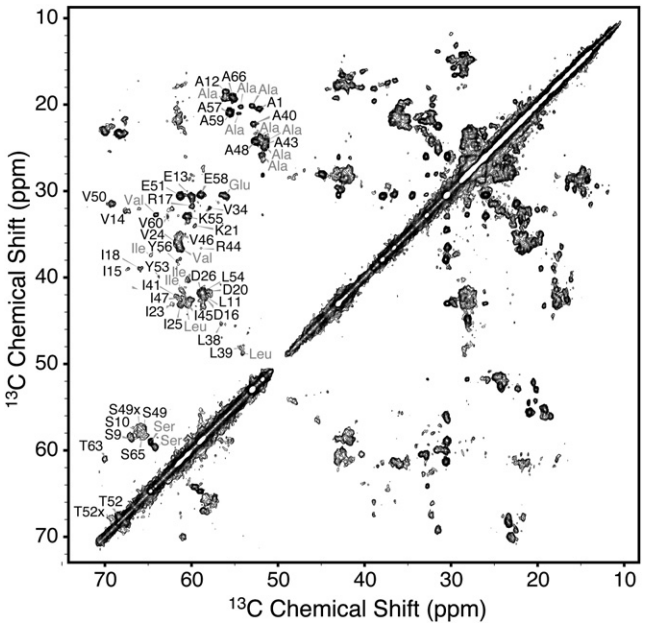

FIGURE 2 ^{13}C-^{13}C correlation spectrum acquired at 750 MHz ^{1}H Larmor frequency, 18 kHz MAS frequency, and 5°C. The experiment used 3.52 ms of RFDR mixing to establish homonuclear connectivities. Cross-peaks corresponding to CA-CB correlations for several assigned residues are indicated.

correlation experiments. In particular, employing CC (Fig. 2), NCACX (Fig. 3), and NCOCX (not shown) 2D correlation spectra, we were able to establish connectivities that in combination permit assignment of ^{15}N and ^{13}C resonances for 57 of the 70 residues, or 81% of the amino acid sequence, of GvpA. A list of assigned ^{15}N and ^{13}C chemical shifts is provided in Table S1 in the Supporting Material. As described and discussed earlier (19), certain resonances are duplicated, in particular those of several Gly and Ala residues, along with S49, T52, Y53, and Y56, and sequential assignment of duplicated resonances was often impeded by the relatively low intensity of these signals.

The N-terminal residue, A1, can be identified due to its unique ^{15}N chemical shift, and sequential correlations can be established to V2 and E3. Residues K4–S8 have not been assigned because they present weak signals that severely limit interresidue correlations. The segment S9–W28 has been assigned, and it gives rise to backbone and side-chain correlations that are generally of medium intensity. The residues V29–L33 again present only weak correlation signals and are not assigned. The segment V34–V66 has been assigned sequentially, and it gives rise to backbone and side-chain signals of medium intensity for the segment V34–I47 and strong intensity for the segment A48–V66. The remaining four C-terminal residues have not been assigned, with the exception of P69, which presents weak signals that can be assigned unambiguously due to the characteristic downfield ^{15}N chemical shift of proline. Several side-chain resonances in the main assigned segments (S9–W28 and V34–V66) either are not observed due to line-broadening dynamics or remain unresolved in highly repetitive spin systems (as expected, given 11 Ala, 8 Ile, and 10 Val residues in a 70-residue protein).

Although backbone ^{13}C and ^{15}N resonances tend to present significant degeneracy, it was possible to resolve assignment ambiguities in many instances by relying on

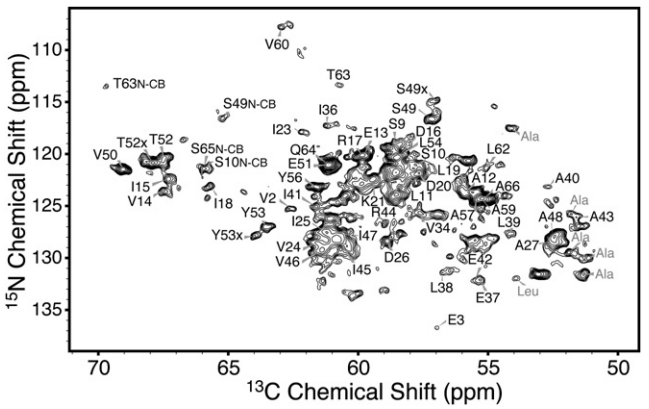

FIGURE 3 ^{13}Cα region of an NCACX spectrum acquired at 900 MHz ^{1}H Larmor frequency, 20.0 kHz MAS frequency, and 5°C. The experiment used ^{15}N-^{13}C cross-polarization followed by ^{13}C-^{13}C RFDR mixing to establish heteronuclear connectivities. Labels correspond to N-CA correlations, unless otherwise noted.

side-chain correlations in NCACX and NCOCX spectra. Resonance intensities vary widely across the GvpA sequence, such that certain parts of the amino acid sequence, especially the C-terminal third, are predominant in the ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra. The segments S9–W28 and V34–I47 generally have weak- or medium-intensity correlations, but especially weak resonances arise from residues such as I15, L19, I36, E37, L38, L39, E42, and R44. The segment A48–V66 generally gives rise to correlations with high intensities, except for the duplicated residues Y53 and Y56, the duplicate form of T52, and residues V60, L62, and T63, which give rise to resonances of moderate intensity.

The side chains beyond CB of several residues in the GvpA sequence are observed in full (in the combined spectra, not necessarily in just one type of spectrum). These residues are L11, V14, I15, I18, I23, V24, I25, L38, L39, V46, I47, V50, T52, V60, L62, and T63. On the other hand, incomplete side chains are observed for I36, V34, I41, I45, and L54. For almost all the acidic residues (Glu and Asp), carboxyl correlations are missing in most spectra. Among the three of each in the sequence, only one Arg and one Lys side chain are observed in full. The Y53 side chain readily shows up in full in RFDR and spin-diffusion spectra with short mixing times, whereas longer mixing times are required for the entire Y56 side chain to be observed. This dissimilarity is likely the result of different ring dynamics in the two Tyr residues. The ring system of W28 has not been observed in full; only a CB-CG correlation has been observed in ^{13}C-^{13}C spectra with long mixing times.

After assignment to the degree possible, two Ser, one Val, one Glu, and eight Ala CA-CB correlations in RFDR spectra remain unassigned. The Ser, Val, and Glu correlations could come from unassigned parts of the protein, but only one Ala remains unassigned in the sequence. Therefore, the rest of the Ala CA-CB resonances are likely to be peak duplications. In the CA regions of NCACX and NCOCX spectra, another ~20 residues could be assigned to the yet unassigned peaks. This number is again higher than the actual number of unassigned residues, indicating duplications that are more extensive than the ones assigned so far.

Secondary-structure analysis

Values of the Cα-Cβ (44) secondary chemical shifts (45,46) are shown in Fig. 4 *A*, and they display strong agreement with the PSIPRED secondary-structure prediction shown in Fig. 4 *B*. A TALOS (47) chemical shift analysis of the assigned residues gives the secondary-structure pattern shown in Fig. 4 *C*, which largely agrees with the secondary chemical shift result and with the PSIPRED prediction. TALOS results are only shown for one form for the residues that display peak duplications, as the TALOS results for the two forms of GvpA do not differ by more than 3–4°, which is within the standard deviation of the TALOS output.

The N-terminal α-helix is predicted by PSIPRED to extend from L11 to I18, and both secondary chemical shifts and the TALOS calculation indicate the beginning of the helix at L11. However, our chemical shift analysis indicates that the helix may continue until D20. It is worth noting that the T5–R17 fragment that has been observed in digestion studies (13) corresponds to cleavage at this transition between the N-terminal α-helix and the N-terminal β-sheet that follows. The latter is predicted by PSIPRED to extend from V24 to V34, and the secondary chemical shifts and TALOS results indicate that it begins at I23 and continues for a few residues. Resonance assignments were not obtained for the second half of this sequence, but there is a secondary chemical shift for V34 that has a β-sheet value. According to our chemical shift analysis the C-terminal β-strand extends from I36 to S49, whereas the segment V50–A59 adopts an α-helical conformation, in close agreement with PSIPRED. It is predicted that the remainder of the sequence adopts a coil conformation, and neither secondary chemical shifts nor TALOS indicate any significant stretches of defined secondary structure for this part of the protein. TALOS results show nonstandard angles for five residues: K21, G22, G35, V60, and G61. Because of their large conformational flexibility, it is not surprising that atypical ϕ and ψ values are observed for Gly residues, especially as they are found in intermediate regions. The unusual torsion angles of Gly neighbors K21 and V60 are also consistent with the Gly residues found in turns between the secondary-structure elements. The Ramachandran plot in Fig. 5 shows that the assigned residues reside in the allowed α-helix and β-sheet regions, except for the few mentioned above.

Amphipathic α-helices

Fig. 6 shows nearest-neighbor-corrected (46,48) secondary ^{15}N backbone chemical shift values for the α-helical segments of GvpA. As all residues in the figure are α-helical, the differences in ^{15}N chemical shifts are mainly governed by differences in NH-CO hydrogen-bond length along the helix, where large secondary chemical shifts indicate short NH-CO hydrogen bonds (45,49). The plots display variations of the secondary chemical shifts, with a period of roughly three to four residues, corresponding to an α-helical turn. A similar periodicity is found in amide proton resonances of curved amphipathic helices, where backbone hydrogen bonds are shorter on the hydrophobic side and longer on the hydrophilic side (46,50). Since amide ^{15}N chemical shifts exhibit trends with respect to hydrogen-bond length that are similar to those of amide ^{1}H shifts, we may speculate that the periodicity observed in the ^{15}N shifts for the N-teminal helix (Fig. 6, *left side*) might be the result of curvature in the helix, whereas the trend in the C-terminal helix is less defined. These observations are consistent with the

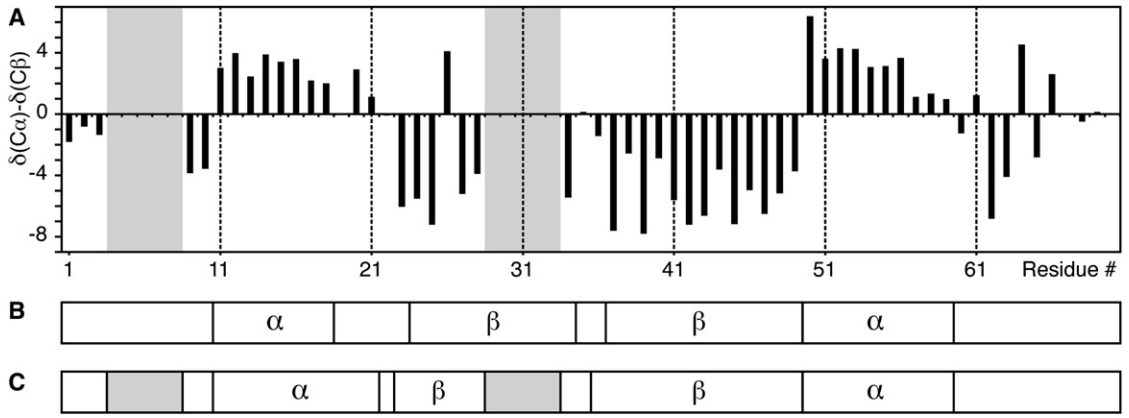

FIGURE 4 Secondary-structure analysis. (*A*) CA-CB secondary chemical shifts (gray areas indicates unassigned sequences). (*B*) PSIPRED prediction (α-helical, β-sheet, and coil regions). (*C*) TALOS results (notation as in *B*, but with unassigned/noncalculated segments in gray).

amphipathic residue patterns seen in helical wheel renderings (see Fig. S1).

Molecular mobility

Fig. 7 shows ^{1}H-^{13}C INEPT spectra acquired at 3°C (Fig. 7 *A*) and −30°C (Fig. 7 *B*). In the solid state, the INEPT sequence excites those ^{13}C sites with sufficient flexibility to average out the dipolar interactions to the point that they do not interfere with the much weaker J-couplings. Typically, only mobile side chains and highly flexible backbone regions give rise to INEPT signals. As Fig. 7 shows, a few signals are observed at 3°C, in most cases corresponding to single sites. About eight CA sites are observed, indicating that at this temperature most of the backbone is not very flexible. The number of side chains corresponds roughly to the number of backbone sites. At −30°C, there is much less mobility, and no CA sites are observed, whereas some side chain sites are still observed, mostly methyl groups. The spectra in Figs. 1 and 7 together indicate that GvpA in intact collapsed vesicles is overall moderately dynamic, but presents only a few sites with solution-like mobility.

To assign the highly mobile residues, a 2D ^{13}C-^{13}C INEPT-TOBSY spectrum was acquired (Fig. 8). In this experiment, after excitation of mobile ^{13}C sites, TOBSY mixing is applied to achieve homonuclear polarization transfer mediated by ^{13}C-^{13}C J-couplings. From the nearly complete ^{13}C spin systems in these data, we were able to identify three Ala, two Val, and one Pro set of resonances that are also observed, at least weakly, in dipolar correlation spectra. In addition, resonance sets were detected for Ser, Thr, Leu, Ile, Glu, Gln, Lys, and Arg. Except for the unique proline, P69, none of the observed INEPT-TOBSY resonances could be assigned sequentially. However, the Thr resonances are likely to correspond to T5, since T52 and T63 are already assigned in dipolar correlation spectra.

Solvent accessibility

To gain information about the GvpA subunit topology that forms the gas vesicle wall, we acquired water-edited correlation spectra and studied a $^{2}H_2O$-exchanged sample.

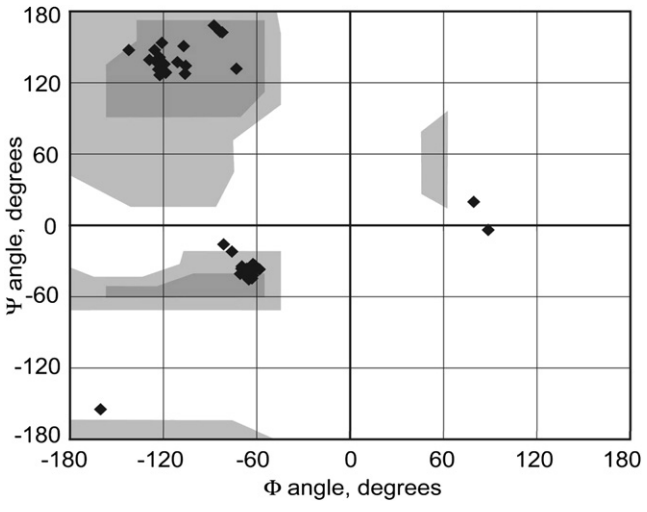

FIGURE 5 Ramachandran plot for GvpA using dihedral angles from TALOS. Gray areas indicate allowed regions for nonglycine residues (*dark gray* indicates core areas).

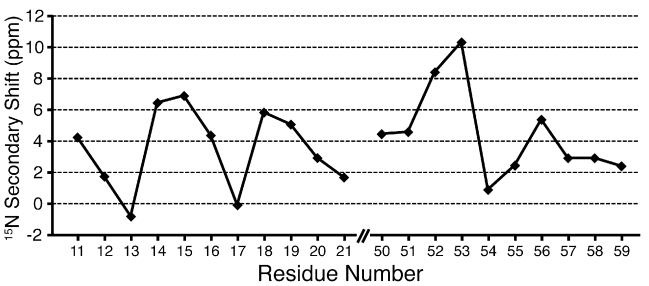

FIGURE 6 Nearest-neighbor corrected ^{15}N secondary shifts, showing periodic patterns characteristic of curved helices.

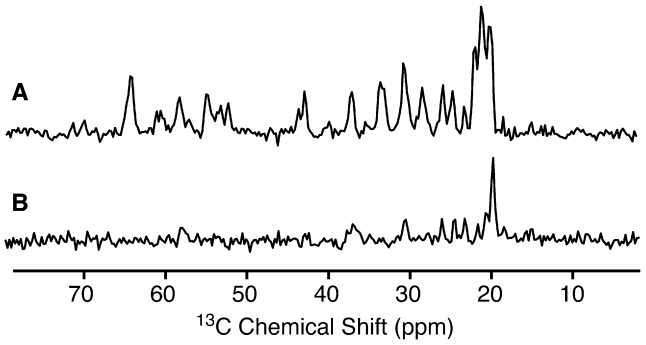

FIGURE 7 ^{1}H-^{13}C INEPT spectra acquired at 700 MHz ^{1}H Larmor frequency and 12.5 kHz MAS at (*A*) 3°C and (*B*) −30°C. Each spectrum is the average of 256 scans. Only highly mobile ^{13}C sites are observable in this type of experiment.

Water-edited spectra can be used to identify solvent-accessible residues by delaying ^{1}H magnetization transfer to ^{13}C until the intrinsic ^{1}H polarization in the protein is relaxed by relatively strong dipolar interactions. However, the spectra we recorded in this fashion show that considerable ^{1}H polarization remains in the protein after the filter period, possibly due to motion of certain segments of the polypeptide chain. Therefore, water-edited 1D and 2D ^{13}C-^{13}C correlation experiments yielded ambiguous information.

On the other hand, a sample in $^{2}H_2O$ allowed the identification of exchangeable amide sites by analyzing ^{15}N-^{13}C double-CP spectra recorded with different contact times for the initial ^{1}H-^{15}N CP step. Short ^{1}H-^{15}N CP periods are sufficient to transfer magnetization only to ^{15}N sites with directly bonded protons, whereas longer CP times allow polarization to build up at neighboring sites. Fig. 9 shows NCA spectra of vesicles collapsed in $^{2}H_2O$ recorded with ^{1}H-^{15}N CP contact times of 0.5 ms (Fig. 9 *A*) and 8.0 ms (Fig. 9 *B*). Most NCA crosspeaks appear at both CP times, but certain peaks that are not observed at 0.5 ms CP time are observed with an 8-ms contact time, including: A1, V2, T52, V60, T63, and several unassigned NCA peaks, some of which appear to correspond to residues with doubled resonances. In a similar way, although Lys Nz-Ce correlations are observed at 8 ms CP time, they are not seen at 0.5 ms CP time (not shown), and thus we can conclude that all three Lys side chains are solvent-accessible. On the other hand, since the majority of NCA correlations are observed at 0.5 ms (Fig. 9 *B*), most amide sites do not exchange with $^{2}H_2O$, either due to participation in secondary-structure hydrogen bonding, or because they are buried within the solvent-inaccessible protein core.

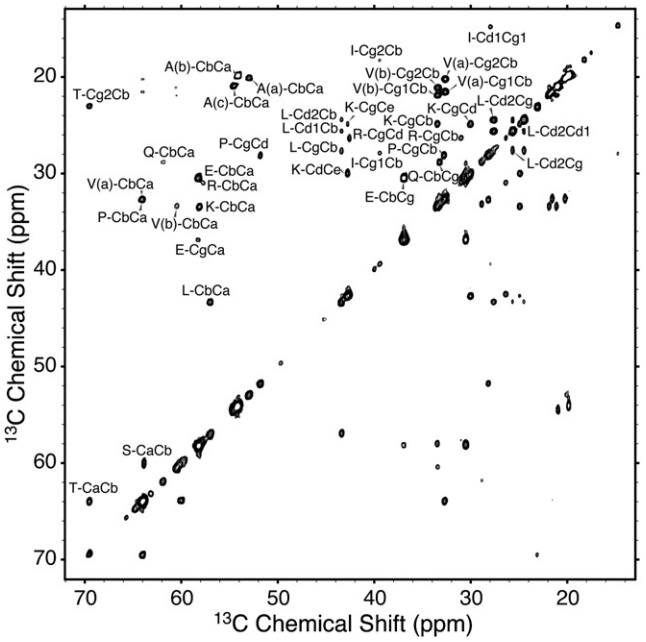

FIGURE 8 Magic-angle spinning ^{13}C-^{13}C INEPT-TOBSY spectrum recorded at 750 MHz ^{1}H Larmor frequency, 8.333 kHz MAS frequency, and 3°C to establish homonuclear correlations between highly mobile ^{13}C nuclei. The labels indicate the likely amino acid types as determined by the chemical shifts of each spin system.

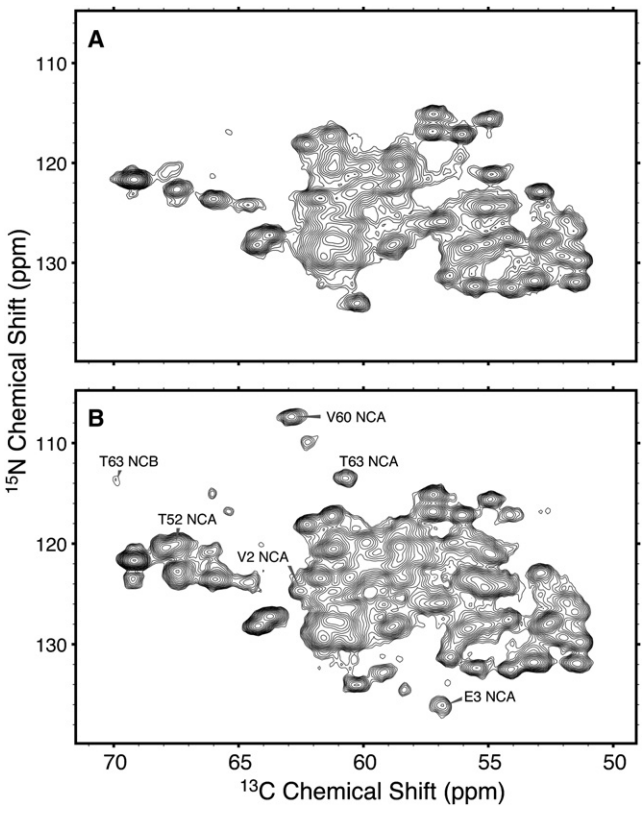

FIGURE 9 NCA spectra of vesicles collapsed in $^{2}H_2O$ with (*A*) 0.5 and (*B*) 8.0 ms ^{1}H-^{15}N CP time. At short CP times, deuterium-exchanged ^{15}N nuclei are not polarized and their NCA crosspeaks are not visible, whereas at long CP times these ^{15}N nuclei are polarized by distant protons. The spectra, acquired at 700 MHz ^{1}H Larmor frequency and 5°C, were processed with 30 Hz Lorentzian to Gaussian line broadening in each dimension.

DISCUSSION

GvpA subunit structure

Solid-state NMR assignments for 81% of the GvpA sequence support the coil-α-β-β-α-coil secondary structure that is predicted for this protein. This new information makes it necessary to reject the so-called paperclip model of gas vesicle subunit structure (1), which assumed GvpA to be in an all-β-sheet conformation. As seen in Fig. 4, the β-sheet region has a maximum length of only 27 residues, and is flanked by α-helices at both ends. It is also clear that the termini beyond the helices are in coil conformation and solvent-accessible, as indicated by ^2H_2O exchange at certain sites near the termini.

Air-water interface formation

Gas vesicles function by forming an interface between the aqueous phase and the air inside the vesicle. In contrast, ordinary proteins function fully immersed in condensed aqueous or lipid phases, and tend to unfold on exposure to an air-liquid interface (51,52), with the hydrophobic core exposed to the air. On the other hand, extended β-sheets with an alternating polar-nonpolar amino acid sequence are stable at an air-water interface, as their amphipathic nature allows them to expose the hydrophobic side chains on one face to the air, and the hydrophilic side chains on the other face to the water. Indeed, amyloid fibrils have been found to self-assemble in air-water interfaces, with their hydrophobic side chains exposed to the air (53). Interfacial amyloids have even been found to occur naturally: the soil bacteria *Streptomyces coelicolor* secretes chaplin proteins (54) and filamentous fungi secrete hydrophobins (55), both proteins that self-assemble into amyloid-like fibrils at the air-water interface, thereby lowering the surface tension of water and presumably facilitating the growth of spore-forming hyphae into the air.

Gas vesicles have two segments that our NMR data indicate to be of β-sheet secondary structure, as predicted. Both segments have alternating polar-nonpolar residue patterns that are highly conserved, although they are otherwise known to be disfavored by nature because of their propensity for amyloid formation (56). It is likely that this portion of GvpA is largely responsible for the air-water interface, as the presence of an inward-facing hydrophobic surface would explain not only the ability of gas vesicles to prevent internal condensation of water, but also the strong 10 Å reflection, typical of stacked β-sheets, that is observed in x-ray diffraction studies of collapsed, but not of inflated, gas vesicles (11).

Perhaps the most intriguing aspect of the functional amyloids is how cells prevent these molecules from accumulating in dysfunctional deposits. In particular, how do aquatic organisms control the assembly-disassembly of gas vesicles and the recycling of GvpA subunits? Chaperones (possibly the GvpO and GvpK proteins that are essential for gas vesicle expression, but are not detected in gas vesicles) are likely to be involved in the process, and the presence of the amphipathic α-helices in GvpA may help prevent catastrophic aggregation and precipitation. A similar issue arises in the secretion of chaplins and hydrophobins.

CONCLUSIONS

We have obtained well-resolved solid-state NMR spectra of intact, collapsed gas vesicles from *Anabaena flos-aquae*, allowing assignment of 81% of the GvpA sequence, followed by secondary chemical shift analysis, indicating a coil-α-β-β-α-coil secondary structure. Nearest-neighbor corrected ^{15}N secondary shifts provide evidence for curvature in the α-helices and ^2H_2O exchange experiments identify solvent-exposed residues. These data provide insights into the topology of GvpA subunits in the gas vesicle shell and how it might stabilize the air-water interface. However, distance constraints are needed to obtain a structural model.

SUPPORTING MATERIAL

A table of resonance assignments for GvpA (Table S1) and helical wheels for the α-helices of GvpA (Fig. S1) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00784-8.

The authors thank Drs. Patrick van der Wel and Anthony Bielecki for helpful discussions and technical support.

Funding for this work was provided by National Institutes of Health grants EB002175, EB001960, and EB002026.

REFERENCES

1. Walsby, A. E. 1994. Gas vesicles. *Microbiol. Rev.* 58:94–144.
2. Walsby, A. E., and B. Buckland. 1969. Isolation and purification of intact gas vesicles from a blue-green alga. *Nature.* 224:716–717.
3. Walsby, A. E. 1976. The buoyancy-providing role of gas vacuoles in an aerobic bacterium. *Arch. Microbiol.* 109:135–142.
4. Offner, S., A. Hofacker, …, F. Pfeifer. 2000. Eight of fourteen gvp genes are sufficient for formation of gas vesicles in halophilic archaea. *J. Bacteriol.* 182:4328–4336.
5. Boucher, Y., C. J. Douady, …, W. F. Doolittle. 2003. Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* 37: 283–328.
6. Bowen, C. C., and T. E. Jensen. 1965. Blue-green algae: fine structure of the gas vacuole. *Science.* 147:1460–1462.
7. Walsby, A. E., and H. H. Eichenberger. 1968. The fine structure of gas-vacuoles released from cells of the blue-green alga *Anabaena flos-aquae*. *Arch. Microbiol.* 60:76–83.
8. Offner, S., U. Ziese, …, F. Pfeifer. 1998. Structural characteristics of halobacterial gas vesicles. *Microbiology.* 144:1331–1342.
9. Hayes, P. K., A. E. Walsby, and J. E. Walker. 1986. Complete amino acid sequence of cyanobacterial gas-vesicle protein indicates a 70-residue molecule that corresponds in size to the crystallographic unit cell. *Biochem. J.* 236:31–36.
10. McGuffin, L. J., K. Bryson, and D. T. Jones. 2000. The PSIPRED protein structure prediction server. *Bioinformatics.* 16:404–405.
11. Blaurock, A. E., and A. E. Walsby. 1976. Crystalline structure of the gas vesicle wall from *Anabaena flos-aquae*. *J. Mol. Biol.* 105:183–199.

12. Belenky, M., R. Meyers, and J. Herzfeld. 2004. Subunit structure of gas vesicles: a MALDI-TOF mass spectrometry study. *Biophys. J.* 86: 499–505.

13. Dunton, P. G., W. J. Mawby, …, A. E. Walsby. 2006. Analysis of tryptic digests indicates regions of GvpC that bind to gas vesicles of *Anabaena flos-aquae*. *Microbiology.* 152:1661–1669.

14. Hayes, P. K., C. M. Lazarus, …, A. E. Walsby. 1988. The protein encoded by *gvpC* is a minor component of gas vesicles isolated from the cyanobacteria *Anabaena flos-aquae* and *Microcystis* sp. *Mol. Microbiol.* 2:545–552.

15. Hayes, P. K., B. Buchholz, and A. E. Walsby. 1992. Gas vesicles are strengthened by the outer-surface protein, GvpC. *Arch. Microbiol.* 157:229–234.

16. Shukla, H. D., and S. DasSarma. 2004. Complexity of gas vesicle biogenesis in *Halobacterium* sp. strain NRC-1: identification of five new proteins. *J. Bacteriol.* 186:3182–3186.

17. McMaster, T. J., M. J. Miles, and A. E. Walsby. 1996. Direct observation of protein secondary structure in gas vesicles by atomic force microscopy. *Biophys. J.* 70:2432–2436.

18. 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.

19. Sivertsen, A. C., M. J. Bayro, …, J. Herzfeld. 2009. Solid-state NMR evidence for inequivalent GvpA subunits in gas vesicles. *J. Mol. Biol.* 387:1032–1039.

20. Murzin, A. G. 2008. Biochemistry. Metamorphic proteins. *Science.* 320:1725–1726.

21. Wishart, D. S., C. G. Bigam, …, B. D. Sykes. 1995. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR.* 6:135–140.

22. Morcombe, C. R., and K. W. Zilm. 2003. Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson.* 162:479–486.

23. Pines, A., M. G. Gibby, and J. S. Waugh. 1973. Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* 59:569–590.

24. Morris, G. A., and R. Freeman. 1979. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* 101:760–762.

25. Bennett, A. E., R. G. Griffin, …, S. Vega. 1992. Chemical shift correlation spectroscopy in rotating solids: radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* 96:8624–8627.

26. Bennett, A. E., C. M. Rienstra, …, R. G. Griffin. 1998. Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* 108:9463–9479.

27. Bayro, M. J., R. Ramachandran, …, R. G. Griffin. 2008. Radio frequency-driven recoupling at high magic-angle spinning frequencies: homonuclear recoupling sans heteronuclear decoupling. *J. Chem. Phys.* 128:052321.

28. Schaefer, J., E. O. Stejskal, …, R. A. McKay. 1984. Quantitative determination of the concentrations of ^{13}C-^{15}N chemical bonds by double cross-polarization NMR. *J. Magn. Reson.* 59:150–156.

29. Baldus, M., A. T. Petkova, …, R. G. Griffin. 1998. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* 95:1197–1207.

30. Szeverenyi, N. M., M. J. Sullivan, and G. E. Maciel. 1982. Observation of spin exchange by two-dimensional Fourier transform ^{13}C cross polarization-magic-angle spinning. *J. Magn. Reson.* 47:462–475.

31. Suter, D., and R. R. Ernst. 1985. Spin diffusion in resolved solid-state NMR spectra. *Phys. Rev. B Condens. Matter.* 32:5608–5627.

32. Bennett, A. E., C. M. Rienstra, …, R. G. Griffin. 1995. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 103:6951–6958.

33. Baldus, M., and B. H. Meier. 1996. Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J. Magn. Reson. A.* 121:65–69.

34. Baldus, M., R. J. Iuliucci, and B. H. Meier. 1997. Probing through-bond connectivities and through-space distances in solids by magic-angle-spinning nuclear magnetic resonance. *J. Am. Chem. Soc.* 119: 1121–1124.

35. Hardy, E. H., R. Verel, and B. H. Meier. 2001. Fast MAS total through-bond correlation spectroscopy. *J. Magn. Reson.* 148:459–464.

36. Delaglio, F., S. Grzesiek, …, A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.

37. Goddard, T. D. and D. G. Kneller. SPARKY 3. University of California, San Francisco.

38. McDermott, A., T. Polenova, …, E. K. Paulsen. 2000. Partial NMR assignments for uniformly (^{13}C, ^{15}N)-enriched BPTI in the solid state. *J. Biomol. NMR.* 16:209–219.

39. Igumenova, T. I., A. E. McDermott, …, A. J. Wand. 2004. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126:6720–6727.

40. Heise, H., W. Hoyer, …, M. Baldus. 2005. Molecular-level secondary structure, polymorphism, and dynamics of full-length α-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. USA.* 102:15871–15876.

41. Kloepper, K. D., D. H. Zhou, …, C. M. Rienstra. 2007. Temperature-dependent sensitivity enhancement of solid-state NMR spectra of α-synuclein fibrils. *J. Biomol. NMR.* 39:197–211.

42. Bayro, M. J., T. Maly, …, R. G. Griffin. 2009. Long-range correlations between aliphatic ^{13}C nuclei in protein MAS NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* 48:5708–5710.

43. Ritter, C., M.-L. Maddelein, …, R. Riek. 2005. Correlation of structural elements and infectivity of the HET-s prion. *Nature.* 435:844–848.

44. Luca, S., D. V. Filippov, …, M. Baldus. 2001. Secondary chemical shifts in immobilized peptides and proteins: a qualitative basis for structure refinement under magic angle spinning. *J. Biomol. NMR.* 20:325–331.

45. Wishart, D. S., and B. D. Sykes. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* 239:363–392.

46. Wishart, D. S., and D. A. Case. 2001. Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* 338:3–34.

47. Cornilescu, G., F. Delaglio, and A. Bax. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR.* 13:289–302.

48. Braun, D., G. Wider, and K. Wüthrich. 1994. Sequence-corrected ^{15}N "random coil" chemical shifts. *J. Am. Chem. Soc.* 116:8466–8469.

49. de Dios, A. C., J. G. Pearson, and E. Oldfield. 1993. Secondary and tertiary structural effects on protein NMR chemical shifts: an ab initio approach. *Science.* 260:1491–1496.

50. Zhou, N. E., B. Y. Zhu, …, R. S. Hodges. 1992. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic. α.-helical peptides. *J. Am. Chem. Soc.* 114:4320–4326.

51. Tronin, A., T. Dubrovsky, …, C. Nicolini. 1996. Role of protein unfolding in monolayer formation on air-water interface. *Langmuir.* 12: 3272–3275.

52. Maa, Y.-F., and C. C. Hsu. 1997. Protein denaturation by combined effect of shear and air-liquid interface. *Biotechnol. Bioeng.* 54: 503–512.

53. Schladitz, C., E. P. Vieira, …, H. Möhwald. 1999. Amyloid-β-sheet formation at the air-water interface. *Biophys. J.* 77:3305–3310.

54. Claessen, D., R. Rink, …, H. A. Wosten. 2003. A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev.* 17:1714–1726.

55. Mackay, J. P., J. M. Matthews, …, M. D. Templeton. 2001. The hydrophobin EAS is largely unstructured in solution and functions by forming amyloid-like structures. *Structure.* 9:83–91.

56. Broome, B. M., and M. H. Hecht. 2000. Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. *J. Mol. Biol.* 296:961–968.