BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

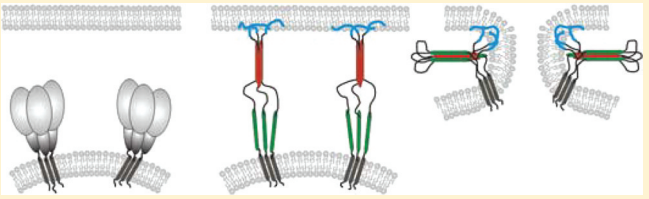
Solid-State Nuclear Magnetic Resonance (NMR) Spectroscopy of Human Immunodeficiency Virus gp41 Protein That Includes the Fusion Peptide: NMR Detection of Recombinant Fgp41 in Inclusion Bodies in Whole Bacterial Cells and Structural Characterization of Purified and Membrane-Associated Fgp41

Erica P. Vogel, Jaime Curtis-Fisk, Kaitlin M. Young, and David P. Weliky*

Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States

Supporting Information

ABSTRACT: Human immunodeficiency virus (HIV) infection of a host cell begins with fusion of the HIV and host cell membranes and is mediated by the gp41 protein, a single-pass integral membrane protein of HIV. The 175 N-terminal residues make up the ectodomain that lies outside the virus. This work describes the production and characterization of an ectodomain construct containing the 154 N-terminal gp41 residues, including the fusion peptide (FP) that binds to target cell membranes. The Fgp41 sequence was derived from one of the African clade A strains of HIV-1 that have been less studied than European/North American clade B strains. Fgp41 expression at a level of ∼100 mg/L of culture was evidenced by an approach that included amino acid type ^{13}CO and ^{15}N labeling of recombinant protein and solid-state NMR (SSNMR) spectroscopy of lyophilized whole cells. The approach did not require any protein solubilization or purification and may be a general approach for detection of recombinant protein. The purified Fgp41 yield was ∼5 mg/L of culture. SSNMR spectra of membrane-associated Fgp41 showed high helicity for the residues C-terminal of the FP. This was consistent with a "six-helix bundle" (SHB) structure that is the final gp41 state during membrane fusion. This observation and negligible Fgp41-induced vesicle fusion supported a function for SHB gp41 of membrane stabilization and fusion arrest. SSNMR spectra of residues in the membrane-associated FP provided evidence of a mixture of molecular populations with either helical or β-sheet FP conformation. These and earlier SSNMR data strongly support the existence of these populations in the SHB state of membrane-associated gp41.



The human immunodeficiency virus (HIV) is enveloped by a membrane obtained during budding from an infected host cell. An early step in HIV infection of a new cell is joining or "fusion" of the HIV and host cell membranes. This process is catalyzed by the ∼350-residue HIV gp41 protein that is an integral membrane protein of the viral envelope.[1] The ∼175 N-terminal residues form the ectodomain that lies outside HIV (see Figure 1a for most of the ectodomain sequence and identification of structural and functional regions). Prior to fusion, gp41 is noncovalently associated with the gp120 protein (Figure 1b). Productive infection begins with binding of gp120 to receptor proteins in a target cell membrane and is followed by dissociation of gp120 from gp41.[2] There are ensuing structural changes of gp41 and likely binding of the ∼20-residue N-terminal "fusion peptide" (FP) region to target cell membranes with concurrent changes in the two membranes, including mixing of lipids, formation of a single hemifusion diaphragm bilayer that separates the HIV and cell contents, and opening of the diaphragm to form a single membrane that encloses HIV and the cell.[3] Although there are no high-resolution structures of full-length gp41, other structural and functional data support (1) trimeric gp41, (2) an early stage "prehairpin intermediate" (PHI) state with a parallel trimer of fully extended ectodomains between the HIV membrane and the FP in the cell membrane, and (3) a final "six-helix bundle" (SHB) state with a gp41 trimer with each gp41 molecule having an N-helix−turn−C-helix hairpin structure and parallel N-helices in the trimer interior and parallel C-helices on the trimer exterior (Figure 1b).[4−7] Studies of cell−cell fusion induced by gp120−gp41 complexes indicate that most membrane fusion steps with the exception of diaphragm opening occur prior to formation of the final SHB state.[8]

The importance of the FP in fusion and infection has been highlighted by decreases in both functions with point mutations in the FP.[9] The current understanding of gp41 is also based on smaller fragments of gp41 where fusogenic function has typically been assayed by fragment-induced perturbation or fusion of membrane vesicles. One such fragment is the HIV fusion peptide (HFP) that corresponds to the 20−30 N-

Received: August 13, 2011
Revised: October 7, 2011
Published: October 10, 2011

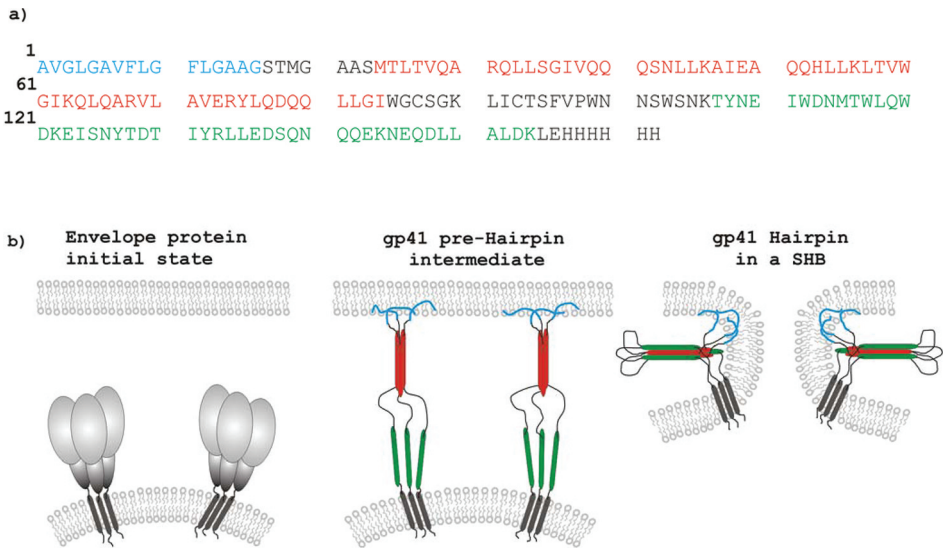


Figure 1. (a) Fgp41 amino acid sequence with regions encompassing the fusion peptide, N-helix, and C-helix colored blue, red, and green, respectively. (b) Conceptual representation of gp41 structural states with time increasing from left to right and the same color coding as panel a. The sequence is from the HIV-1 clade A isolate, AY288087 or Q45D5. This is a primary isolate from a female patient in Kenya. The last eight residues of the sequence are non-native. The red and green regions were based on continuous helical regions in high-resolution structures of different soluble ectodomain gp41 constructs with the hairpin conformation (shown schematically in panel b). In these SHB structures, the red and green helices are antiparallel to one another.

terminal residues of gp41 and has moderate fusogenicity.[10] The functional significance of the PHI trimeric topology (Figure 1b) has been supported by the high fusogenicity of (1) a cross-linked HFP trimer (HFPtr) and (2) "N70", the 70 N-terminal residues of gp41.[10−13] The higher fusogenicity of N70 relative to that of the HFP may also have a contribution from the N-helix residues that are C-terminal of the FP. Much larger ectodomain constructs have also been produced with N-helix and C-helix regions and form the thermostable SHB structure (Figure 1b), which is the final gp41 state. Different approaches were used to obtain these FP-containing hairpin constructs. In one approach, the FP and hairpin regions were produced separately by chemical synthesis and bacterial expression, respectively, and the "FP-hairpin" was then made by native chemical ligation.[11,12] In another approach, a chimera was expressed in *Escherichia coli* and contained an N-terminal molecular carrier protein (e.g., glutathione *S*-transferase) followed by the gp41 ectodomain.[14−16] The carrier was cleaved during purification. There are conflicting results from different studies of the fusogenicity of such hairpin constructs with reports of both very high levels of fusion and no fusion. There were some differences among the studies, including (1) deletion of loop residues between the N- and C-helices in some constructs, (2) lipid compositions of the vesicles including different fractions of negatively charged lipid, (3) the use of smaller and less stable sonicated vesicles versus larger and more stable extruded vesicles, and (4) pH values that ranged between 3.0 and 7.5.[17]

Structural studies have also been conducted for some of the aforementioned fragments. A helical monomer HFP has been observed in detergent with one report of a continuous helix between residues 4 and 22.[18−22] However, to the best of our knowledge, the HFP does not induce fusion between detergent micelles. The structure of the membrane-associated HFP has been probed mostly by solid-state nuclear magnetic resonance (SSNMR) spectroscopy with supporting data from other techniques such as infrared spectroscopy.[23−25] SSNMR spectra of the HFP associated with membranes lacking cholesterol show distinct populations of predominant β-sheet and predominant α-helical molecules, while the HFP associated with membranes with ~30 mol % cholesterol shows only the β-sheet conformation with an antiparallel alignment of adjacent hydrogen-bonded HFPs.[10,26−30] The biological relevance of membrane cholesterol is supported by the ~25 mol % cholesterol in host cell membranes and the ~45 mol % cholesterol in HIV membranes.[31] The FP structure appears to be similar in the highly fusogenic HFP trimer and in N70, whereas the FP-hairpin construct with SHB structure showed approximately equal populations of molecules with either β-sheet or helical FP structures.[10,12,29] For membrane-associated N70, the N-helix residues appear to be predominantly helical and N70 is recognized by an antibody specific for trimeric coiled-coil N-helices.[32] There are several high-resolution structures of hairpin constructs without the FP that show (1) the hairpin structure of individual molecules and (2) molecular trimers with SHB structure.[4−7]

SSNMR requires production of multimilligram quantities of isotopically labeled protein, and protein yields may be reduced by ligation and/or cleavage steps. This was the reason behind one of the goals of this study, expression of the FP-containing gp41 ectodomain ("Fgp41") in bacteria without a chimera or ligation. This goal seemed reasonable because recently developed protocols yielded 20 mg of protein/L of culture for the full-length "FHA2" ectodomain of the influenza virus fusion protein.[33−35] There is considerable diversity among HIV protein sequences in patient sera and in cell cultures. This was the reason behind a second goal of this study, functional and structural experiments on a gp41 ectodomain sequence that differed from the sequence of the earlier studies to address the generality of the functional and structural findings across strains of HIV. In these earlier studies, the sequence was from the HXB2 strain of HIV-1 that was first created in a cell culture in 1984 and is grouped with "clade B" HIV-1 prevalent in patients in North America and Europe.[36] In some contrast, the gp41

sequence in this study is from the primary HIV-1 isolate Q45D5 from the sera of a newly infected Kenyan woman.[37] The Q45D5 isolate is grouped with clade A HIV-1 that is prevalent in Central and East Africa. The HXB2 and Q45D5 gp41 ectodomain sequences are provided in the Supporting Information. A third goal of this study was to conduct functional and SSNMR structural studies to compare with the FP-hairpin construct in which 46 contiguous residues, including the native loop, were replaced with the non-native SGGRGG sequence.[11,12,17] The FP-hairpin did not induce vesicle fusion and inhibited fusion by constructs such as N70. The deletion of these residues in the FP-hairpin may be important as a 35-residue peptide that included the native loop region induced vesicle fusion under some conditions.[38] The Fgp41 construct of the present study has the native sequence without deletion.

MATERIALS AND METHODS

Materials. The Fgp41 plasmid was constructed within the pET24a(+) vector using the DNA sequence of the Q45D5 primary isolate that is grouped with clade A of HIV-1. The plasmid was transformed into BL21(DE3) chemically competent *E. coli* cells (Novagen, Gibbstown, NJ). The Fgp41 amino acid sequence is given in Figure 1a, and the correlative sequence from plasmid DNA extracted from the cells is provided in the Supporting Information. Other reagents and their sources were as follows: Luria-Bertani broth (LB) medium (Acumedia, Lansing, MI), *n*-octyl β-D-thioglucopyranoside (bTOG) and isopropyl β-D-thiogalactopyranoside (IPTG) (Anatrace, Maumee, OH), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (sodium salt) (POPG) (Avanti Polar Lipids, Alabaster, AL), and isotopically labeled amino acids (Cambridge Isotope Laboratories, Andover, MA). Other materials were obtained from Sigma-Aldrich (St. Louis, MO).

Protein Production and Purification. The protocol was based on a previous protocol for the influenza virus fusion protein ectodomain.[34] One key feature was initial bacterial growth in rich medium to high cell densities. Relative to initial growth in minimal medium, protein production was augmented by the cell densities and by the larger number of ribosomes per cell. Bacterial cell cultures were grown in media containing 15 mg/L kanamycin because the pET24a(+) vector contains a gene for kanamycin resistance. Bacterial cells in 1 mL of an 80/20 (v/v) H_2O/glycerol mixture were added to two 2.8 L baffled fernbach flasks that each contained 1 L of LB and were capped with a foam plug. Bacterial growth to an OD_{600} of ≈4 occurred during overnight incubation at 37 °C with shaking at 140 rpm. The cell suspensions were centrifuged (10000*g* for 10 min), and the cell pellets were harvested and then resuspended in a single flask containing 1 L of fresh medium with M9 minimal salts, 2.0 mL of 1.0 M $MgSO_4$, and 5.0 mL of a 50% glycerol solution. Growth resumed after incubation at 37 °C for approximately 1 h. At this time, 100 mg of 1-^{13}C-labeled amino acid per liter and 100 mg of ^{15}N-labeled amino acid per liter (or 100 mg of 1-^{13}C- and ^{15}N-labeled amino acid per liter) were added to the medium. IPTG was then added to a final concentration of 2 mM, which induced expression of Fgp41 (6 h at 23 °C). The cell pellet was harvested after centrifugation and stored at −80 °C. The wet cell mass was ∼8 g.

A variant of this procedure was used to make whole cell NMR samples and included (1) 50 mL culture volumes in 250 mL flasks, (2) overall reagent concentrations similar to those described above, including ∼10 mg of labeled amino acid(s), and (3) in some cases lyophilization of the cells.

The Fgp41 purification was based on the C-terminal His tag, and the following protocol is for ∼3 g of wet cells. Most Fgp41 was in inclusion bodies, so cells were lysed in a 40 mL solution containing 1% sodium dodecyl sulfate (SDS) as well as 50 mM sodium phosphate (pH 8.0), 300 mM NaCl, and 1 mM imidazole. Lysis was accomplished by sonication (4 × 1 min) with a tip sonifier (Branson, Danbury, CT), and the soluble and insoluble portions of the lysate were then separated by centrifugation (48000*g* for 20 min at 4 °C). The soluble lysate was mixed with 0.25 mL of cobalt His-Select resin (prepared using the manufacturer's instructions) and then agitated at ambient temperature for 1 h. The resin with bound Fgp41 was separated from the lysate by flow through a fritted column. Purification was accomplished by passing buffer [50 mM sodium phosphate (pH 8.0), 300 mM NaCl, and 0.5% SDS] through the resin with stepwise increases in imidazole concentration. Weakly bound protein was first eluted with buffer containing 1, 20, or 50 mM imidazole, and tightly bound protein (predominantly Fgp41) was subsequently eluted with buffer containing 250 mM imidazole. Excess SDS was precipitated by overnight incubation at 4 °C and removed by centrifugation (16000*g* for 1 min). Fgp41 was typically dialyzed into "HEPES/MES" buffer [5 mM HEPES and 10 mM MES (pH 7.4)]. Fgp41 concentrations were quantified using an ε_{280} of 46000 M^{-1} cm^{-1}, which was estimated from the sum of the ε_{280} values of individual tryptophans and tyrosines.

Circular Dichroism (CD) Spectroscopy. Spectra were recorded using a CD instrument (Chirascan, Applied Photophysics, Surrey, U.K.), a 1 mm path length, a 260−200 nm spectral window, wavelength points separated by 0.5 nm, and 1 s signal averaging per point. Fgp41 samples were prepared by precipitation of excess SDS followed by overnight dialysis into HEPES/MES buffer (pH 7.4) with DTT added at 2 times the molar concentration of Fgp41 to prevent disulfide bond formation. Most spectra were obtained with 20 μM Fgp41. For each sample, a reference spectrum was also taken of buffer without Fgp41 and the relevant Fgp41 spectrum was the difference between the spectra for Fgp41 and buffer and buffer only.

Lipid Mixing. One early step in fusion between the HIV and target cell membranes is mixing of lipids between the two membranes. This aspect of Fgp41 fusogenicity was probed by fluorescence-detected Fgp41-induced mixing of lipids between membrane vesicles. A set of vesicles that contained POPC and POPG lipids in a 4:1 molar ratio was prepared, and another set of "labeled" vesicles that contained an additional 2 mol % of the fluorescent lipid *N*-NBD-PE and 2 mol % of the quenching lipid *N*-Rh-PE was prepared. Large unilamellar vesicles (LUVs) were prepared by (1) dissolving lipids in chloroform and then removing chloroform by nitrogen gas and overnight vacuum, (2) formation of pH 7.5 aqueous lipid dispersions with ≈5 mM total lipid and 25 mM HEPES, including five freeze−thaw cycles, and (3) ∼20-fold extrusion through a polycarbonate filter with 0.1 μm diameter pores. The assay was conducted at 37 °C with continuous stirring in HEPES buffer using a mixture of unlabeled vesicles (135 μM total lipid) and labeled vesicles (15 μM total lipid). After measurement of the initial fluorescence F_0, an aliquot of 30 μM Fgp41 in HEPES/MES buffer was added to the vesicle solution so that the final Fgp41 concentration was 3 μM and the Fgp41:total lipid ratio was 0.02. Fgp41-induced fusion between labeled and unlabeled

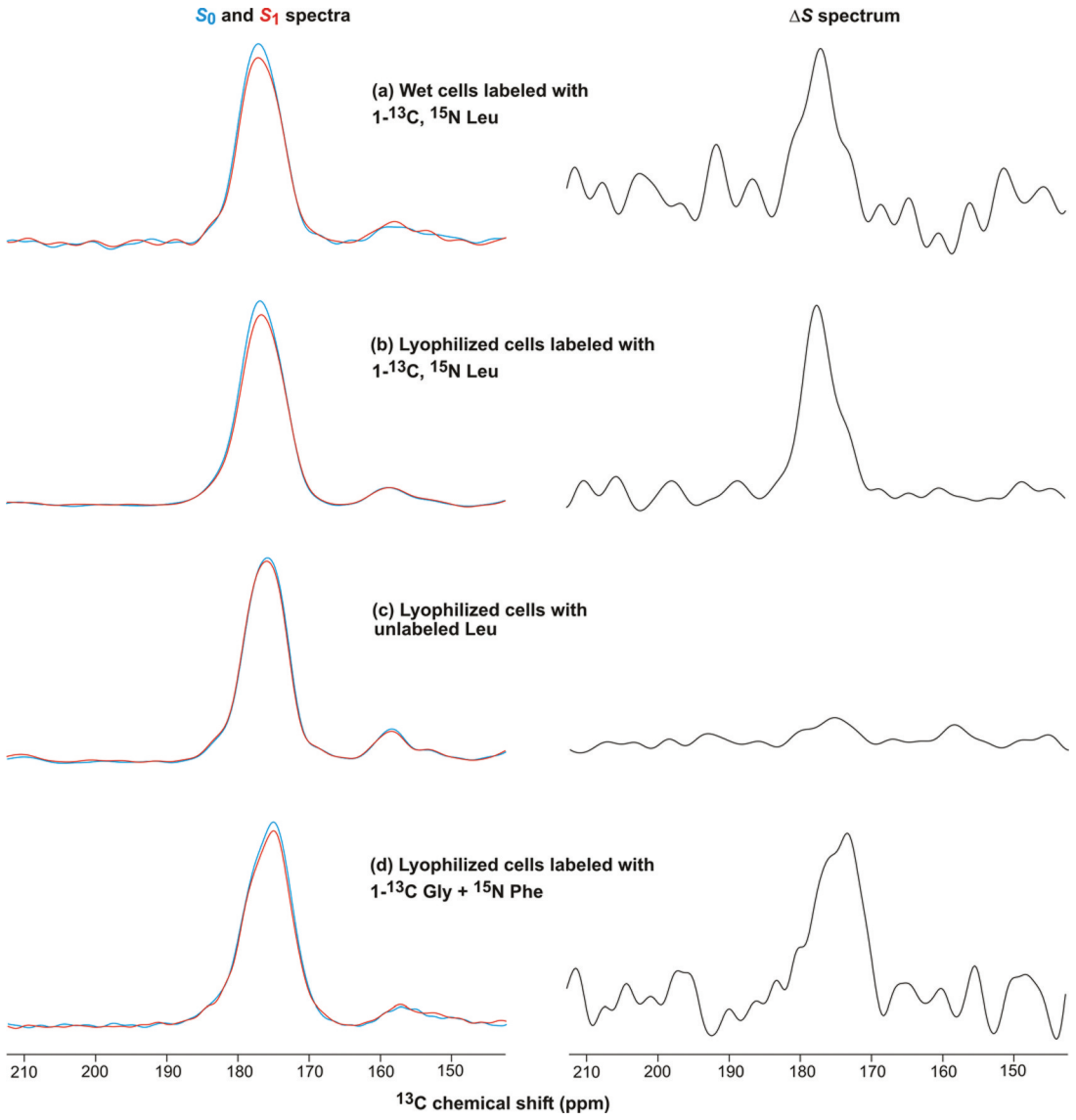


Figure 2. REDOR ^{13}CO NMR spectra of whole bacterial cells produced by sequential steps: (1) growth in rich medium, (2) growth in minimal medium, (3) addition of labeled or unlabeled amino acid(s), (4) induction of Fgp41 expression, and (5) centrifugation. The induction temperature and duration were either (a–c) 23 °C and ~2 h or (d) 37 °C and ~5 h, respectively. The left panels display S_0 (blue) and S_1 (red) spectra and the right panels the $\Delta S \equiv S_0 - S_1$ spectra. The REDOR dephasing time was either (a–c) 1 or (d) 2 ms. For panels a, b, and d, the dominant contribution to each ΔS spectrum was from residues with labeled ^{13}CO groups that were directly bonded to labeled ^{15}N atoms. These residues were (a and b) L33, L44, L54, L81, L134, and L149 of the LL sequential pairs of Fgp41 and (d) G10 of the G10-F11 unique sequential pair. Each S_0 or S_1 spectrum was processed with 100 Hz Gaussian line broadening, and each ΔS spectrum was processed with either (a and d) 200 or (b and c) 100 Hz line broadening. Polynomial baseline correction (typically fifth order) was applied to each spectrum. Each S_0 or S_1 spectrum was the sum of (a) 100000, (b) 100000, (c) 127222, or (d) 48448 scans.

vesicles resulted in a longer fluorophore–quencher distance and increased fluorescence. The fluorescence increase ΔF_{Fgp41} was compared to the maximal fluorescence increase (ΔF_{max}) obtained after subsequent addition of the Triton X-100 detergent that solubilized the vesicles. Assay parameters included (1) the fluorimeter from Photon Technology International, (2) excitation and emission wavelengths of 465 and 530 nm, respectively, with 4 nm bandwidths, and (3) 1.8 mL of an initial vesicle solution, a 0.2 mL aliquot of Fgp41, and an ~20 μL aliquot of 10% Triton X-100.

Membrane Reconstitution. A homogeneous mixture of the POPC (28 mg) and POPG (7 mg) lipids and the bTOG (140 mg) detergent was made by (1) dissolution in chloroform, (2) removal of chloroform via nitrogen gas and overnight vacuum, and (3) dissolution in HEPES/MES buffer. Fgp41 (~10 mg) was added to the solution and had been in affinity column eluents from which excess SDS had been removed. Dialysis of the bTOG/lipid/Fgp41 solution against HEPES/MES buffer removed bTOG with consequent liposome formation with bound Fgp41. Dialysis parameters included (1) a bTOG/lipid/Fgp41 solution in 10 kDa molecular mass cutoff tubing (initial volume of ~15 mL), (2) a 3 L buffer volume, and (3) a 3 day duration with one buffer change. The proteoliposome pellet was harvested after centrifugation (50000*g* for 3 h), and unbound Fgp41 did not pellet under these conditions. The pellet was packed into a 4 mm diameter magic angle spinning (MAS) rotor with ~5 mg of Fgp41 and ~20 mg of total lipid in the 40 μL active sample volume.

dx.doi.org/10.1021/bi201292e | *Biochemistry* 2011, 50, 10013–10026

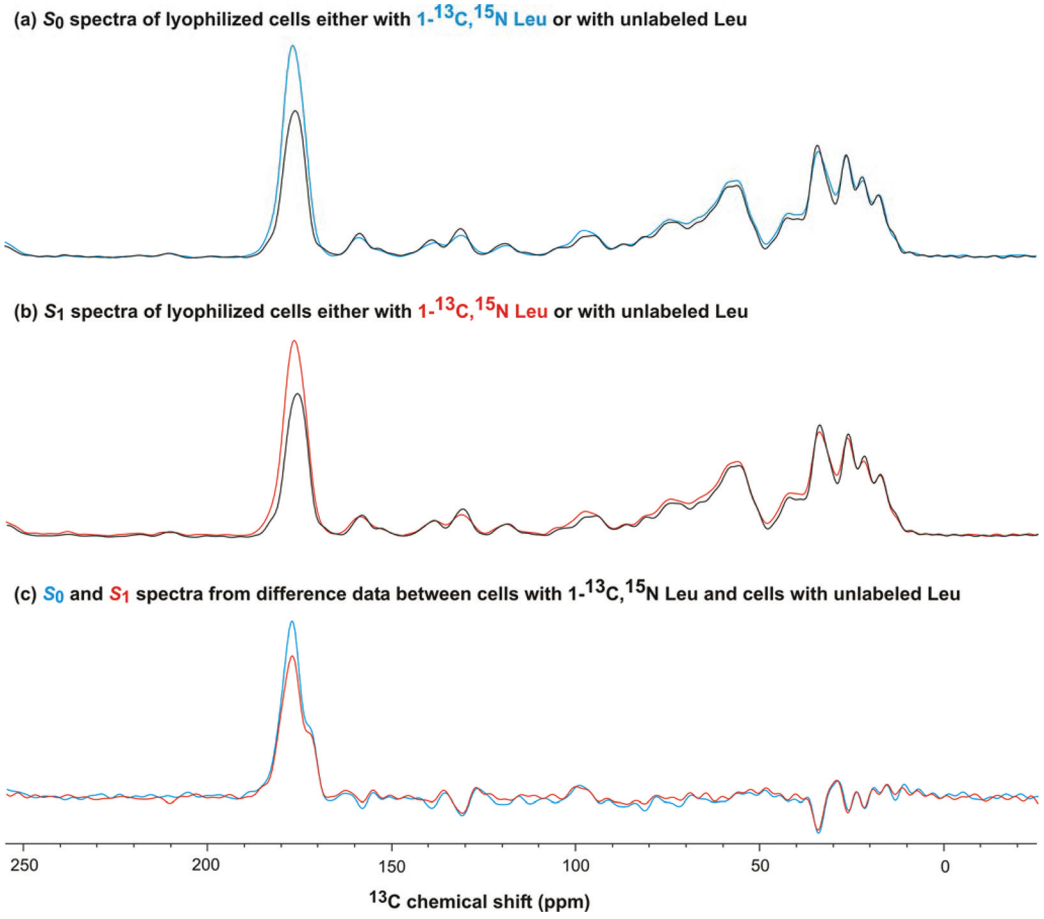


Figure 3. REDOR ^{13}C NMR spectra of lyophilized whole bacterial cells produced with either [1-^{13}C,^{15}N]Leu or unlabeled Leu. The cell production and NMR parameters are described in the legend of Figure 2b,c. Panel a displays the S_0 spectra of the (blue) labeled and (black) unlabeled cells with the relative intensities adjusted to yield the best agreement in the 0−90 ppm region. The spectral intensity in this region should not be affected by the labeling. The incorporation of the labeled Leu into protein synthesized during the induction period is evidenced by the larger ^{13}CO intensity for the labeled cell spectrum. Panel b displays the S_1 spectra of the (red) labeled and (black) unlabeled cells. Panel c displays the S_0 (blue) and S_1 (red) spectra processed from the difference NMR data: labeled cells − 0.75 × unlabeled cells. The 0.75 factor reflects the ratio of the number of scans summed for the labeled cells relative to the number for the unlabeled cells and resulted in a minimal signal in the 0−90 ppm region. The spectra in panel c are representative of the [1-^{13}C,^{15}N]Leu incorporated into the cellular protein.

Solid-State NMR Spectroscopy. Data were obtained with a 9.4 T instrument (Agilent Infinity Plus) and a triple-resonance MAS probe whose rotor was cooled with nitrogen gas at −10 °C. Because of heating from MAS and rf radiation, we expect that water in the sample was liquid rather than solid. Experimental parameters included (1) an 8.0 kHz MAS frequency, (2) a 5 μs ^{1}H $\pi/2$ pulse and 2 ms cross-polarization time with a 50 kHz ^{1}H field and a 70−80 kHz ramped ^{13}C field, (3) a 1−2 ms rotational-echo double-resonance (REDOR) dephasing time with a 9 μs ^{13}C π-pulse at the end of each rotor period except the last period and, for some data, a 12 μs ^{15}N π-pulse at the center of each rotor period, (4) ^{13}C detection with 90 kHz two-pulse phase modulation ^{1}H decoupling (which was also on during the dephasing time), and (5) a 0.8 s pulse delay.[39] Data were acquired without (S_0) and with (S_1) ^{15}N π-pulses during the dephasing time and represented the full ^{13}C signal and the signal of ^{13}C atoms not directly bonded to ^{15}N nuclei, respectively. The $S_0 - S_1$ (ΔS) difference signal was therefore dominated by the labeled ^{13}CO groups in the sequential pairs targeted by the labeling. Spectra were externally referenced to the methylene carbon of adamantane at 40.5 ppm, so that the ^{13}CO shifts could be directly compared to those of soluble proteins.[40]

RESULTS

Detection of Fgp41 in Whole Cells by Solid-State NMR. The bacterial growth and Fgp41 expression conditions were very similar to those used for FHA2, a construct corresponding to the full-length ectodomain (including the fusion peptide) of the influenza virus HA2 fusion protein.[34] Like Fgp41, FHA2 had a C-terminal hexahistidine tag, and bacterial cell lysis and protein solubilization in buffer containing *N*-lauroylsarcosine detergent followed by affinity chromatography resulted in 10 mg of purified FHA2/L of culture. In some contrast, application of this protocol to cells induced to synthesize Fgp41 gave only 0.1 mg of Fgp41/L of culture. It was unclear whether the poor yield was due to a low level of Fgp41 expression or to poor Fgp41 solubilization by the detergent.

The FHA2 solubilized by the detergent was likely initially associated with the cell membrane. It was also shown that a much larger fraction of FHA2 was not solubilized by detergent and was likely constituted in inclusion bodies. The poor yield of Fgp41 might therefore be due to dominant incorporation in inclusion bodies. The molecular structure of FHA2 in inclusion bodies had been probed by (1) adding specific ^{13}CO- and ^{15}N-

labeled amino acids immediately prior to induction, (2) recording REDOR SSNMR spectra of the whole cells after induction so that the filtered ΔS signal corresponded to the ^{13}CO group of a targeted residue in FHA2, and (3) correlation of the experimental peak ^{13}CO shift with the local conformation at this residue.[41] A modified approach was applied to cells induced to express Fgp41 with the goal of assessing Fgp41 production. Addition of 10 mg of [1-^{13}C,^{15}N]Leu to a 50 mL culture just prior to induction of expression targeted the 24 Leu residues and six LL repeats in the Fgp41 sequence (Figure 1a). The isotropic ^{13}CO regions of the REDOR S_0, S_1, and ΔS spectra of the cells are displayed in Figure 2a,b. The Figure 2a sample was an aliquot of the wet cell pellet obtained after the induction period and subsequent centrifugation, and the Figure 2b sample was an aliquot of this whole cell pellet that had been lyophilized. The spectra were similar for both wet and lyophilized cells with ~4 times greater signal per scan in the lyophilized cell sample because this sample had a higher fraction of nonaqueous cell mass. For either sample type, the intensity of the S_1 spectrum was reduced relative to S_0. This supported the presence of LL repeats in the protein produced during the induction period and correlated with the six LL repeats in the Fgp41 sequence. The ΔS spectra had prominent signals in the ^{13}CO region, and these were the only signals that could be detected above the noise (see the Supporting Information for full spectra). Control cells were produced using unlabeled rather than labeled Leu. The resultant NMR spectra are displayed in Figure 2c and had comparable S_0 and S_1 intensities with little ^{13}CO ΔS signal. This provided further support for the idea that the ^{13}CO ΔS signal from the labeled cells could be ascribed to LL repeats in protein produced during expression. Cells were also labeled with [1-^{13}C]Gly and [^{15}N]Phe that targeted the 11 Gly residues in the Fgp41 sequence and the single GF pair at G10 and F11. The resulting NMR spectra are displayed in Figure 2d and included a prominent ^{13}CO ΔS signal that was consistent with Fgp41 production.

Both labeled and natural abundance ^{13}CO groups contribute to the S_0 and S_1 NMR signals of the labeled whole cells. Panels a and b of Figure 3 provide quantitative assessment of these two contributions and show the full S_0 and S_1 spectra, respectively, of the Leu-labeled and unlabeled cells. In each panel, the two spectra were scaled to have equal intensity in the 0−90 ppm region because this region should be unaffected by labeling. The ratio of the unlabeled to labeled scaling factors was ~0.75 and matched the ratio of numbers of scans summed for the labeled versus unlabeled samples. This matching was expected because the signal intensities of individual scans were approximately equal to each other, so the sum signal intensity increased linearly with the number of scans. For panels a and b, the difference between the intensities in the ^{13}CO region was the labeled Leu contribution to the S_0 and S_1 signal, respectively. For these labeled Leu ^{13}CO groups, there was a smaller S_1 intensity relative to that of S_0. This is shown more clearly in Figure 3c, which displays the S_0 and S_1 spectra processed from labeled cell data − 0.75 × unlabeled cell data. For labeled Leu in the cells, the normalized experimental dephasing $(\Delta S/S_0)^{exp}$ equals 0.13 ± 0.01 and was determined from the ^{13}CO S_0 and S_1 intensities in panel c.

The following model and analysis support the idea that most of the labeled Leu was in Fgp41; i.e., Fgp41 was the dominant protein produced during expression. Consider the model. (1) The 24 Leu residues of Fgp41 are labeled with ^{13}CO and ^{15}N. (2) The ^{13}CO groups of the N-terminal Leu residues of the six LL repeats (directly bonded to ^{15}N) have an S_1/S_0 intensity ratio of 0.3. (3) The other 18 Leu ^{13}CO groups have an S_1/S_0 ratio of 1.0. Points 2 and 3 are based on earlier experiments and simulations.[42] For the Fgp41 Leu ^{13}CO groups, $(\Delta S/S_0)^{calc} = [24 - 18 - (6)(0.3)]/24 = 0.17$, which is close to $(\Delta S/S_0)^{exp}$ and supports dominant production of Fgp41 during the induction period. Additional evidence of the production of Fgp41 from SSNMR spectra is described later in the paper and resulted in an estimate of the ratio of the mass of Fgp41 to the mass of lyophilized cells.

Optimization of Fgp41 Solubilization. Initial cell lysis buffers contained either 8 M urea, 0.5% *N*-lauroylsarcosine, 0.5% Triton X-100, or 10% SDS (see Materials and Methods for other buffer components). The solubilization efficiency of the buffer was assessed using detection of a band at ~19 kDa (assigned to Fgp41) in the SDS−PAGE gel of the soluble lysate and then consideration of the absolute intensity of this band as well as its intensity relative to those of other bands in the gel lane. A dark Fgp41 band that was intense relative to other proteins was observed with lysis in buffer containing SDS (see the Supporting Information for a representative gel). Bands that may be Fgp41 were also apparent for lyses in either urea or *N*-lauroylsarcosine; however, purifications of these lysates consistently yielded <1 mg of Fgp41/L of culture, whereas purifications of SDS lysates yielded >1 mg of Fgp41/L of culture. Subsequent lyses were therefore conducted with SDS. The effect of SDS concentration on Fgp41 solubilization was further investigated by comparison of lysis in buffer containing either 0.5, 1, 3, or 5% SDS. For 1%, a dark band that was intense relative to other proteins was observed (see the Supporting Information). Subsequent lyses were conducted using 1% SDS. The effect of different sonication conditions during lysis on Fgp41 solubilization was also investigated. The darkest Fgp41 band was observed using four 1 min cycles at 80% amplitude with 0.8 s on and 0.2 s off. Increasing the number of cycles did not result in a darker band.

Optimization of Fgp41 Expression. Experiments were conducted to investigate the effects on Fgp41 expression of (1) the glycerol concentration in the expression medium, (2) the IPTG concentration, and (3) the induction time. The protocol included (1) overnight cell growth at 37 °C from a glycerol stock in 2 L of LB, (2) cell pelleting by centrifugation followed by resuspension in 1 L of LB, (3) growth at 37 °C for 1 h, (4) transferring 100 mL aliquots of medium into separate flasks, (5) addition of glycerol and then IPTG with the concomitant induction of expression at 23 °C, (6) cell pelleting by centrifugation followed by lysis in buffer with 1% SDS, and (7) SDS−PAGE of the soluble cell lysates with visual comparison of their Fgp41 band intensities. In general, only one parameter, e.g., IPTG concentration, was varied among a group of aliquots. Results were as follows. (1) Comparison among 0.2, 1.0, or 2.0 mM IPTG showed the darkest band at 2.0 mM. (2) Comparison among 0.1, 0.25, and 0.5% (v/v) glycerol showed the darkest bands for 0.1 and 0.25%. (3) Comparison among induction times of 2, 4, and 6 h showed the darkest band for 6 h. Subsequent experiments were conducted using 2 mM IPTG, 0.25% glycerol, and a 6 h induction.

Optimization of Fgp41 Purification. The basis for the development of the Fgp41 purification protocol was an earlier protocol developed in our lab for FHA2.[34] The Fgp41 band was observed with a modified protocol using buffers that contained 50 mM sodium phosphate (pH 8.0), 0.5% SDS, 300

mM NaCl, and imidazole at different concentrations. Relative to only washing with buffer containing 20 mM imidazole, SDS–PAGE showed that sequential washes with buffers containing 1, 20, and 50 mM imidazole was more effective at washing non-Fgp41 proteins from the resin while leaving most Fgp41 bound to the resin. After the washes, the Fgp41 was eluted from the resin using buffer containing 250 mM imidazole. The eluent was incubated overnight at 4 °C with consequent precipitation of excess SDS. Negligible Fgp41 precipitated as evidenced by very similar A_{280} values for the eluent before and after incubation. SDS–PAGE showed that the eluent contained highly pure Fgp41 (Figure 4a) and that

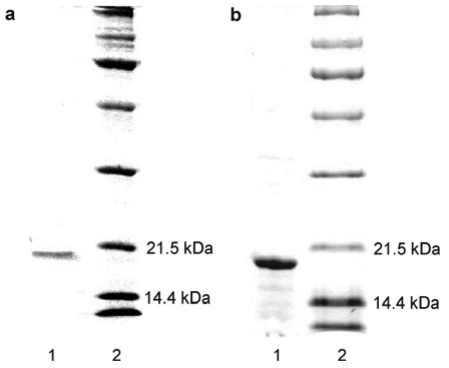


Figure 4. (a) SDS–PAGE gel of (lane 1) an aliquot of the purification eluent with 250 mM imidazole and (lane 2) molecular mass standards. (b) SDS–PAGE gel of (lane 1) an aliquot of the proteoliposome complexes formed during membrane reconstitution of Fgp41 and (lane 2) molecular mass standards. The samples were boiled prior to being loaded onto the gel.

Fgp41 could be membrane-reconstituted (Figure 4b). The final purified yield of Fgp41 (as determined by A_{280}) was ~5 mg/L of culture. This yield was obtained using a 1 h initial mixing of the lysate and resin with a similar yield obtained for a 2 h mixing and a reduced yield 3 mg/L for a 4 h mixing. Enhanced proteolysis is one explanation for a reduced yield with a longer mixing time.

CD Spectroscopy. Figure 5 (black trace) displays the CD spectrum of purified Fgp41 after dialysis into HEPES/MES buffer (pH 7.4). Minima near 208 and 222 nm were diagnostic of α-helical conformation as might be expected from the hairpin structure (Figure 1). The magnitude of θ_{222} showed a small linear decrease over the range of 25–100 °C (Figure 6), where $|\theta_{222}|^{100\text{oC}} \approx 0.8|\theta_{222}|^{25\text{oC}}$. The CD spectra at 25 °C were very similar before and after heating (Figure 5) and showed that the temperature-dependent changes were reversible. This behavior was very similar to the temperature dependencies of the CD spectra of the shorter hairpin and FP-hairpin constructs whose sequence was from the laboratory HXB2 strain of HIV-1.[11,12] For these constructs, 46 contiguous residues including the native loop were replaced with six non-native residues. Subsequent differential scanning calorimetry experiments showed an unfolding transition centered at 110 °C for both constructs. Consideration of other CD measurements on this unfolded state indicates that for Fgp41, $|\theta_{222}|^{\text{unfolded}} \approx 0.2|\theta_{222}|^{25\text{oC}}$ so even at 100 °C, Fgp41 appears to retain hyperthermostable hairpin structure.

Vesicle Fusion. Fgp41 induced negligible intervesicle fusion at pH 7.5 as assayed by lipid mixing (Figure 7). The fluorescence increase was ~2% of that observed for Triton X-

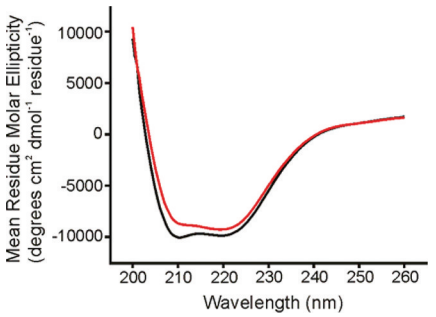


Figure 5. CD spectra of Fgp41 at 25 °C. The black trace is for a sample that had not been heated, and the red trace was obtained after the sample had been heated to 100 °C with subsequent cooling to ambient temperature. Each trace is the difference between the CD spectrum of Fgp41 with buffer and the spectrum of buffer alone. The similarity of the red and black spectra supports the reversibility of any thermal denaturation of Fgp41. Fgp41 samples were prepared by precipitation of excess SDS, subsequent dialysis in HEPES/MES buffer (pH 7.4), and addition of DTT at 2 times the molar concentration of Fgp41 to inhibit disulfide bond formation. For these spectra, the Fgp41 concentration was 20 μM. Spectra for other Fgp41 samples were similar with minima near 208 and 222 nm that were diagnostic of α-helical structure. In some spectra, the θ_{222} could be as low as -15000 deg cm^2 dmol^{-1}.

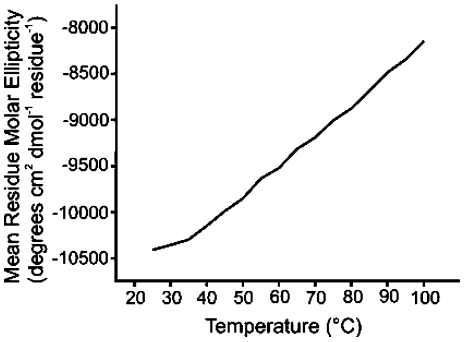


Figure 6. Plot of CD θ_{222} vs temperature for Fgp41. No unfolding transition is apparent for temperatures up to 100 °C. Sample conditions were the same as those described in the legend of Figure 5.

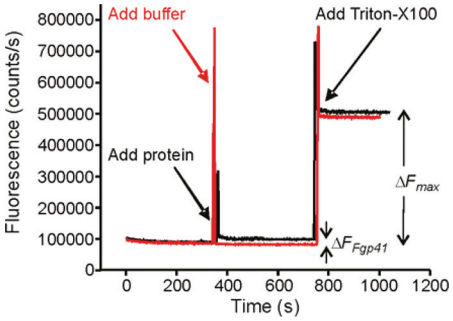


Figure 7. Vesicle fusion assayed by fluorescence. An aliquot of either Fgp41 with buffer (black trace) or buffer alone (red trace) was added to a vesicle solution at 350 s. Fgp41-induced vesicle fusion was evidenced by the fluorescence increase (ΔF_{Fgp41}) of the black trace. In either trace, Triton X-100 detergent was added at 750 s and solubilized the vesicles, with a resultant maximal fluorescence and fluorescence increase (ΔF_{max}). The spikes at 350 and 750 s were artifacts caused by transient exposure to stray light. Assay parameters included vesicles with a 4:1 POPC:POPG composition, a Fgp41:total lipid molar ratio of 0.02, pH 7.5, and 37 °C.

dx.doi.org/10.1021/bi201292e | *Biochemistry* 2011, 50, 10013–10026

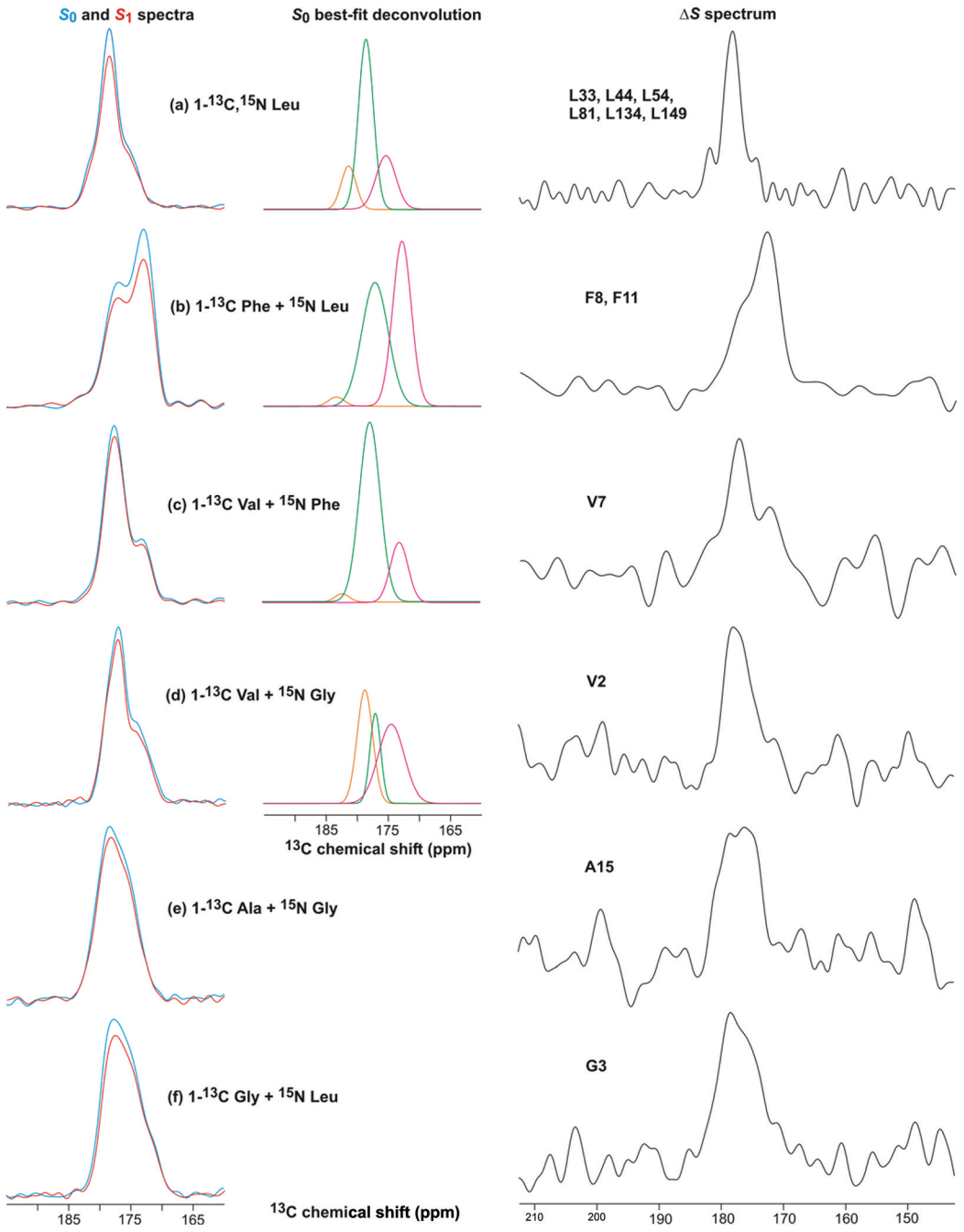


Figure 8. REDOR ^{13}CO NMR spectra of Fgp41 reconstituted in membranes. The labeled amino acids in the expression medium are shown. The left panels display S_0 (blue) and S_1 (red) spectra; the middle panels display the best-fit Gaussian deconvolutions of the S_0 spectra, and the right panels display the $\Delta S \equiv S_0 - S_1$ spectra. The REDOR dephasing time was either (a) 1 or (b–f) 2 ms, and the dominant contribution to each ΔS spectrum was from residues with labeled ^{13}CO groups that were directly bonded to labeled ^{15}N atoms. These residues were (a) L33, L44, L54, L81, L134, and L149, (b) F8 and F11, (c) V7, (d) V2, (e) A15, and (f) G3. Each S_0 or S_1 spectrum was processed with 100 Hz Gaussian line broadening, and each ΔS spectrum was processed with (a and b) 100 or (c–f) 200 Hz line broadening. Polynomial baseline correction (typically fifth order) was applied to each spectrum. Each S_0 or S_1 spectrum was the sum of (a) 93424, (b) 115610, (c) 109504, (d) 110736, (e) 165216, or (f) 103717 scans.

100 detergent, where Triton is commonly considered to induce 100% lipid mixing.

SSNMR of Membrane-Reconstituted Fgp41. Figure 8 displays S_0, S_1, and ΔS REDOR SSNMR spectra of membrane-reconstituted Fgp41 labeled with different amino acids. Many of these spectra were deconvolved into a few Gaussian line shapes (see Figure 8 and the Supporting Information). Table 1 presents the best-fit peak chemical shifts, line widths, and integrated intensities of the individual line shapes of the S_0 spectra, and Table 2 presents the line shape parameters of the ΔS spectra. All fits were excellent as judged by the close agreement between the line shape sum and the experimental intensity (see the Supporting Information). These fittings were used to understand whether the N-helix and C-helix structures of the six-helix bundle were retained in the membrane-associated Fgp41 and to assess the distribution of conformations in the FP region.

Table 1. Analysis and Deconvolution of S_0 and S_1 Solid-State NMR Spectra of Membrane-Reconstituted Fgp41

Fgp41 labeling	$(\Delta S/S_0)^{exp}$ (integrated)[c]	S_0 spectral deconvolution[a]				fraction of calculated S_0 intensity[b]				$(\Delta S/S_0)^{calc}$ [f]
		peak shift (ppm)[d]		peak width (ppm)[e]	intensity (fraction of total)	labeled	natural abundance Fgp41	natural abundance lipid	labeled in N- and C-helices	
[1-^{13}C,^{15}N]Leu	0.12	181.3	helix	2.8	0.15	0.86	0.07	0.07	0.68	0.15
		178.5		2.8	0.60					
		175.3	β	3.7	0.25					
[1-^{13}C]Phe with [^{15}N]Leu	0.15	183.2	helix	3.0	0.02	0.42	0.31	0.27	0	0.24
		177.1		5.0	0.51					
		172.7	β	3.5	0.47					
[1-^{13}C]Val with [^{15}N]Phe	0.07	182.1	helix	2.7	0.02	0.69	0.19	0.12	0.43	0.07
		177.7		3.9	0.76					
		173.1	β	3.3	0.22					
[1-^{13}C]Val with [^{15}N]Gly	0.08	178.6	helix	2.9	0.37	0.64	0.17	0.19	0.40	0.07
		177.0		2.0	0.20					
		174.4	β	4.9	0.43					
[1-^{13}C]Ala with [^{15}N]Gly	0.08					0.79	0.12	0.09	0.34	0.05
[1-^{13}C]Gly with [^{15}N]Leu	0.11					0.77	0.15	0.08	0.14	0.06

[a]Spectral deconvolution was conducted with three Gaussian line shapes whose peak shifts, line widths, and intensities were independently varied until there was minimal difference between the sum of the line shapes and the experimental line shape. For all cases, there was excellent agreement between the best-fit deconvolution sum line shape and the experimental line shape (see the Supporting Information). Deconvolution was not meaningful for the [1-^{13}C]Ala and [1-^{13}C]Gly samples because the S_0 spectra were broad and relatively featureless and the deconvolutions were dominated by a line shape with an ~7 ppm line width. [b]Spectral intensities were calculated with the following considerations: (1) 100% labeling of the Fgp41 residues corresponding to the labeled amino acid(s) with no scrambling to other amino acid types, (2) 1.0 relative intensity for each labeled ^{13}CO, (3) 0.011 relative intensity for each natural abundance ^{13}CO, (4) the Fgp41 natural abundance signal being the sum from backbone ^{13}CO groups and Asn, Asp, Gln, and Glu side chain ^{13}CO groups, and (5) the lipid natural abundance signal calculated using the experimental Fgp41:total lipid molar ratios. The specific ratio in each sample was as follows: [1-^{13}C,^{15}N]Leu, 0.011; [1-^{13}C]Phe with [^{15}N]Leu, 0.012; [1-^{13}C]Val with [^{15}N]Phe, 0.016; [1-^{13}C]Val with [^{15}N]Gly, 0.009; [1-^{13}C]Ala with [^{15}N]Gly, 0.013; and [1-^{13}C]Gly with [^{15}N]Leu, 0.019. The labeled ^{13}CO fraction in N- and C-helices was based on the red and green regions in Figure 1a. [c]The typical uncertainty of $(\Delta S/S_0)^{exp}$ was ±0.02 as determined from the standard deviation of integrals of regions of the S_0 and S_1 spectra that contained noise rather than signal. [d]The reasons for assignment of peaks to specific conformations are provided in the text. [e]Full width at half-maximal line width. [f]$(\Delta S/S_0)^{calc}$ values were based on (1) the fraction of the S_0 signal from labeled ^{13}CO groups directly bonded to labeled ^{15}N atoms and (2) an S_1/S_0 intensity ratio for these ^{13}CO groups of 0.70 (1 ms dephasing time) or 0.85 (2 ms dephasing time). These ratios were based on experimental REDOR data of crystalline amino acid samples as well as simulations. The 1 ms dephasing time was used for the Fgp41 sample labeled with [1-^{13}C,^{15}N]Leu and the 2 ms dephasing time for the other Fgp41 samples.

dx.doi.org/10.1021/bi201292e|*Biochemistry* 2011, 50, 10013–10026

Table 2. Deconvolution of ΔS Spectra of Membrane-Reconstituted Fgp41[a]

Fgp41 labeling	peak shift (ppm)[b]		peak width (ppm)[c]	intensity (fraction of total)
[1-^{13}C,^{15}N]Leu	182.1	helix	1.4	0.10
	178.3		3.1	0.82
	174.6	β	1.5	0.08
[1-^{13}C]Phe with [^{15}N]Leu	176.9	helix	4.9	0.37
	172.6	β	4.3	0.63
[1-^{13}C]Val with [^{15}N]Phe	181.8	helix	3.3	0.12
	177.4		3.9	0.57
	172.2	β	4.2	0.31

[a]Spectral deconvolution was conducted with two or three Gaussian line shapes whose peak shifts, line widths, and intensities were independently varied until there was minimal difference between the sum of the line shapes and the experimental line shape. For all cases, there was excellent agreement between the best-fit deconvolution sum line shape and the experimental line shape (see the Supporting Information). Deconvolution was not meaningful for the other samples because the ΔS spectra were broad and relatively featureless. [b]The reasons for assignment of peaks to specific conformations are provided in the text. [c]Full width at half-maximal line width.

Figure 8a displays the ^{13}CO spectra of the [1-^{13}C,^{15}N]Leu-labeled sample. The S_0 spectrum targeted the 24 Leu residues in the Fgp41 sequence, and the ΔS spectrum targeted the L33, L44, L54, L81, L134, and L149 ^{13}CO groups, which are the N-terminal Leu residues in LL repeats. The ^{13}CO signal was the only discernible feature in the ΔS spectrum (see the Supporting Information). Both the S_0 and ΔS spectra had high signal-to-noise ratios and were fitted well to the sum of three components. In both cases, the two higher shift components comprised >75% of the integrated intensity and were assigned to the helical conformation because their peak shifts were much closer to the characteristic shifts of helical Leu residues (Gaussian distribution of 178.5 ± 1.3 ppm) than to β-strand Leu residues (175.7 ± 1.5 ppm).[43] The ^{13}CO S_0 spectrum had contributions from the labeled Fgp41 Leu residues, as well as natural abundance sites in Fgp41 and lipids. Calculated relative fractional contributions are listed in Table 1 and show that the Fgp41 Leu residues dominate the spectrum. Using an S_1/S_0 intensity ratio of 0.3 for the N-terminal Leu residues of the LL pairs and a ratio of 1.0 for other ^{13}CO groups (based on model compound studies and simulations), the $(\Delta S/S_0)^{calc}$ for the sample was 0.15 and correlated reasonably well with the $(\Delta S/S_0)^{exp}$ of 0.12 ± 0.02.[42]

If the SHB structure were retained in membrane-associated Fgp41, then the fractional contribution to the S_0 ^{13}CO intensity of Leu residues in the N- and C-helices would be 0.68. This correlated well with the experimental fractional S_0 intensity of 0.75 in the helical conformation and supports retention of SHB structure upon membrane binding. Further support for this structure was the correlation between the experimental helical fractional intensity of 0.92 in the ΔS spectrum and the location of the six LL repeats in the N- and C-helices.

Spectra of the remaining labeled samples provided information about structure in the putative SHB region as well as in the FP. Figure 8b displays spectra from a sample with [1-^{13}C]Phe and [^{15}N]Leu labeling. There are three Phe residues in the sequence: F8 and F11 in the FP and F96, which would be in the loop region of a SHB structure. There was ∼0.4 fractional contribution of the labeled Phe ^{13}CO groups to the S_0 spectrum and ∼0.3 contribution each from natural abundance ^{13}CO groups in Fgp41 and the lipid. The S_0 spectrum was well-fitted to the sum of three line shapes. The two line shapes with higher peak shifts comprised ∼0.5 fractional contribution of the total intensity, and the shifts were generally consistent with helical protein conformation. The peak shift of the other line shape was consistent with β-strand protein conformation and with lipid shifts. The labeled F8 and F11 ^{13}CO groups in the FP were directly bonded to labeled Leu ^{15}N atoms with an S_1/S_0 of ∼0.15 for a 2 ms dephasing time.[42] The other ^{13}CO groups had S_1/S_0 values of ∼1. $(\Delta S/S_0)^{calc}$ was close to $(\Delta S/S_0)^{exp}$, and the ΔS spectrum was dominated by the F8 and F11 ^{13}CO signals. The ΔS spectrum was well-fitted to two line shapes with the higher (lower) peak shifts consistent with helical (β-strand) Phe ^{13}CO shift distributions of 177.1 ± 1.4 ppm (174.3 ± 1.6 ppm).[43] The lower ∼173 ppm experimental peak shift matched well with the 173 ppm peak shifts measured for F8 and F11 of the membrane-associated HFP fragment.[26−28] This peptide has been shown to form small oligomers with antiparallel β-sheet structure.[30] For membrane-bound Fgp41, the ratio for F8 and F11 of helical to β-strand/sheet intensities was ∼1:2 and was consistent with two Fgp41 populations with different FP conformations.

Figure 8c displays the spectra and analysis for a sample labeled with [1-^{13}C]Val and [^{15}N]Phe. The approach to the analysis was the same as in the previous paragraph. The eight labeled Val ^{13}CO groups made a fractional contribution of ∼0.7 to the S_0 signal. The S_0 spectrum was well-fitted to three line shapes, and the two higher-shift line shapes comprised ∼0.8 fraction of the total intensity and had shifts that correlated with the helical rather than the β-strand Val ^{13}CO distribution (177.7 ± 1.4 ppm vs 174.8 ± 1.4 ppm).[43] The line shape with the lowest peak shift correlated with β-strand/sheet conformation. The high helical content was consistent with SHB structure for membrane-bound Fgp41. The $(\Delta S/S_0)^{calc}$ matched $(\Delta S/S_0)^{exp}$. The ΔS spectrum was dominated by V7 and was well-fitted to three line shapes that indicated a ratio of helical to β-strand/sheet populations of ∼2:1. This ΔS spectrum confirmed two Fgp41 populations with different FP conformations, while the difference in population ratio relative to the Figure 8b ΔS spectrum may reflect the lower signal-to-noise ratio of the Figure 8c spectrum, sample-to-sample variation, and/or conformational differences between V7 and F8 and F11.

Figure 8d displays the spectra and analysis for a sample labeled with [1-^{13}C]Val and [^{15}N]Gly. As with Figure 8c, analysis of the S_0 spectrum of Figure 8d supported a dominant helical conformation consistent with a six-helix bundle structure. Comparison of the two spectra provided insight into sample-to-sample variation and the robustness of the S_0 deconvolution. The $(\Delta S/S_0)^{calc}$ matched $(\Delta S/S_0)^{exp}$. The ΔS spectrum was dominated by V2 and extended broadly over the 170−180 ppm region, so that deconvolution was not meaningful. As noted in the previous paragraph, this shift range includes the helical and β-strand/sheet shift distributions, and the ΔS spectrum was therefore consistent with a mixture of Fgp41 populations with helical and β-strand/sheet conformations at V2 in the FP. We note that the V2 ^{13}CO signal of the membrane-associated HFP was also broader than signals from residues 6−12 in the interior hydrophobic region.[26]

Table 3. Deconvolution of Spectra of Lyophilized Cells Induced To Produce Fgp41[a]

sample/spectrum	spectrum type	peak shift (ppm)[b]		peak width (ppm)[c]	intensity (fraction of total)
[1-^{13}C,^{15}N]Leu cells	ΔS	182.1	helix	2.2	0.03
		177.4		4.5	0.82
		173.1	β	3.2	0.15
[1-^{13}C,^{15}N]Leu cells − 0.75 × unlabeled cells	S_0	180.8	helix	8.9	0.18
		177.6		5.6	0.69
		172.1	β	3.1	0.13

[a]Spectral deconvolution was conducted with three Gaussian line shapes whose peak shifts, line widths, and intensities were independently varied until there was minimal difference between the sum of the line shapes and the experimental line shape. For both cases, there was excellent agreement between the best-fit deconvolution sum line shape and experimental line shape (see the Supporting Information). [b]The reasons for the assignment of peaks to specific conformations are provided in the text. [c]Full width at half-maximal line width.

Panels e and f of Figure 8 display spectra from samples that were labeled with [1-^{13}C]Ala and [^{15}N]Gly or [1-^{13}C]Gly and [^{15}N]Leu. The analyses are presented together because of the similar results. The S_0 spectra were broad and featureless over the 170−185 ppm range, so that deconvolution was not meaningful. This spectral breadth was understood by considering that although the fractional contribution of the labeled ^{13}CO groups to the total S_0 intensity was ∼0.8, the labeled contribution from N- and C-helices in a SHB structure would be ∼0.25. Approximately half of the S_0 intensity would be from labeled ^{13}CO groups in the FP and loop regions. The earlier Figure 8a−d analyses supported a mixture of helical and β-strand/sheet shifts for FP ^{13}CO groups, and broad signals are also expected from ^{13}CO groups in the less-ordered loop region. For the spectra in panels e and f of Figure 8, there were relatively good agreements between $(\Delta S/S_0)^{calc}$ and $(\Delta S/S_0)^{exp}$ and the ΔS spectra were dominated by the ^{13}CO groups of A15 and G3, respectively. These ΔS spectra extended over 170−180 ppm, and as with the V2 ΔS spectrum, the breadth correlated with being near one end of the FP region and with the spectral breadth observed for the corresponding residues in the membrane-associated HFP.[26]

Analysis of the SSNMR Spectra of Lyophilized Whole Cells. Comparison of the ΔS spectrum of [1-^{13}C,^{15}N]Leu-labeled cells (Figure 2b) to the ΔS spectrum of the unlabeled cells (Figure 2c) shows a clear effect from using labeled Leu. The "labeled cell difference" S_0 (S_1) spectrum (Figure 3c) is the difference between the S_0 (S_1) spectra of the labeled and unlabeled cells and shows only the contribution of the labeled Leu. Deconvolution was applied to the labeled cell difference S_0 spectrum and to the labeled cell ΔS spectrum. Both spectra were well-fitted to the sum of three Gaussian line shapes (see Table 3 and the Supporting Information) and were dominated by the [1-^{13}C,^{15}N]Leu incorporated into the cell protein produced during the expression period. To assess the fraction of Fgp41 in this protein, we compared the deconvolutions of (1) the ΔS spectrum of labeled cells and the ΔS spectrum of membrane-reconstituted Fgp41 and (2) the S_0 spectrum of labeled cell difference and the S_0 spectrum of membrane-reconstituted Fgp41. For either case, there were striking similarities in the deconvolutions, including the peak chemical shifts and the large fraction of the total intensity in the two high-shift peaks corresponding to the helical conformation. These similarities as well as the detection of large ΔS signals provide additional strong evidence that Fgp41 is the predominant labeled protein in the cells. This result was used to conservatively estimate that there was at least 3 mg of Fgp41 in the lyophilized labeled cell NMR sample. Other inputs for this estimate were as follows. (1) The mass of Fgp41 in the membrane-reconstituted sample was ∼5 mg. (2) The membrane and whole cell data were acquired on the same spectrometer and were the sums of approximately the same numbers of scans. (3) For the membrane-reconstituted and whole cell samples, the integrated ^{13}CO intensities of the ΔS spectra were within 20% agreement and there was a similar agreement for the S_0 spectra. There was ∼50 mg of total cell mass in the whole cell NMR sample, so the ratio of the mass of Fgp41 to the total dry cell mass was ∼0.05. There was ∼2 g of dry cell mass/L of culture, so prior to solubilization and purification, there was ∼100 mg of Fgp41/L of culture. The much smaller purified yield of ∼5 mg of Fgp41/L of culture points to solubilization and purification rather than expression as the limiting factors in Fgp41 production.

Because relatively harsh conditions were needed to solubilize Fgp41 in the cells, it seems likely that most Fgp41 was in inclusion bodies. Detection of the predominant helical conformation for the Leu residues in Fgp41 in the lyophilized cells, including those in the N- and C-helices of a putative SHB structure, suggests that this structure is retained in inclusion bodies.

DISCUSSION

To the best of our knowledge, earlier reports of bacterial production of large quantities of the ectodomain of HIV gp41 that incorporated the FP have always included a N-terminal carrier protein like glutathione *S*-transferase that was subsequently cleaved. This study describes production of Fgp41 that has a much shorter six-residue C-terminal histidine tag rather than a N-terminal carrier protein. The ∼5 mg of purified yield Fgp41/L of culture is approximately half of that of the highest reported yield with a carrier protein.[16]

Initial solubilization and purification attempts gave only 0.1 mg of Fgp41/L of culture, and SSNMR was applied to detect whether this poor yield was due to a very low level of expression. The successful approach included identifying an abundant amino acid in Fgp41 (24 Leu residues) that was the first amino acid of an abundant sequential pair (six LL repeats). The procedure included (1) inducing cells in minimal medium with either [1-^{13}C,^{15}N]Leu or unlabeled Leu, (2) cell pellet lyophilization, and (3) recording ^{13}C REDOR SSNMR spectra of the lyophilized whole cells with a short dephasing time. As expected, the spectra of the labeled and unlabeled cells were very similar in the aliphatic ^{13}C shift region, but the labeled cells had greater intensity in the ^{13}CO region. The labeled cell −

unlabeled cell difference spectra were therefore assigned to Leu ^{13}CO groups incorporated into protein produced during the expression period. The $(\Delta S/S_0)^{exp}$ of the difference spectra was close to the $(\Delta S/S_0)^{calc}$ estimated for Fgp41 and provided strong evidence that Fgp41 was expressed in large quantities in the bacteria. Calibration with standard spectra led to the estimate of ~100 mg of Fgp41/L of culture. The harsh conditions needed to solubilize Fgp41 in the cells and the ~5 mg of purified Fgp41/L of culture provided evidence that most Fgp41 is in inclusion bodies in the bacteria. This approach to detection of recombinant protein in whole cells by SSNMR has several strengths, including (1) small (~50 mL) culture volumes, (2) small (~10 mg) quantities of isotopically labeled amino acids, and (3) a simple sample preparation protocol without protein solubilization or purification. The main drawback might be the few days of SSNMR spectrometer time. Future studies with other proteins are needed to test the generality of the approach.

Interpretation of the SSNMR spectra using this approach will likely not be greatly affected by some "scrambling", i.e., conversion of the labeled amino acids into other amino acids. For example, transfer of the ^{15}N from the labeled amino acid to other amino acids would likely result in a larger number of labeled ^{13}CO−^{15}N sequential pairs and therefore larger ΔS signals and more sensitive detection of the recombinant protein. Support for minimal scrambling of the Fgp41 sample labeled with [1-^{13}C,^{15}N]Leu included (1) expression at lower temperature for a short 2 h duration, (2) $(\Delta S/S_0)^{exp}$ values for both the whole cell and membrane-reconstituted samples that were close to the values calculated using models without scrambling, and (3) deconvolutions of the S_0 and ΔS ^{13}CO spectra of these samples that agreed nearly quantitatively with the expected secondary structure distributions of the 24 Leu residues and the six N-terminal Leu residues in LL pairs, respectively (see Tables 1−3).[44]

The CD spectra and melting curves of purified Fgp41 support thermostable SHB structure, and this structure was retained upon membrane binding as evidenced by a predominant sharp (3 ppm) helical ^{13}CO feature in the ΔS spectrum of Fgp41 produced with [1-^{13}C,^{15}N]Leu. This feature was assigned to the sum of ^{13}CO signals from six Leu residues that are in N- and C-helices in the SHB structure. The SHB was also observed for the membrane-associated FP-hairpin construct whose sequence was from an HIV clade different from that of Fgp41 and for which 46 contiguous residues including the native loop were replaced with six non-native residues. By contrast, Fgp41 had the full native sequence of its clade. The similar results for Fgp41 and FP-hairpin support the SHB as the final stable structure for membrane-associated gp41 (Figure 1b).

There are two Cys residues in the Fgp41 sequence that are separated by five residues. These Cys residues are likely on either side of the tip of the loop in the hairpin structure and therefore positioned to form an intramolecular disulfide bond.[4] For the laboratory strain HXB2 sequence, these Cys residues are replaced with Ala residues (see the Supporting Information). The unfolding temperature of the HXB2 hairpin structure is 105 °C, which should be within a few degrees of that of Fgp41 (Figure 6).[12,14] It is therefore unlikely that the disulfide bond of Fgp41 contributes appreciably to the thermostability of the hairpin structure of Fgp41.

Fgp41 induced negligible intervesicle lipid mixing at pH 7.5, which correlated with the same result for FP-hairpin. gp41 in the final SHB state may therefore be fusion-inactive at least with respect to lipid mixing that occurs early in either fusion of membranes of HIV and host cells or in gp41-mediated cell−cell fusion. This view is supported by other fusion data showing that most membrane changes occur prior to formation of the final gp41 SHB state.[8] For negatively charged vesicles, FP-hairpin and related SHB gp41 constructs induce lipid mixing at pH values much lower than pH 7 (e.g., pH 4), and the pH-dependent functional difference has been correlated with changes in protein−membrane electrostatics.[17] It is therefore likely that Fgp41 will also induce lipid mixing at these lower pH values. Over the past 25 years, there have been a series of experimental studies by different groups to determine whether HIV infects cells through direct fusion at the plasma membrane or through an endocytic mechanism.[45,46] In our view, the preponderance of data for either route supports HIV−cell fusion at pH ≈7 where SHB gp41 is fusion-inactive. There may be some differences among enveloped viruses as there is significant evidence of fusion activity of folded influenza virus fusion protein ectodomain FHA2.[33,35]

Relative to the sharp 3 ppm ΔS ^{13}CO signal from six Leu residues in the SHB, broader (4−10 ppm) ΔS ^{13}CO signals were observed from (typically) one residue in the FP. These breadths indicate conformational heterogeneity in the FP.[24,25] This point was further supported by the ΔS spectra of V7, F8, and F11, which were reasonably deconvolved into helical and β-sheet signals and indicated two populations of Fgp41 with distinct FP conformations. Helical and β-sheet FP signals were also observed for membrane-associated FP-hairpin samples even though there were differences between the Fgp41 and FP-hairpin samples, including (1) two of the 16 FP residues being different, (2) lipids being ester-linked (Fgp41) versus ether-linked (FP-hairpin), (3) membrane reconstitution being based on detergent dialysis (Fgp41) versus simple mixing of protein and vesicle solutions (FP-hairpin), and (4) unfrozen Fgp41 versus frozen FP-hairpin samples.[12] Detection of helical and β-sheet FP populations in both sample types strongly supports the existence of these populations in membrane-associated gp41 in its final SHB state. In the future, it would be very interesting to study a larger gp41 construct that contains the transmembrane domain and for which there may be close contact between the FP and transmembrane domains.

For most nonbacterial proteins produced in bacteria, a large fraction of the protein in the cells is found in "inclusion bodies", which are macroscopic noncrystalline solid aggregates.[41,47,48] Inclusion body formation appears to be largely independent of protein sequence. There are few data about the structure(s) of recombinant protein molecules in inclusion bodies. In this study, deconvolutions of the S_0 and ΔS spectra of the [1-^{13}C,^{15}N]Leu-labeled inclusion body Fgp41 in cells resulted in line shapes with peak shifts and relative intensities similar to those of membrane-associated Fgp41 with folded SHB structure. It therefore seems likely that at least the SHB fold exists for most Fgp41 molecules in inclusion bodies. This is the second protein for which there is SSNMR evidence of a stable fold in inclusion bodies. This work highlights the potential of SSNMR for probing protein structure and aggregation in inclusion bodies.

MAS SSNMR structural studies of proteins are generally conducted by one of two approaches: (1) uniform ^{13}C and ^{15}N labeling, unambiguous assignment of most cross-peaks in multidimensional NMR spectra, and structural interpretation of the peak shifts and the cross-peak intensities of nuclei far apart

in the sequence or (2) specific (often residue-type or at least amino acid-type) labeling and quantitative SSNMR measurements (e.g., shifts or dipolar couplings) to test specific structural models.[49−51] The choice of approach for a particular protein depends on protein size and quantity as well as NMR line widths. Approach 1 is more feasible for smaller proteins, high protein concentrations, and narrow (<1 ppm) line widths. This study is an example of approach 2, which was appropriate given the 162 residues, the ≈0.01 Fgp41:lipid ratio (with additional dilution of Fgp41 in the sample from water), the 3−10 ppm ^{13}CO line widths, and the possibility of FP conformational heterogeneity (shown to be true in this study). The approach considered a model based on the existing high-resolution SHB structures of gp41 fragments and the extensive residue specific SSNMR data for membrane-associated HFP.

■ ASSOCIATED CONTENT

Supporting Information

Fgp41 amino acid and DNA sequences and comparison to the HXB2 amino acid sequence, additional SDS−PAGE gels, full ΔS spectra of [1-^{13}C,^{15}N]Leu-labeled samples, and spectral deconvolutions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (517) 355-9715. Fax: (517) 353-1793. E-mail: weliky@chemistry.msu.edu.

Funding

The research was supported by National Institutes of Health Grant A147153.

■ ACKNOWLEDGMENTS

Dr. William Wedemeyer is acknowledged for the gp41 cDNA. Dr. Jun Sun is acknowledged for Fgp41 vector construction. Dr. Lisa Lapidus is acknowledged for the CD spectrometer and fluorimeter. Dr. Kelly Sackett is acknowledged for assistance with the lipid mixing assay.

■ ABBREVIATIONS

BOG, *n*-octyl β-D-glucopyranoside; bTOG, *n*-octyl β-D-thioglucopyranoside; C_8E_5, pentaethylene glycol monooctyl ether; CD, circular dichroism; C-helix, C-terminal helix; ΔS, $S_0 - S_1$; $(\Delta S/S_0)^{calc}$, calculated $\Delta S/S_0$; $(\Delta S/S_0)^{exp}$, experimental $\Delta S/S_0$; DTT, dithiothreitol; FP, fusion peptide; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HEPES/MES, 5 mM HEPES and 10 mM MES (pH 7.4); HFP, HIV fusion peptide; HIV-1, human immunodeficiency virus type 1; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria-Bertani broth; LUV, large unilamellar vesicle; MAS, magic angle spinning; MES, 2-(*N*-morpholino)ethanesulfonic acid; MW, molecular weight; N-helix, N-terminal helix; *N*-NBD-PE, *N*-(7-nitro-2,1,3-benzoxadiazol-4-yl)phosphatidylethanolamine; *N*-Rh-PE, *N*-(lissamine Rhodamine B sulfonyl)phosphatidylethanolamine; POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; POPG, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (sodium salt); PHI, prehairpin intermediate; REDOR, rotational-echo double-resonance; SDS, sodium dodecyl sulfate; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SHB, six-helix bundle; SIV, simian immunodeficiency virus; SSNMR, solid-state nuclear magnetic resonance.

■ REFERENCES

(1) White, J. M., Delos, S. E., Brecher, M., and Schornberg, K. (2008) Structures and mechanisms of viral membrane fusion proteins: Multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol. 43*, 189−219.

(2) Melikyan, G. B. (2008) Common principles and intermediates of viral protein-mediated fusion: The HIV-1 paradigm. *Retrovirology 5*, 111.

(3) Chernomordik, L. V., Zimmerberg, J., and Kozlov, M. M. (2006) Membranes of the world unite! *J. Cell Biol. 175*, 201−207.

(4) Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J. 17*, 4572−4584.

(5) Yang, Z. N., Mueser, T. C., Kaufman, J., Stahl, S. J., Wingfield, P. T., and Hyde, C. C. (1999) The crystal structure of the SIV gp41 ectodomain at 1.47 Å resolution. *J. Struct. Biol. 126*, 131−144.

(6) Eckert, D. M., and Kim, P. S. (2001) Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem. 70*, 777−810.

(7) Buzon, V., Natrajan, G., Schibli, D., Campelo, F., Kozlov, M. M., and Weissenhorn, W. (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog. 6*, e1000880.

(8) Markosyan, R. M., Cohen, F. S., and Melikyan, G. B. (2003) HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol. Biol. Cell 14*, 926−938.

(9) Freed, E. O., Delwart, E. L., Buchschacher, G. L. Jr., and Panganiban, A. T. (1992) A mutation in the human immunodeficiency virus type 1 transmembrane glycoprotein gp41 dominantly interferes with fusion and infectivity. *Proc. Natl. Acad. Sci. U.S.A. 89*, 70−74.

(10) Yang, R., Prorok, M., Castellino, F. J., and Weliky, D. P. (2004) A trimeric HIV-1 fusion peptide construct which does not self-associate in aqueous solution and which has 15-fold higher membrane fusion rate. *J. Am. Chem. Soc. 126*, 14722−14723.

(11) Sackett, K., Nethercott, M. J., Shai, Y., and Weliky, D. P. (2009) Hairpin folding of HIV gp41 abrogates lipid mixing function at physiologic pH and inhibits lipid mixing by exposed gp41 constructs. *Biochemistry 48*, 2714−2722.

(12) Sackett, K., Nethercott, M. J., Epand, R. F., Epand, R. M., Kindra, D. R., Shai, Y., and Weliky, D. P. (2010) Comparative analysis of membrane-associated fusion peptide secondary structure and lipid mixing function of HIV gp41 constructs that model the early Pre-Hairpin Intermediate and final Hairpin conformations. *J. Mol. Biol. 397*, 301−315.

(13) Pan, J. H., Lai, C. B., Scott, W. R. P., and Straus, S. K. (2010) Synthetic fusion peptides of tick-borne encephalitis virus as models for membrane fusion. *Biochemistry 49*, 287−296.

(14) Lev, N., Fridmann-Sirkis, Y., Blank, L., Bitler, A., Epand, R. F., Epand, R. M., and Shai, Y. (2009) Conformational stability and membrane interaction of the full-length ectodomain of HIV-1 gp41: Implication for mode of action. *Biochemistry 48*, 3166−3175.

(15) Cheng, S. F., Chien, M. P., Lin, C. H., Chang, C. C., Lin, C. H., Liu, Y. T., and Chang, D. K. (2010) The fusion peptide domain is the primary membrane-inserted region and enhances membrane interaction of the ectodomain of HIV-1 gp41. *Mol. Membr. Biol. 27*, 31−44.

(16) Lin, C. H., Lin, C. H., Chang, C. C., Wei, T. S., Cheng, S. F., Chen, S. S. L., and Chang, D. K. (2011) An efficient production and characterization of HIV-1 gp41 ectodomain with fusion peptide in *Escherichia coli* system. *J. Biotechnol. 153*, 48−55.

(17) Sackett, K., TerBush, A., and Weliky, D. P. (2011) HIV gp41 six-helix bundle constructs induce rapid vesicle fusion at pH 3.5 and little fusion at pH 7.0: Understanding pH dependence of protein aggregation, membrane binding, and electrostatics, and implications for HIV-host cell fusion. *Eur. Biophys. J. 40*, 489−502.

(18) Chang, D. K., Cheng, S. F., and Chien, W. J. (1997) The amino-terminal fusion domain peptide of human immunodeficiency virus

type 1 gp41 inserts into the sodium dodecyl sulfate micelle primarily as a helix with a conserved glycine at the micelle-water interface. *J. Virol. 71*, 6593−6602.
(19) Morris, K. F., Gao, X. F., and Wong, T. C. (2004) The interactions of the HIV gp41 fusion peptides with zwitterionic membrane mimics determined by NMR spectroscopy. *Biochim. Biophys. Acta 1667*, 67−81.
(20) Jaroniec, C. P., Kaufman, J. D., Stahl, S. J., Viard, M., Blumenthal, R., Wingfield, P. T., and Bax, A. (2005) Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry 44*, 16167−16180.
(21) Li, Y. L., and Tamm, L. K. (2007) Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. *Biophys. J. 93*, 876−885.
(22) Gabrys, C. M., and Weliky, D. P. (2007) Chemical shift assignment and structural plasticity of a HIV fusion peptide derivative in dodecylphosphocholine micelles. *Biochim. Biophys. Acta 1768*, 3225−3234.
(23) Pereira, F. B., Goni, F. M., Muga, A., and Nieva, J. L. (1997) Permeabilization and fusion of uncharged lipid vesicles induced by the HIV-1 fusion peptide adopting an extended conformation: Dose and sequence effects. *Biophys. J. 73*, 1977−1986.
(24) Grasnick, D., Sternberg, U., Strandberg, E., Wadhwani, P., and Ulrich, A. S. (2011) Irregular structure of the HIV fusion peptide in membranes demonstrated by solid-state NMR and MD simulations. *Eur. Biophys. J. 40*, 529−543.
(25) Tristram-Nagle, S., Chan, R., Kooijman, E., Uppamoochikkal, P., Qiang, W., Weliky, D. P., and Nagle, J. F. (2010) HIV fusion peptide penetrates, disorders, and softens T-cell membrane mimics. *J. Mol. Biol. 402*, 139−153.
(26) Yang, J., Gabrys, C. M., and Weliky, D. P. (2001) Solid-state nuclear magnetic resonance evidence for an extended β strand conformation of the membrane-bound HIV-1 fusion peptide. *Biochemistry 40*, 8126−8137.
(27) Zheng, Z., Yang, R., Bodner, M. L., and Weliky, D. P. (2006) Conformational flexibility and strand arrangements of the membrane-associated HIV fusion peptide trimer probed by solid-state NMR spectroscopy. *Biochemistry 45*, 12960−12975.
(28) Qiang, W., Bodner, M. L., and Weliky, D. P. (2008) Solid-state NMR spectroscopy of human immunodeficiency virus fusion peptides associated with host-cell-like membranes: 2D correlation spectra and distance measurements support a fully extended conformation and models for specific antiparallel strand registries. *J. Am. Chem. Soc. 130*, 5459−5471.
(29) Qiang, W., and Weliky, D. P. (2009) HIV fusion peptide and its cross-linked oligomers: Efficient syntheses, significance of the trimer in fusion activity, correlation of β strand conformation with membrane cholesterol, and proximity to lipid headgroups. *Biochemistry 48*, 289−301.
(30) Schmick, S. D., and Weliky, D. P. (2010) Major antiparallel and minor parallel β sheet populations detected in the membrane-associated human immunodeficiency virus fusion peptide. *Biochemistry 49*, 10623−10635.
(31) Brugger, B., Glass, B., Haberkant, P., Leibrecht, I., Wieland, F. T., and Krasslich, H. G. (2006) The HIV lipidome: A raft with an unusual composition. *Proc. Natl. Acad. Sci. U.S.A. 103*, 2641−2646.
(32) Sackett, K., Wexler-Cohen, Y., and Shai, Y. (2006) Characterization of the HIV N-terminal fusion peptide-containing region in context of key gp41 fusion conformations. *J. Biol. Chem. 281*, 21755−21762.
(33) Curtis-Fisk, J., Preston, C., Zheng, Z. X., Worden, R. M., and Weliky, D. P. (2007) Solid-state NMR structural measurements on the membrane-associated influenza fusion protein ectodomain. *J. Am. Chem. Soc. 129*, 11320−11321.
(34) Curtis-Fisk, J., Spencer, R. M., and Weliky, D. P. (2008) Isotopically labeled expression in *E. coli*, purification, and refolding of the full ectodomain of the influenza virus membrane fusion protein. *Protein Expression Purif. 61*, 212−219.
(35) Kim, C. S., Epand, R. F., Leikina, E., Epand, R. M., and Chernomordik, L. V. (2011) The final conformation of the complete ectodomain of the HA2 subunit of influenza hemagglutinin can by itself drive low pH-dependent fusion. *J. Biol. Chem. 286*, 13226−13234.
(36) Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wongstaal, F. (1985) Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature 313*, 277−284.
(37) Painter, S. L., Biek, R., Holley, D. C., and Poss, M. (2003) Envelope variants from women recently infected with clade A human immunodeficiency virus type 1 confer distinct phenotypes that are discerned by competition and neutralization experiments. *J. Virol. 77*, 8448−8461.
(38) Pascual, R., Moreno, M. R., and Villalain, J. (2005) A peptide pertaining to the loop segment of human immunodeficiency virus gp41 binds and interacts with model biomembranes: Implications for the fusion mechanism. *J. Virol. 79*, 5142−5152.
(39) Gullion, T., and Schaefer, J. (1989) Rotational-echo double-resonance NMR. *J. Magn. Reson. 81*, 196−200.
(40) Morcombe, C. R., and Zilm, K. W. (2003) Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson. 162*, 479−486.
(41) Curtis-Fisk, J., Spencer, R. M., and Weliky, D. P. (2008) Native conformation at specific residues in recombinant inclusion body protein in whole cells determined with solid-state NMR spectroscopy. *J. Am. Chem. Soc. 130*, 12568−12569.
(42) Yang, J. (2003) Ph.D. Thesis, Michigan State University, East Lansing, MI.
(43) Zhang, H. Y., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR 25*, 173−195.
(44) Tong, K. I., Yamamoto, M., and Tanaka, T. (2008) A simple method for amino acid selective isotope labeling of recombinant proteins in *E. coli*. *J. Biomol. NMR 42*, 59−67.
(45) Grewe, C., Beck, A., and Gelderblom, H. R. (1990) HIV: Early virus-cell interactions. *J. AIDS 3*, 965−74.
(46) Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G. B. (2009) HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. *Cell 137*, 433−444.
(47) Wang, L. (2009) Towards revealing the structure of bacterial inclusion bodies. *Prion 3*, 139−145.
(48) Gatti-Lafranconi, P., Natalello, A., Ami, D., Doglia, S. M., and Lotti, M. (2011) Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology. *FEBS J. 278*, 2408−2418.
(49) Tycko, R. (2006) Molecular structure of amyloid fibrils: Insights from solid-state NMR. *Q. Rev. Biophys. 39*, 1−55.
(50) McDermott, A. (2009) Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu. Rev. Biophys. 38*, 385−403.
(51) Fowler, D. J., Weis, R. M., and Thompson, L. K. (2010) Kinase-active signaling complexes of bacterial chemoreceptors do not contain proposed receptor-receptor contacts observed in crystal structures. *Biochemistry 49*, 1425−1434.

Supporting Information for: “Solid-State Nuclear Magnetic Resonance Spectroscopy of Human Immunodeficiency Virus gp41 Protein that Includes the Fusion Peptide: NMR Detection of Recombinant Fgp41 in Inclusion Bodies in Whole Bacterial Cells and Structural Characterization of Purified and Membrane-Associated Fgp41”

Fgp41 amino acid and DNA sequences

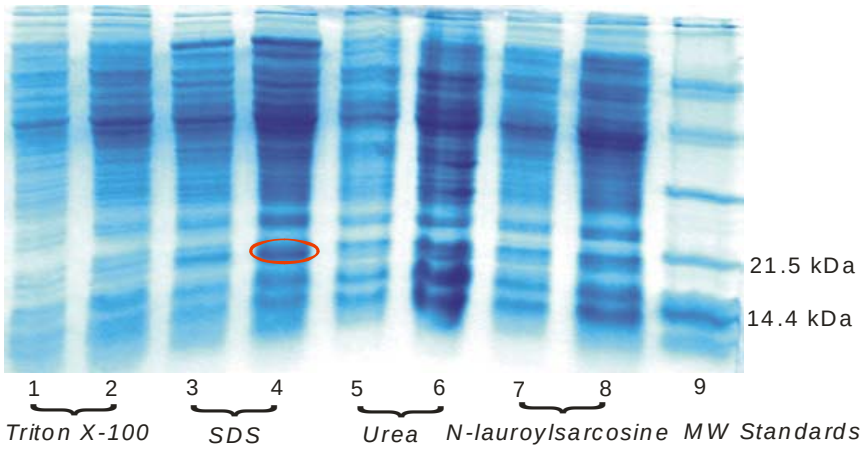
Start A V G L G A V F L G F L G A A G S T M G A A S M T L T V Q A R Q L L S G I V Q Q Q S N L L K A I E A Q Q H L L K L T V W G I K Q L Q A R V L A V E R Y L Q D Q Q L L G I W G C S G K L I C T S F V P W N N S W S N K T Y N E I W D N M T W L Q W D K E I S N Y T D T I Y R L L E D S Q N Q Q E K N E Q D L L A L D K L E H H H H H H **Stop**

atggcagttggactaggagctgtcttccttgggttcttgggagcagcagggagcactatgggcgcggcgtcaatgacgctgacg
gtacaggccagacaattattgtctggcatagtgcaacagcaaagcaatttgctgaaggctatagaggctcaacagcatctgttga
aactcacggtctggggtattaaacagctccaggcaagagtcctggctgtggaaagatacctacaggatcaacagctcctgggaa
tttggggctgctctggaaaactcatctgcacctcttttgtgccctggaacaatagttggagtaacaagacttataatgagatttggg
acaacatgacctggttgcaatgggataaagaaattagcaattacacagacacaatatacaggctacttgaagactcgcagaacca
gcaggaaaagaatgaacaagacttattggcattagataaactcgagcaccaccaccaccaccactga

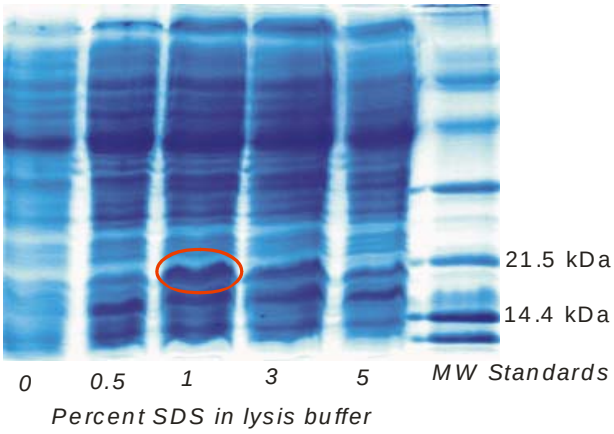
HIV-1 Sequences: Fgp41 (black) and HXB2 (blue)

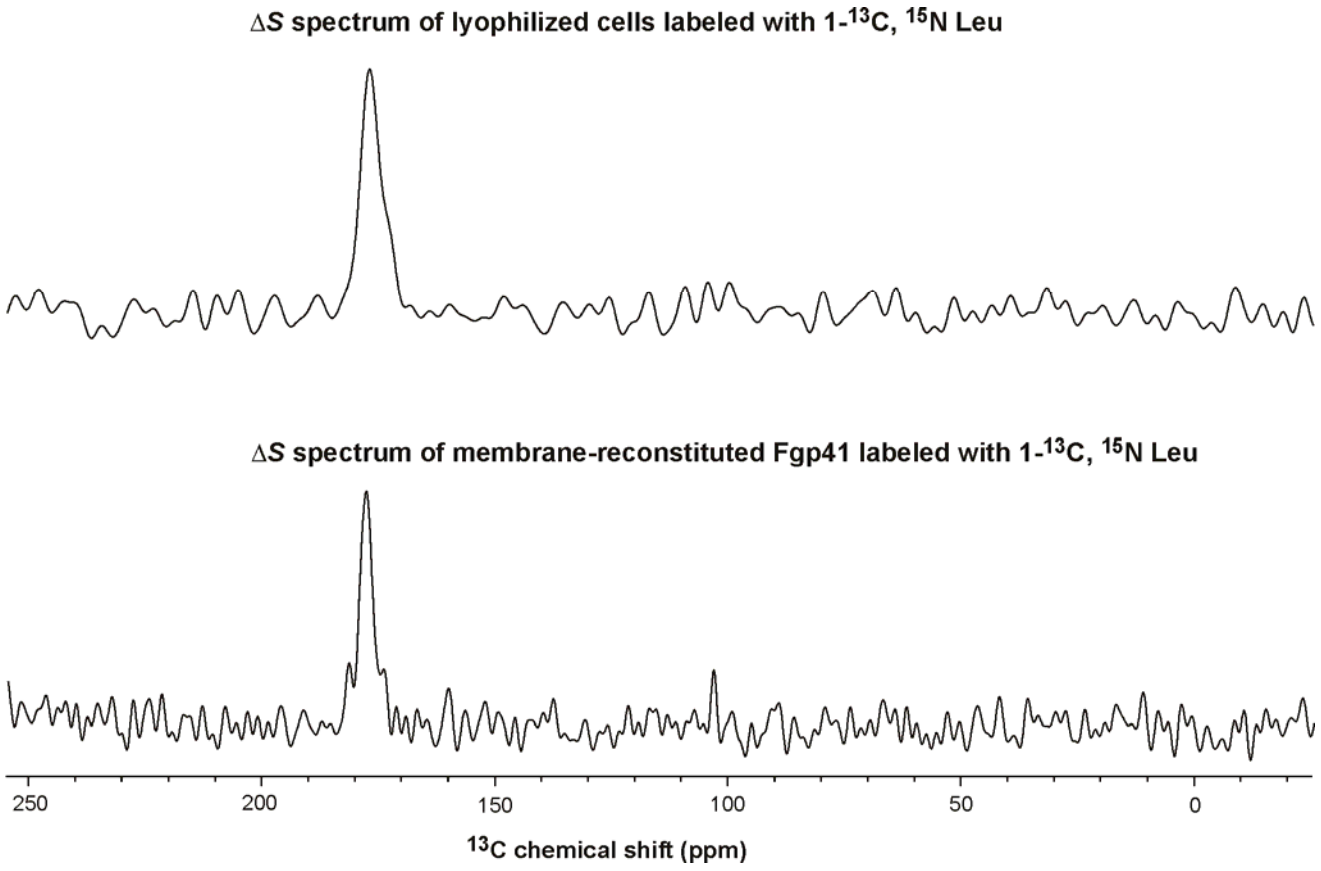
```
1
AVGLGAVFLG  FLGAAGSTMG  AASMTLTVQA  RQLLSGIVQQ  QSNLLKAIEA  QQHLLKLTVW
AVGIGALFLG  FLGAAGSTMG  AASMTLTVQA  RQLLSGIVQQ  QNNLLRAIEA  QQHLLQLTVW
61
GIKQLQARVL  AVERYLQDQQ  LLGIWGASGK  LIATSFVPWN  NSWSNKTYNE  IWDNMTWLQW
GIKQLQARIL  AVERYLKDQQ  LLGIWGCSGK  LICTTAVPWN  ASWSNKSLWQ  IWNHTTWMEW
121
DKEISNYTDT  IYRLLEDSQN  QQEKNEQDLL  ALDKLEHHHH  HH
DREINNYTSL  IHSLIEESQN  QQEKNEQELL  ELDK------  --
```

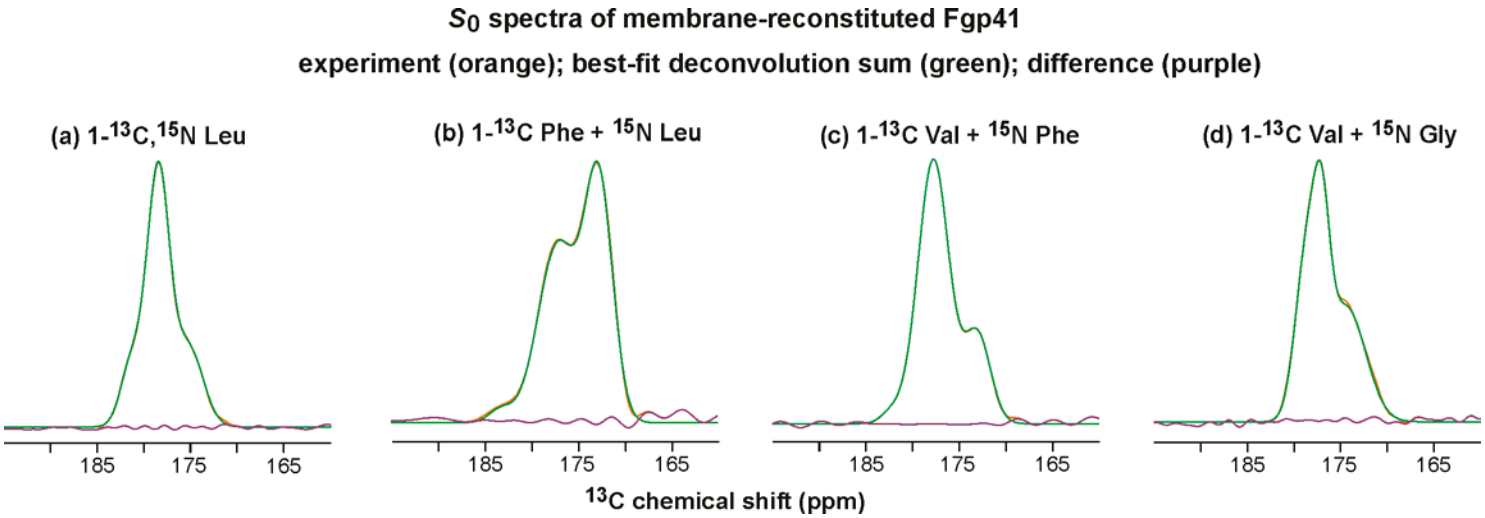
Representative SDS-PAGE of soluble cell lysates produced using buffers with different detergents or urea. For each buffer, the left and right lanes respectively correspond to 2 and 5 μL aliquots of lysate. The ~19 kDa band apparent in some lanes is assigned to Fgp41. One example is circled in red in lane 4 for lysis in SDS. Bands that may be Fgp41 were also apparent for lyses in either urea or *N*-lauroylsarcosine but purifications of these lysates consistently yielded <1 mg Fgp41/L culture whereas purifications of SDS lysates yielded >1 mg Fgp41/L culture. Subsequent lyses were therefore done with SDS.

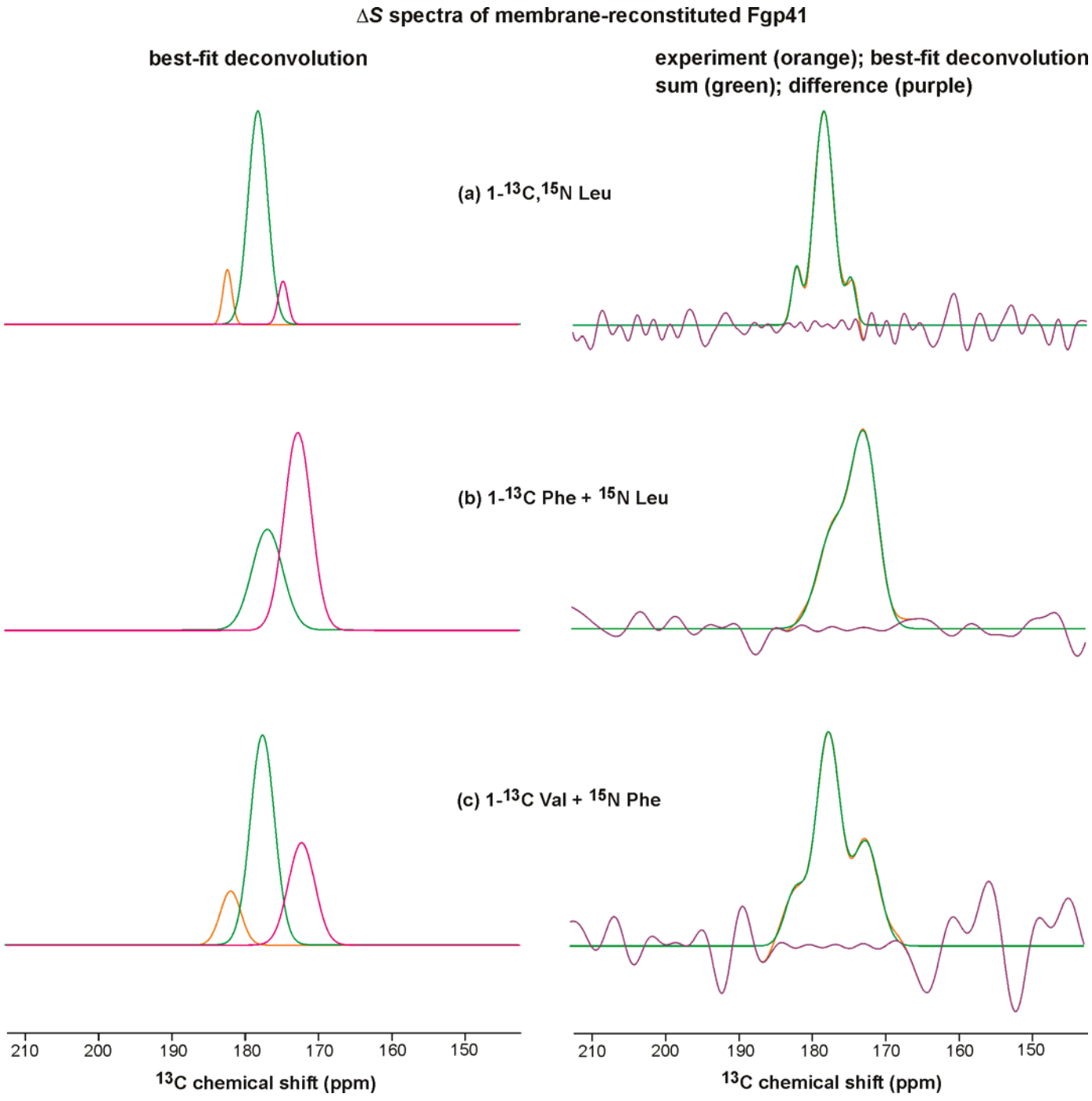


SDS-PAGE of soluble cell lysates produced using buffers with different [SDS].

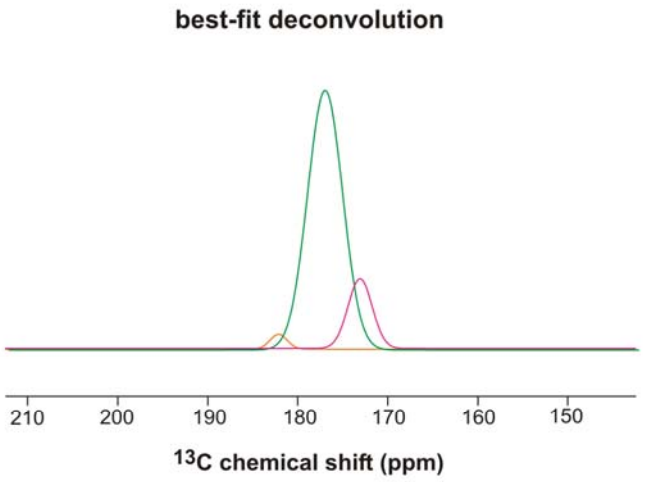


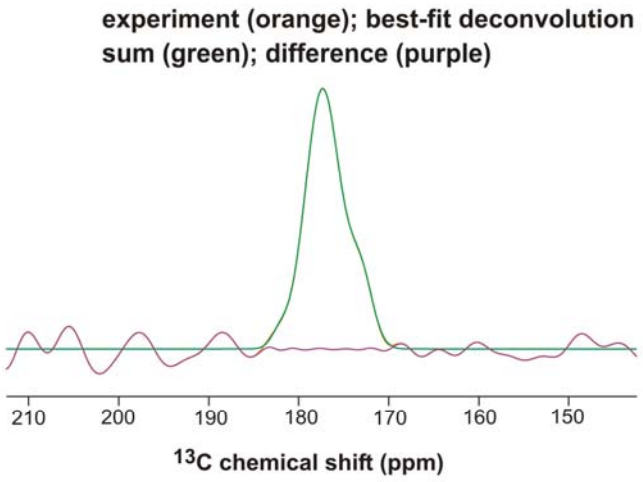

ΔS spectrum of lyophilized cells labeled with 1-13C, 15N Leu
ΔS spectrum of membrane-reconstituted Fgp41 labeled with 1-13C, 15N Leu
250
200
150
100
50
0
13C chemical shift (ppm)


S0 spectra of membrane-reconstituted Fgp41
experiment (orange); best-fit deconvolution sum (green); difference (purple)
(a) 1-13C,15N Leu
(b) 1-13C Phe + 15N Leu
(c) 1-13C Val + 15N Phe
(d) 1-13C Val + 15N Gly
185
175
165
13C chemical shift (ppm)


ΔS spectra of membrane-reconstituted Fgp41
best-fit deconvolution
experiment (orange); best-fit deconvolution sum (green); difference (purple)
(a) 1-13C,15N Leu
(b) 1-13C Phe + 15N Leu
(c) 1-13C Val + 15N Phe
210
200
190
180
170
160
150
13C chemical shift (ppm)

Analysis of ΔS spectrum of lyophilized cells labeled with 1-^{13}C, ^{15}N Leu





Analysis of S_0 spectrum from difference data between lyophilized cells with 1-^{13}C, ^{15}N Leu and cells with unlabeled Leu