



# Solid-State NMR Spectroscopy of the HIV gp41 Membrane Fusion Protein Supports Intermolecular Antiparallel $\beta$ Sheet Fusion Peptide Structure in the Final Six-Helix Bundle State

Kelly Sackett, Matthew J. Nethercott, Zhaoxiong Zheng and David P. Weliky

Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA

Correspondence to David P. Weliky: [weliky@chemistry.msu.edu](mailto:weliky@chemistry.msu.edu)

<http://dx.doi.org/10.1016/j.jmb.2013.11.010>

Edited by E. O. Freed

## Abstract

The HIV gp41 protein catalyzes fusion between viral and target cell membranes. Although the  $\sim 20$ -residue N-terminal fusion peptide (FP) region is critical for fusion, the structure of this region is not well characterized in large gp41 constructs that model the gp41 state at different times during fusion. This paper describes solid-state NMR (SSNMR) studies of FP structure in a membrane-associated construct (FP-Hairpin), which likely models the final fusion state thought to be thermostable trimers with six-helix bundle structure in the region C-terminal of the FP. The SSNMR data show that there are populations of FP-Hairpin with either  $\alpha$  helical or  $\beta$  sheet FP conformation. For the  $\beta$  sheet population, measurements of intermolecular  $^{13}\text{C}$ - $^{13}\text{C}$  proximities in the FP are consistent with a significant fraction of intermolecular antiparallel  $\beta$  sheet FP structure with adjacent strand crossing near L7 and F8. There appears to be negligible in-register parallel structure. These findings support assembly of membrane-associated gp41 trimers through interleaving of N-terminal FPs from different trimers. Similar SSNMR data are obtained for FP-Hairpin and a construct containing the 70 N-terminal residues of gp41 (N70), which is a model for part of the putative pre-hairpin intermediate state of gp41. FP assembly may therefore occur at an early fusion stage. On a more fundamental level, similar SSNMR data are obtained for FP-Hairpin and a construct containing the 34 N-terminal gp41 residues (FP34) and support the hypothesis that the FP is an autonomous folding domain.

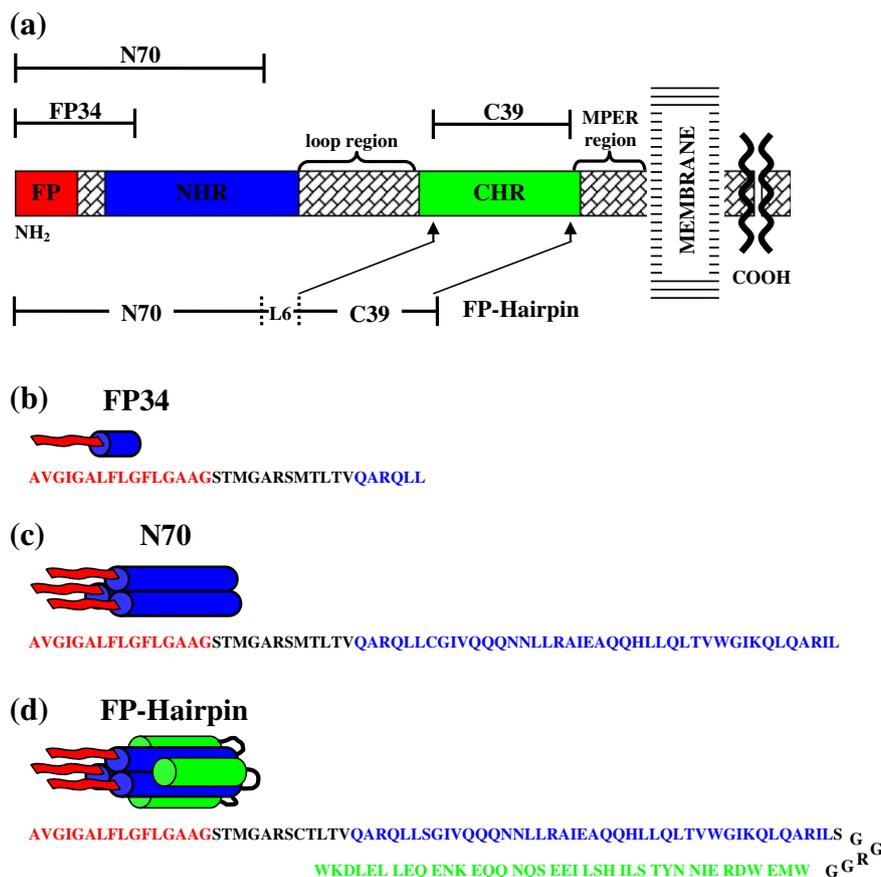
© 2013 Elsevier Ltd. All rights reserved.

## Introduction

The initial step of infection by HIV involves fusion between viral and host cell membranes at or near physiologic pH [1]. Membrane fusion is facilitated/catalyzed by the HIV gp120/gp41 protein complex where gp41 is a monotopic integral membrane protein and gp120 is non-covalently associated with the gp41 ectodomain, that is, the  $\sim 175$  N-terminal residues that are outside the virus (Fig. 1). It is likely that gp41 is assembled as a trimer and each trimer has three associated gp120s (Fig. 2) [2]. Binding of gp120 to target cell receptors triggers a change in the gp120/gp41 structural topology followed by putative binding of the  $\sim 20$ -residue N-terminal “fusion peptide” (FP) residues of gp41 to the target cell membrane and gp41 structural changes [3].

gp41-mediated membrane fusion is often pictured as a time series of structural states for the “soluble

ectodomain” that is C-terminal of the FP. These states include an early “pre-hairpin intermediate” (PHI) and the final “six-helix bundle” (SHB) in which the soluble ectodomain is either extended or compact, respectively (Fig. 2). The PHI state has been postulated from peptide inhibitors of fusion and interpretations of electron density in cryoelectron tomography [1,4]. The SHB state is supported by atomic-resolution structures of parts of the soluble ectodomain that show a hairpin motif for individual gp41 molecules, that is, *N*-helix– $180^\circ$  turn–*C*-helix motif (Fig. 1) [1,5–7]. In Figs. 1 and 2, “NHR” and “CHR” respectively mean *N*-heptad and *C*-heptad repeats and are an alternate terminology for the *N*- and *C*-helices. MPER means “membrane-proximal external region”. In addition, gp41 molecules are assembled as trimers with three interior *N*-helices in parallel coiled-coil configuration and three exterior *C*-helices antiparallel with the *N*-helices. Antibody-binding studies of FP + *N*-helix



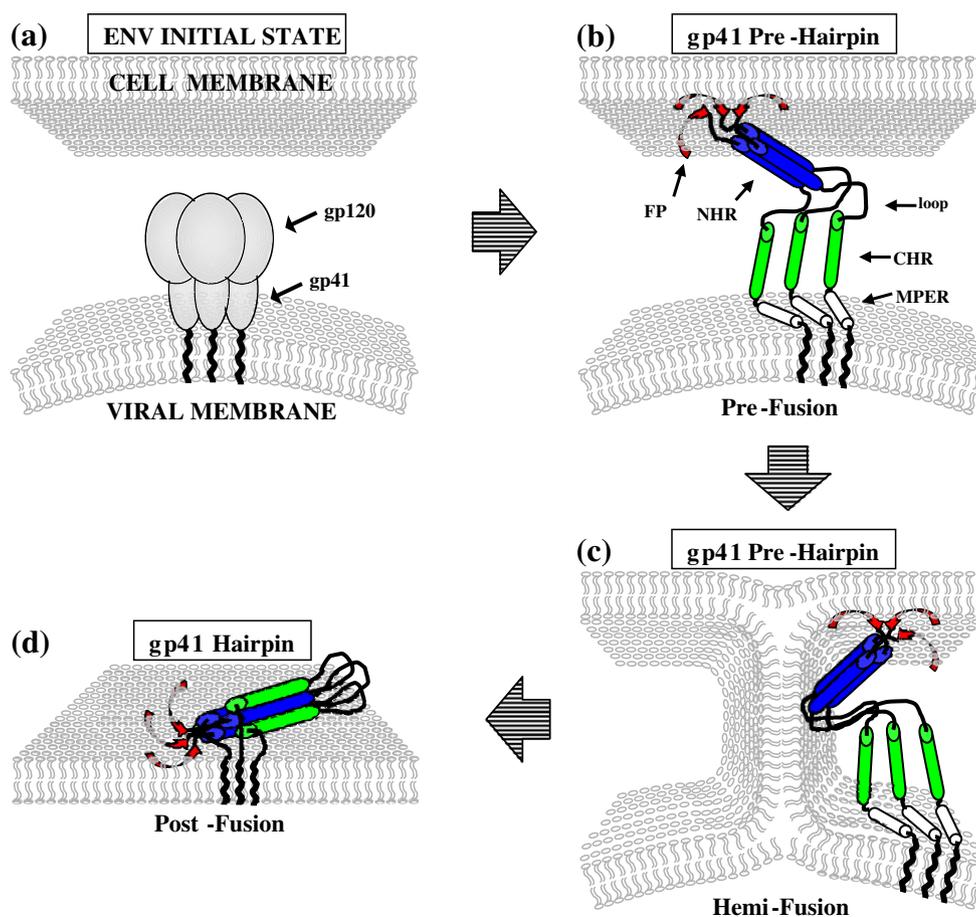
**Fig. 1.** (a) Schematic of the HIV gp41 ectodomain. Primary functional regions are designated by colored boxes and additional functional regions are specified above in braces. Above and below in brackets are the gp41 fragments under study. (b–d) Structural models of FP34, N70, and FP-Hairpin [N70(L6)C39], with primary sequence given below each and color coded to match functional regions in (a). “NHR” and “CHR” respectively mean *N*-heptad and *C*-heptad repeats and are an alternate terminology for the *N*- and *C*-helices. MPER means “membrane-proximal external region”.

constructs that model the *N*-PHI were interpreted to support a significant fraction of molecules with parallel trimeric coiled-coil *N*-helices (Fig. 1c) [8]. In our view, most functional and electron microscopy data support a requirement of multiple gp41 trimers for fusion [9–11]. There are little higher-resolution structural data about these assemblies of trimers and one contribution of the present work is insight into gp41 inter-trimer structure.

There are conflicting ideas about the roles of the putative early-stage PHI and final SHB gp41 states in fusion catalysis. One hypothesis is membrane fusion contemporaneous with the PHI → SHB structural change, although in our view, there are few supporting experimental data [12]. The underlying thinking is that some of the free energy released during this change provides the work to move lipid molecules along the fusion reaction coordinate. Figure 2 presents an alternate model in which most membrane fusion steps occur while gp41 is in the putative PHI state.

These steps include inter-membrane lipid mixing (hemifusion) and initial fusion pore formation. The final SHB state stabilizes the fused membranes including the fusion pore and induces fusion arrest. This model is supported by experimental results including the following: (1) the nascent fusion pores formed in gp41-mediated cell/cell fusion are closed upon addition of a peptide that likely prevents the PHI → SHB structural change; (2) HIV infection and gp41-mediated cell/cell fusion are inhibited by soluble ectodomain constructs of gp41 with SHB structure; (3) gp41 ectodomain constructs that model the PHI induce rapid vesicle fusion at physiologic pH, whereas those with SHB structure induce negligible fusion; and (4) vesicle fusion induced by shorter FP constructs (including a PHI model) is inhibited by SHB constructs [12–15].

The functional significance of the FP has been demonstrated by reduction/elimination of HIV/cell fusion and HIV infection with some FP mutations [9].



**Fig. 2.** Qualitative working model of HIV gp41-mediated fusion with colors matched to Fig. 1. “NHR” and “CHR” respectively mean *N*-heptad and *C*-heptad repeats and are an alternate terminology for the *N*- and *C*-helices. MPER means “membrane-proximal external region”. (a) Prior to interaction with cell receptors, gp41 and gp120 form a non-covalent complex that is likely trimeric. (b–d) Following gp120/receptors interaction, gp41 mediates membrane fusion through protein/membrane interaction. There is substantial experimental evidence supporting multiple trimers at the fusion site, but for clarity, only one trimer is shown. The model includes an early extended PHI (b and c) and final compact SHB (d) state of gp41. Studies of HIV/cell, cell/cell, and vesicle fusion support gp41 in the PHI state for most fusion steps. Membrane-inserted  $\beta$  structure is shown for the FP region and is based on (primarily SSNMR) data for the 23-residue HFP. Figure 9 provides a more biophysically realistic model of antiparallel  $\beta$  sheet structure with two trimers.

Similar functional effects of these mutations are detected in gp41-mediated cell/cell fusion [16]. The functional importance of the FP has been further supported with observations of vesicle fusion induced by some FP-containing gp41 ectodomain constructs in the absence of gp120 and cellular receptors [17,18]. The mutation–fusion activity relationships are similar to what is observed in HIV/cell and cell/cell fusion [19,20]. For FP-containing gp41 constructs that model the PHI and SHB states, native gels/Western blots show higher-molecular-weight bands that may correspond to oligomers of trimers whereas there is often only a single trimer band in the absence of FP [8]. Thus, one FP function may be oligomerization of trimers. Membrane insertion of short FP constructs

may be similar to insertion of the gp41 FP in the target cell membrane during fusion [20,21].

Higher-resolution structural data have been obtained for short FP constructs. For example, solid-state NMR (SSNMR) has been applied to “HFP” that contains the 23 N-terminal residues of gp41. HFP is bound to membranes that contain physiological fractions of anionic lipids and cholesterol [22]. SSNMR data including chemical shifts and internuclear proximities are consistent with small oligomers with intermolecular antiparallel  $\beta$  sheet structure [23,24]. A significant fraction of adjacent strands cross near L7 and F8. The  $\beta$  sheet HFP oligomer inserts into but does not traverse the membrane bilayer with deeper insertion of residues in the  $\beta$  sheet

center (e.g., A6 and L9) and membrane surface location of residues near the  $\beta$  sheet termini (e.g., A1 and A14) [21]. In membranes lacking cholesterol, there are two populations with either  $\alpha$  helical or  $\beta$  sheet structure [20,25,26]. Other biophysical data are also consistent with these two populations [19,27–31].

There are little FP structural data for larger gp41 constructs that model either the early PHI or final SHB gp41 states, and it is therefore not known whether the detailed HFP structure is a good model for the PHI or SHB FP. This motivates the present study that provides extensive data for the FP-Hairpin construct that contains most of the *N*- and *C*-helix regions and a short non-native loop (Fig. 1d). Earlier circular dichroism, calorimetric, and SSNMR data for FP-Hairpin are consistent with highly helical structure for the region C-terminal of the FP and therefore support SHB structure [32]. The  $^{13}\text{C}$  chemical shift distribution of L7 in the FP is consistent with distinct populations of molecules with either a  $\beta$  sheet or an  $\alpha$  helical FP. The present study provides shift data for four additional FP residues. Data are also provided about the fractional populations of parallel and antiparallel  $\beta$  sheet FPs of FP-Hairpin with insight into the assembly of multiple gp41 trimers during HIV/host cell fusion. We think that this is the first information about  $\beta$  sheet FP of gp41 in its final SHB state.

This paper also provides comparative data for the FP34 and N70 constructs that are the N-terminal 34 and 70 residues of gp41, respectively (Fig. 1b and c). N70 has the FP and much of the *N*-helix and is a model of the *N*-PHI. Earlier SSNMR work showed large populations of FP34 and N70 with  $\beta$  sheet FPs [32]. There are large differences in the rates and extents of vesicle fusion induced by FP34, N70, and FP-Hairpin at physiologic pH: moderate, high, and negligible fusogenicities, respectively. The dramatic functional difference between the PHI model, N70, and the SHB model, FP-Hairpin, supports earlier work showing that (1) most HIV/cell fusion happens with gp41 in the PHI state and that (2) SHB gp41 is an inhibitor of HIV fusion and infection. A fraction of N70 likely forms parallel trimers and larger oligomers, and the fusion rate of N70 is similar to that of a trimeric cross-linked HFP [17]. FP trimers/oligomers may therefore have an important role in the rapid vesicle fusion induced by N70 and in the HIV/cell fusion catalyzed by PHI gp41. However, there are also trimers/oligomer FPs in FP-Hairpin; thus, the negligible fusion induced by this construct likely reflects a counteracting effect of the appended SHB such as charge repulsion from a negatively charged membrane [18]. Although  $\beta$  and  $\alpha$  FP populations are observed for many FP-containing constructs, in our view, there is not yet a clear experimental correlation between fusogenicity and FP conformation. Overall, understanding HIV/cell fusion requires more structural data (such as those described in this

paper) on large, membrane-associated gp41 constructs that model the PHI and SHB states.

## Results and Discussion

### Sample preparation

FP-Hairpin had a single labeled (*lab*) and directly bonded  $^{13}\text{C}/^{15}\text{N}$  spin pair in the FP region, that is, at I4/G5, L7/F8, F8/L9, F11/L12, or L12/G13 (Fig. 1). For both N70 and FP34, the *lab* spin pair was L7/F8. “Membrane” samples are prepared by initially solubilizing protein in 10 or 20 mM formate buffer at pH 3 (“Buffer”) followed by dropwise addition to a vesicle suspension buffered at pH 7, with a final pH of 7. The typical vesicle composition was DTPC: DTPG:cholesterol (8:2:5 mole ratio) that reflects dominant choline headgroup, significant negatively charged lipid, and approximate lipid:cholesterol ratio found in viral and host cell plasma membranes [22]. “Membrane + D-malt” samples are similarly prepared except that the initial protein solution is Buffer + decyl maltoside (D-malt) detergent that is non-ionic and non-denaturing. D-malt was added with the intent of reducing protein aggregation. FP-Hairpin binds vesicles at pH 7 but does not induce inter-vesicle lipid mixing (i.e., no vesicle fusion) [18]. The thermostable SHB structure and lack of vesicle fusion of FP-Hairpin at pH 7 are not affected by D-malt. The protein + vesicle suspension is centrifuged, and the hydrated pellet is transferred to an SSNMR rotor. In the manuscript, spectra and corresponding data for samples without D-malt are shown in black and those with D-malt are shown in red.

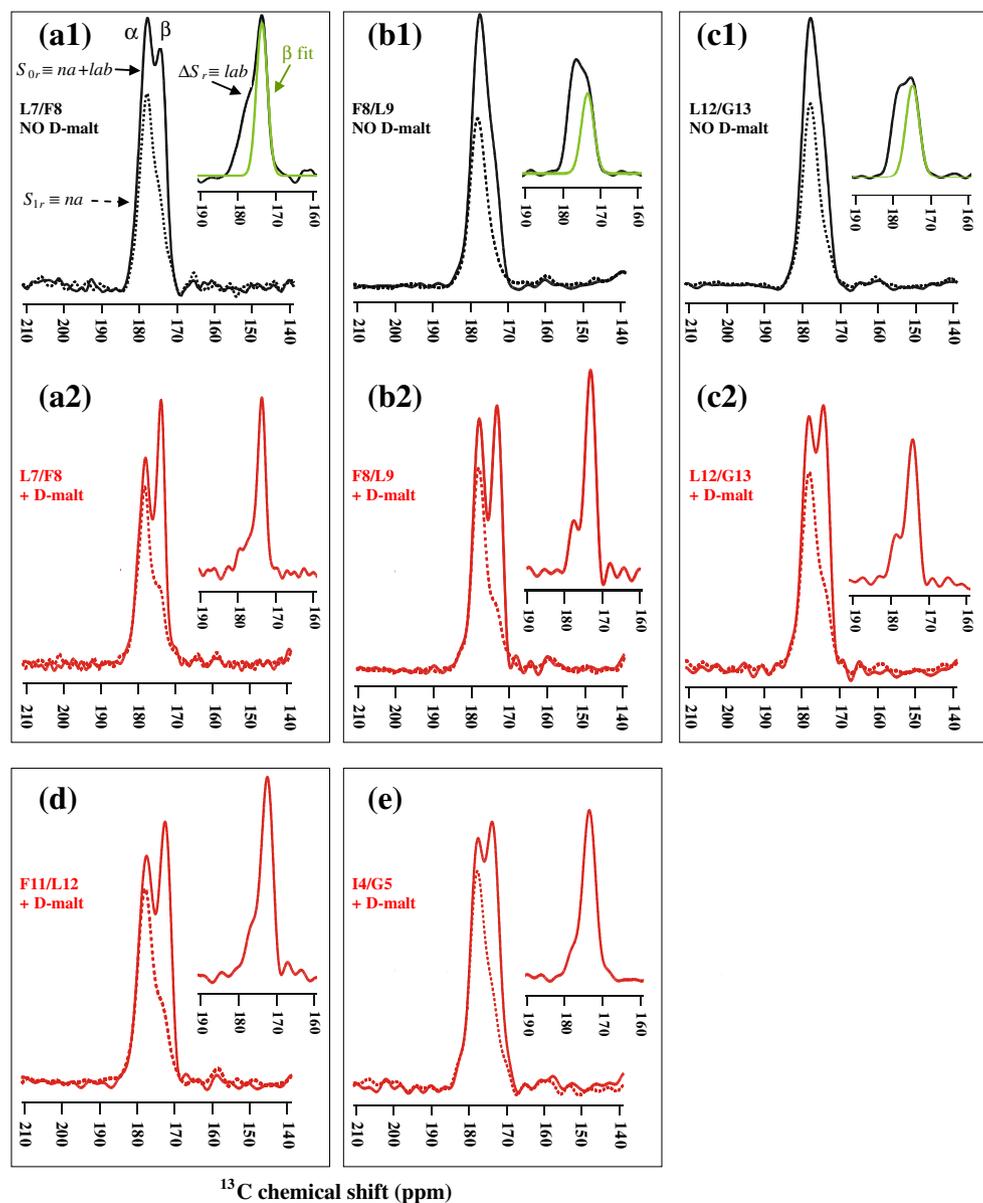
### $\beta$ and $\alpha$ FP populations of FP-Hairpin

As noted in Introduction, an earlier SSNMR study showed approximately equal populations of FP-Hairpin molecules with either  $\beta$  or  $\alpha$  structure at residue L7 in the FP region. One of the goals of the present study is delineation of these populations at other residues in the FP region. These populations provide a more global view of membrane-associated FP structure with C-terminal SHB, which likely models the final fusion state of gp41 (Fig. 2).

Each FP-Hairpin molecule contains 1 *lab* and 144 unlabeled CO sites. Although there is only 0.011  $^{13}\text{C}$  natural abundance (*na*) at each unlabeled site, an unfiltered  $^{13}\text{C}$  spectrum will have a total 0.61 fractional contribution from all of the *na* sites and corresponding 0.39 fractional contribution from the *lab* site. The unfiltered spectrum is therefore not useful for determination of the fractional  $\beta$  and  $\alpha$  populations at the *lab* residue. This determination requires selective detection of the *lab*  $^{13}\text{C}$  signal and was performed

using the rotational-echo double-resonance (REDOR) pulse sequence [33–36]. The interpretation and analysis of the resulting REDOR spectra are illustrated with Fig. 3a-1 from the membrane sample with L7/F8 ( $^{13}\text{CO}/^{15}\text{N}$ ) labeling. The lower-left black and broken traces are the “ $S_{0r}$ ” and “ $S_{1r}$ ” spectra, respectively, and the upper-right black trace is the  $\Delta S_r = S_{0r} - S_{1r}$  difference spectrum. The  $S_{0r}$ ,  $S_{1r}$ , and  $\Delta S_r$  spectra respectively represent the total  $na + lab$ ,  $na$ , and  $lab$

$^{13}\text{CO}$  signals. Interpretation of the peaks in these spectra relies on the well-described correlation between local  $\beta$  or  $\alpha$  structure and respective lower or higher backbone  $^{13}\text{CO}$  shift [25,37]. The higher shift peak in the  $S_{0r}$  spectrum is therefore considered to have a large contribution from  $na$  sites in the highly helical SHB domain while the lower shift peak is dominated by the  $lab$  signal from molecules with  $\beta$  conformation at the  $lab$  residue. The  $S_{1r}$  spectrum



**Fig. 3.**  $^{13}\text{CO}$  REDOR spectra of samples containing FP-Hairpin with a single directly bonded  $^{13}\text{CO}/^{15}\text{N}$  spin pair at (a) L7/F8, (b) F8/L9, (c) L12/G13, (d) F11/L12, and (e) I4/G5. Samples (a-1), (b-1), and (c-1) (black spectra) were prepared with initial protein solubilization in Buffer, whereas samples (a-2), (b-2), (c-2), (d), and (e) (red spectra) were prepared with solubilization in Buffer + D-malt. The continuous and dotted lines are the total  $na + lab$  ( $S_{0r}$ ) and the  $na$  ( $S_{1r}$ ) spectra, respectively. Inset in each figure is the  $lab$  spectrum ( $\Delta S_r$ ) with deconvoluted  $\beta$  sheet peak in green. Spectral processing included Gaussian line broadening of 150 and 100 Hz, respectively, for (a-1) and (c-2) and polynomial baseline correction for all spectra. Each  $S_{0r}$  and  $S_{1r}$  spectrum is the sum of (a-1) 76,128, (a-2) 9328, (b-1) 8928, (b-2) 36,597, (c-1) 19,856, (c-2) 68,308, (d) 20,000, and (e) 51,904 scans.

reflects the *na* signal and the higher shift  $\alpha$  peak is dominant. By contrast, the  $\Delta S_r$  spectrum reflects the *lab* signal and the lower shift  $\beta$  peak is dominant.

The determination of the fractional  $\beta$  and  $\alpha$  populations at the *lab* residue is based on deconvolution of the  $\Delta S_r$  spectrum into lineshapes that are assigned to  $\beta$  and  $\alpha$  structures based on peak shift. The fractional  $\beta$  and  $\alpha$  populations are the fractional integrated intensities of their respective lineshapes. The green traces in Fig. 3a-1, b-1, and c-1 are the deconvolved  $\beta$  lineshapes, and complete deconvolutions for all spectra are provided in Supplementary Information. Each sum of lineshapes matches well to the experimental spectrum, and the peak shifts, linewidths, and fractional integrated intensities from the deconvolutions are listed in Table 1.

### Higher $\beta$ population in membrane + D-malt samples

The effect of the absence *versus* the presence of D-malt in the initial protein solubilization buffer is examined with comparison of the  $\Delta S_r$  fittings of samples with the same labeled protein, for example, Fig. 3a-1 *versus* a-2, Fig. 3b-1 *versus* b-2, and Fig. 3c-1 *versus* c-2. The peak shifts agree to within  $\pm 0.3$  ppm while linewidths in D-malt samples are narrower by 0.4–0.9 ppm. For membrane samples, there are approximately equal populations of molecules with either  $\beta$  or  $\alpha$  structure at the *lab* site, as highlighted by comparison of the green ( $\beta$  lineshape) and black (full  $\Delta S_r$ ) traces in Fig. 3a-1, b-1, and c-1 insets. Membrane + D-malt samples have a larger  $\beta$ : $\alpha$  population ratio ( $\sim 3$ ). The narrower linewidths and higher  $\beta$  population with D-malt are clearly observed in

the  $S_{0r}$  spectra. D-malt is expected to reduce aggregation of FP-Hairpin; thus, the  $\beta$  and  $\alpha$  FP populations may respectively correlate with smaller and larger FP-Hairpin oligomers/aggregates in membranes, although we currently have no data about these oligomer sizes.

The  $\beta$  peak shift of either a membrane or membrane + D-malt sample agrees to within 0.8 ppm with the corresponding shift of the membrane-associated HFP (which lacks the SHB) [23]. A large body of SSNMR data support  $\beta$  sheet structure for HFP in membranes with cholesterol, and this structure may be a reasonable model for the  $\beta$  sheet FP of membrane-associated FP-Hairpin.

The  $\beta$  and  $\alpha$  FP populations observed for FP-Hairpin have also been observed for HFP and larger gp41 constructs (including a 154-residue ectodomain construct) with a variety of membrane compositions [38]. One of these compositions specifically reflected the typical lipid headgroup and cholesterol content of membranes of host cells of the virus [39]. The populations appear to reflect comparable free energies of the two structures, and these populations have also been observed for FPs from other viruses [40–43]. It is possible, although we think unlikely, that the free-energy difference between the  $\beta$  and  $\alpha$  states is dramatically changed by inclusion of the transmembrane domain of gp41 [30].

### Predominant antiparallel FP of membrane-associated FP-Hairpin

As noted in Introduction,  $\beta$  sheet FP structure has been observed for the 23-residue HFP bound to membranes with cholesterol content comparable to

**Table 1.** Fitting of FP-Hairpin  $\Delta S_r$  spectra

Sample/( $^{13}\text{C}$ label)	$\beta$ Sheet <sup>a</sup>			$\alpha$ Helix <sup>a</sup>		
	Peak shift (ppm)	Linewidth <sup>b</sup> (ppm)	Fractional intensity	Peak shift (ppm)	Linewidth <sup>b</sup> (ppm)	Fractional intensity
Membrane/(L7)	173.8	3.0	0.60	177.3	4.0	0.35
<b>Membrane + D-malt/(L7)</b>	<b>174.0</b>	<b>2.4</b>	<b>0.73</b>	<b>177.1</b>	<b>3.6</b>	<b>0.22</b>
Membrane/(F8)	173.4	3.9	0.43	177.0	4.1	0.57
<b>Membrane + D-malt/(F8)</b>	<b>173.3</b>	<b>3.0</b>	<b>0.80</b>	<b>177.5</b>	<b>3.1</b>	<b>0.20</b>
Membrane/(L12)	174.7	4.1	0.52	178.7	4.2	0.48
<b>Membrane + D-malt/(L12)</b>	<b>174.2</b>	<b>3.3</b>	<b>0.75</b>	<b>178.5</b>	<b>3.5</b>	<b>0.25</b>
<b>Membrane + D-malt/(F11)</b>	<b>172.6</b>	<b>3.3</b>	<b>0.76</b>	<b>176.3</b>	<b>3.9</b>	<b>0.24</b>
<b>Membrane + D-malt/(I4)</b>	<b>173.7</b>	<b>4.0</b>	<b>0.88</b>	<b>178.2</b>	<b>3.5</b>	<b>0.12</b>

Spectra are displayed in Fig. 3 insets and are fitted to the sum of two or three Gaussian lineshapes. The black (boldface) colors match the colors of the inset spectra. Fits are shown in Supplementary Information.

<sup>a</sup> Conformation is assigned by similarity between the peak  $^{13}\text{C}$  shift and the peak shift of the database distribution of the amino acid in a specific conformation. Database peak shifts (standard deviations) in ppm include the following: Leu,  $\beta$  strand, 175.7 (1.5); Leu, helix, 178.5 (1.3); Phe,  $\beta$  strand, 174.3 (1.6); Phe, helix, 177.1 (1.4); Ile,  $\beta$  strand, 174.9 (1.4); Ile, helix, 177.7 (1.3). Each  $\beta$  sheet shift of FP-Hairpin is within 1 ppm of the corresponding  $\beta$  sheet HFP shift where HFP residue (shift) include L7 (174.2), F8 (173.8), L12 (174.4), F11 (173.3), and I4 (174.5).

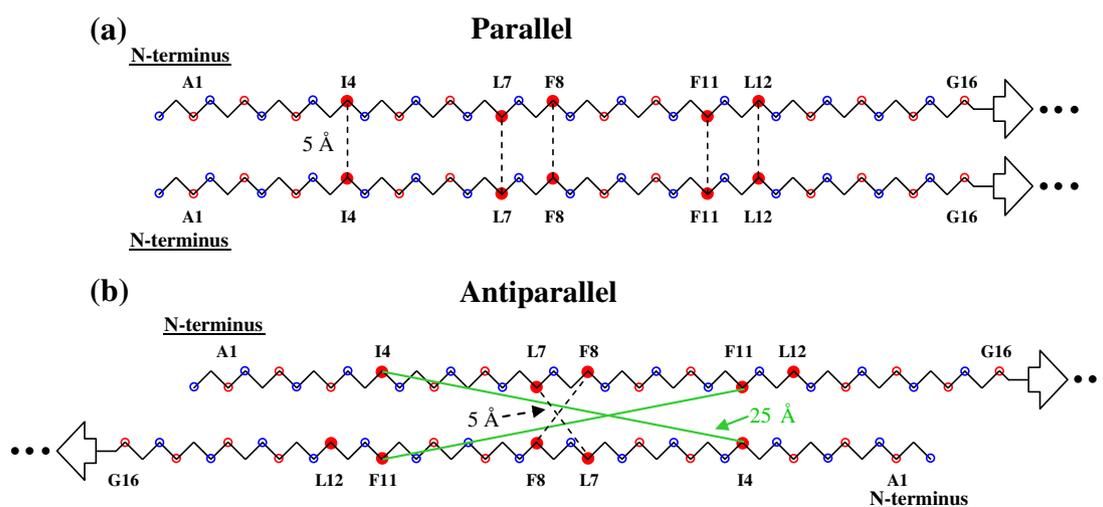
<sup>b</sup> Full-width at half-maximum.

that of host cell and viral membranes. There are also significant populations of molecules with  $\beta$  sheet FP for N70 and FP-Hairpin that are models of putative intermediate and final states of gp41, respectively. There is often rapid fusion of cholesterol-rich vesicles by FP constructs; thus, the  $\beta$  sheet FP is likely fusogenic and it is worthwhile to make high-resolution measurements for this  $\beta$  structure in membranes. For small constructs such as HFP, earlier SSNMR studies were consistent with small oligomers of HFP with antiparallel  $\beta$  structure and  $\sim 0.4$  fraction of adjacent molecules crossing near L7 and F8 [24]. There was no population of parallel  $\beta$  structure for HFP. However, it seems possible that parallel  $\beta$  FP structure could be promoted in FP-Hairpin as an extension of the parallel arrangement of the NHR regions (Fig. 1d). This motivated the SSNMR experiments of the present study to detect populations of membrane-associated FP-Hairpin molecules with either parallel or antiparallel FPs.

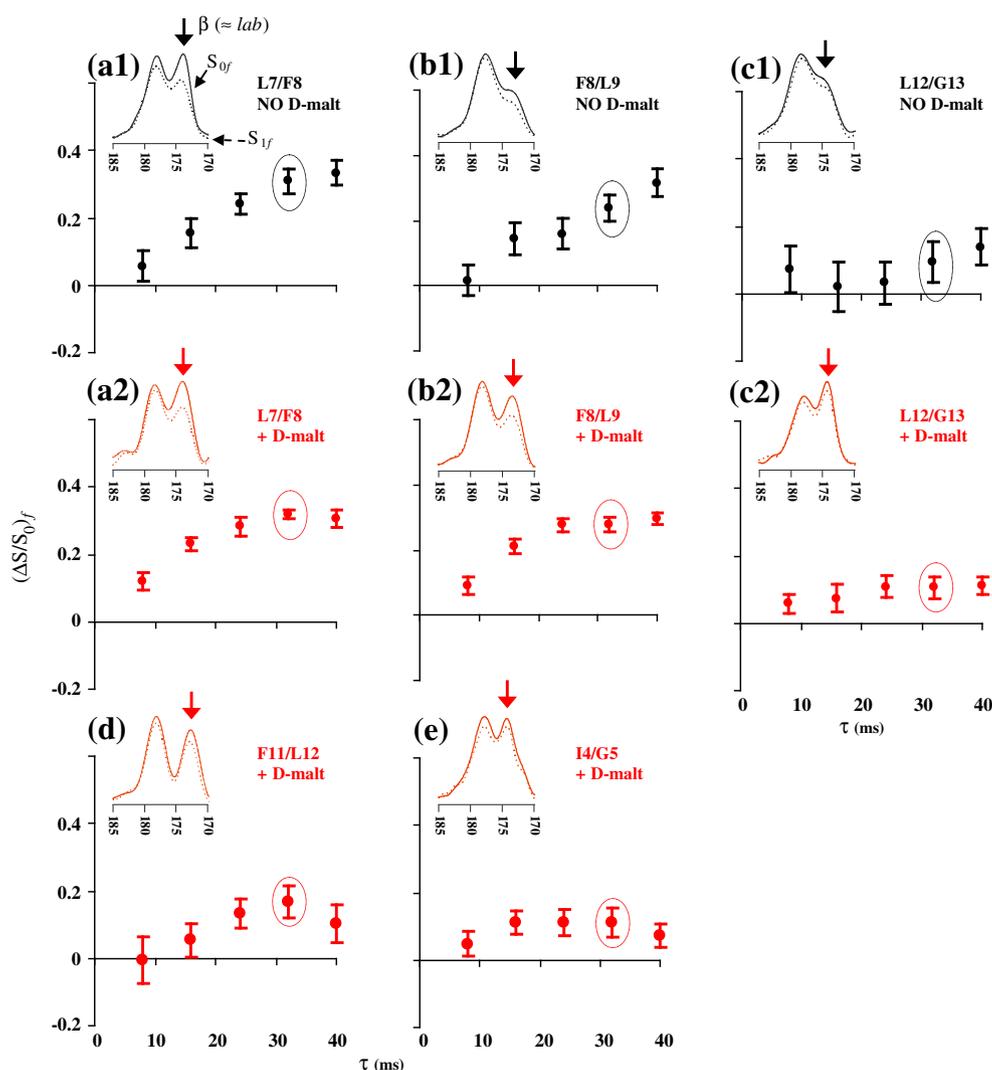
The samples for these experiments were the same as those described above and had a single *lab*  $^{13}\text{C}$  nucleus in the FP region. SSNMR was applied to probe proximities between the *lab*  $^{13}\text{C}$  nuclei on adjacent FP strands in the intermolecular  $\beta$  sheet. These proximities were then compared to those predicted for the parallel and antiparallel models displayed in Fig. 4. For the parallel model in Fig. 4a, the intermolecular distance ( $\equiv r_{ij}$ ) between *lab*  $^{13}\text{C}$  nuclei is  $\sim 5$  Å and also independent of the identity of the *lab* residue, that is,  $\sim 5$  Å for the I4, L7, F8, F11,

and L12 samples. For the antiparallel model in Fig. 4b, the  $r_{ij}$  is strongly dependent on the identity of the *lab* residue, that is,  $\sim 5$  Å for the L7 and F8 samples and  $>20$  Å for the I4, F11, and L12 samples. The dependence of the SSNMR-derived  $r_{ij}$  on *lab* site is used to assess the fractions of parallel and antiparallel strand arrangements.

The  $r_{ij}$  of a sample was probed using the finite-pulse constant-time double-quantum build-up (fpCTDQBU) SSNMR technique [25,44–47]. The  $^{13}\text{C}$ – $^{13}\text{C}$  dipole–dipole coupling scales as  $r^{-3}$  and the fpCTDQBU method detects such couplings between proximal nuclei, that is, those with  $^{13}\text{C}$ – $^{13}\text{C}$  distance  $\leq 10$  Å. Some of the fpCTDQBU  $^{13}\text{C}$  spectra and analysis are displayed in Figs. 5 and 6 with spectra denoted by  $S_{0f}$  or  $S_{1f}$  and respectively displayed with continuous or broken lines. The notation was consistent with the REDOR notation and the “f” subscript refers to fpCTDQBU. Similar to the REDOR  $S_{0r}$  spectra of Fig. 3, each fpCTDQBU  $^{13}\text{C}$   $S_{0f}$  spectrum has a lower shift  $\beta$  peak dominated by the *lab* FP signal and a higher shift  $\alpha$  peak dominated by the *na* SHB signal. The respective assignments of the  $\beta$  and  $\alpha$  peaks to dominant *lab* FP and *na* SHB contributions are supported by good agreement between the  $\beta$  peak shifts of the  $S_{0f}$  and  $\Delta S_f$  spectra, typically different by  $\leq 0.3$  ppm, and small sample-to-sample variation of the  $\alpha$  peak shift, typically  $\leq 0.5$  ppm (Table 2). Additional support for the assignments were much larger attenuations of the  $\beta$  signals relative to the  $\alpha$  signals in the REDOR  $S_{1f}$  spectra of Fig. 3. We also use  $S_{0f}^b$  and



**Fig. 4.** Backbone models of two strands of the apolar (residues 1  $\rightarrow$  16) FP region with either (a) in-register parallel 1  $\rightarrow$  16/1  $\rightarrow$  16 or (b) antiparallel 1  $\rightarrow$  15/15  $\rightarrow$  1  $\beta$  sheet structure. Blue and red open circles are N and CO nuclei, respectively; red filled circles are the  $^{13}\text{C}$  labels used in this study. Some residues are identified. The following distances are used: (1) 3.35 Å between CO nuclei of adjacent residues in a single strand; (2) 5.0 Å between CO nuclei at the same residue position on adjacent strands in the parallel structure; and (3) 4.1 Å between C and N in a  $\text{CO}\cdots\text{HN}$  hydrogen bond. For either model, the length of each broken line represents 5 Å. For the antiparallel model, the intercarbonyl intermolecular distances for F11–F11 and I4–I4 are 23 Å and 25 Å, respectively, and are indicated by green lines. The L12–L12 distance is 30 Å.



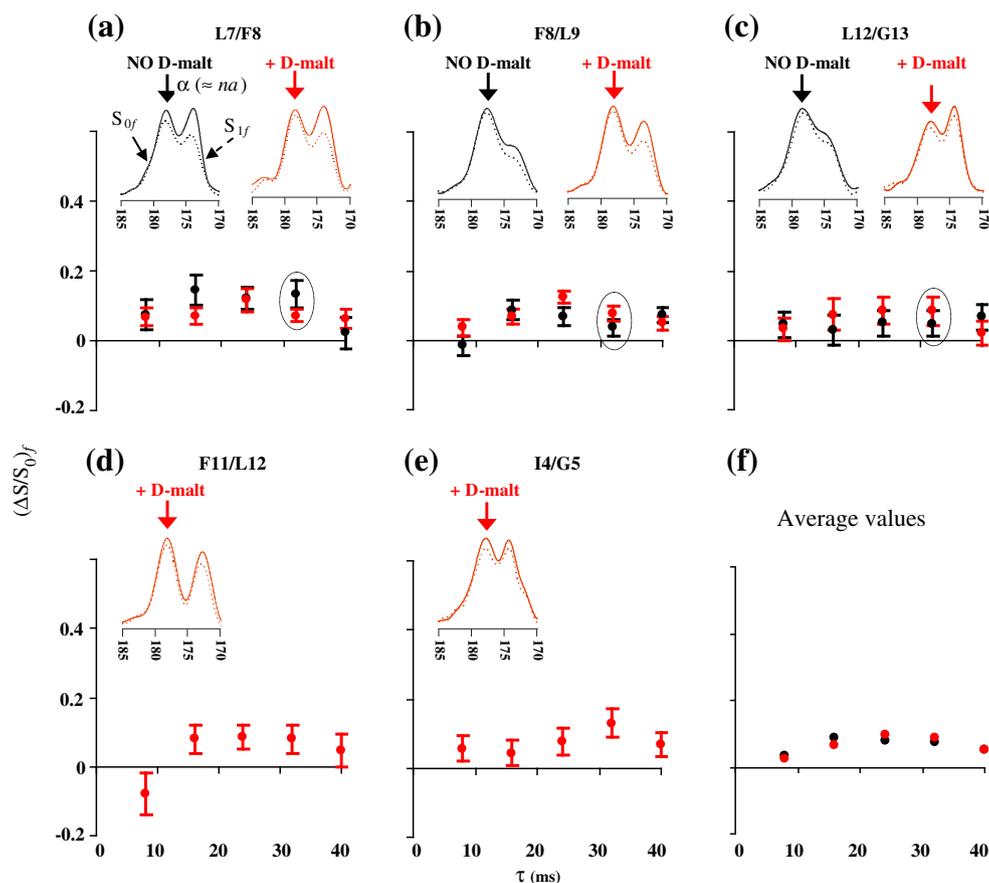
**Fig. 5.** fpCTDQBU  $(\Delta S/S_0)_f^\beta$  versus dephasing time  $\tau$  for samples containing FP-Hairpin  $^{13}\text{C}$  labeled at (a) L7, (b) F8, (c) L12, (d) F11, and (e) I4. (a-1), (b-1), and (c-1) (black circles) are membrane samples and (a-2), (b-2), (c-2), (d), and (e) (red circles) are membrane + D-malt samples. (a–e) Insets are experimental  $S_{0f}$  (continuous line) and  $S_{1f}$  (dotted line)  $^{13}\text{C}$  spectra for  $\tau = 32$  ms. Relative to  $S_{0f}$ , attenuation of the  $S_{1f}$  signal is due to  $^{13}\text{C}$ – $^{13}\text{C}$  couplings. The  $(\Delta S/S_0)_f^\beta$  are calculated using the  $^{13}\text{C}$  spectral intensities integrated in a 1.0-ppm window centered at the  $\beta$  peak shift (denoted by an arrow). Spectral processing was typically performed without line broadening and with polynomial baseline correction (see Supplementary Information). Each  $S_{0f}$  and  $S_{1f}$  spectrum represents the sum of (a-1) 15,000, (a-2) 12,000, (b-1) 12,000, (b-2) 11,000, (c-1) 14,000, (c-2) 12,000, (d) 16,000, and (e) 14,000 scans.

$S_{0f}^\alpha$  to respectively denote the  $\beta$  and  $\alpha$  peak intensities of the  $S_{0f}$  spectrum and  $S_{1f}^\beta$  and  $S_{1f}^\alpha$  to denote the corresponding intensities in the  $S_{1f}$  spectrum.

For the  $\beta$  fpCTDQBU signals dominated by *lab*  $^{13}\text{C}$  nuclei, larger *lab*  $^{13}\text{C}$ –*lab*  $^{13}\text{C}$  couplings and corresponding shorter  $r_{ij}$  are indicated by higher attenuation of  $S_{1f}^\beta$  relative to  $S_{0f}^\beta$ . For example, the Fig. 5a-1 and a-2 spectra show significant  $\beta$  signal attenuation for the L7 sample that is consistent with adjacent strands with proximal L7 residues. Attenuation is quantitated using the intensity ratio  $(S_{0f}^\beta - S_{1f}^\beta)/S_{0f}^\beta = (\Delta S/S_0)_f^\beta$ . The fpCTDQBU spectra were obtained for five different values of the dephasing time ( $\tau$ )

and significant build-up of  $(\Delta S/S_0)_f^\beta$  with  $\tau$  was an indicator of proximal *lab*  $^{13}\text{C}$  nuclei on adjacent strands. Figure 5 displays plots of  $(\Delta S/S_0)_f^\beta$  versus  $\tau$  for the different samples with inset spectra corresponding to  $\tau = 32$  ms. For a particular *lab* FP-Hairpin, the  $(\Delta S/S_0)_f^\beta$  build-up is typically comparable for the membrane and membrane + D-malt samples. Uncertainties are usually smaller for the membrane + D-malt samples because of the higher fractional population with  $\beta$  FP and because of narrower  $\beta$  linewidths.

The  $(\Delta S/S_0)_f^\beta$  data are first interpreted in terms of the dominant contribution of the *lab*  $^{13}\text{C}$  nuclei to



**Fig. 6.** fpCTDQBU  $(\Delta S/S_0)_f^\alpha$  versus dephasing time for samples containing FP-Hairpin  $^{13}\text{C}$  labeled at (a) L7, (b) F8, (c) L12, (d) F11, and (e) I4. Black circles are for membrane samples and red circles are for membranes + D-malt samples. Data are overlaid in (a)–(c) to compare samples with the same label prepared without or with D-malt. (f) The average  $(\Delta S/S_0)_f^\alpha$  of membrane samples (black dots) or membranes + D-malt samples (red dots). (a–e) The insets are experimental  $S_{0f}$  (continuous line) and  $S_{1f}$  (dotted line)  $^{13}\text{C}$  spectra for  $\tau = 32$  ms. The  $(\Delta S/S_0)_f^\alpha$  are calculated using the  $^{13}\text{C}$  spectral intensities integrated in a 1.0-ppm window centered at the  $\alpha$  peak shift (denoted by an arrow).

the  $\beta$  signal and intermolecular  $lab$   $^{13}\text{C}$ – $lab$   $^{13}\text{C}$  couplings due to  $r_{ij}$ . This basic interpretation is not changed by more detailed analysis that also considers the minor contribution of  $na$   $^{13}\text{C}$  nuclei to the  $\beta$  signal and couplings to  $na$   $^{13}\text{C}$  nuclei. This

detailed analysis is presented in a subsequent section of the paper.

The  $(\Delta S/S_0)_f^\beta$  data can be separated into two categories: (1) L7 and F8 samples for which  $(\Delta S/S_0)_f^\beta$  reaches 0.3 – 0.4 for  $\tau \geq 32$  ms; and (2) I4, F11, and

**Table 2.**  $S_{0f}$  and  $S_{0f}$   $^{13}\text{C}$  peak shifts (ppm) of FP-Hairpin

Sample/ $^{13}\text{C}$ label)	$S_{0f}$		$S_{0f}^a$	
	$\beta$ Sheet	$\alpha$ Helix	$\beta$ Sheet	$\alpha$ Helix
Membrane/(L7)	174.1	177.8	174.1	178.2
<b>Membrane + D-malt/(L7)</b>	<b>174.0</b>	<b>178.2</b>	<b>174.1</b>	<b>178.5</b>
Membrane/(F8)	173.4	177.7	173.5	177.8
<b>Membrane + D-malt/(F8)</b>	<b>174.4</b>	<b>178.0</b>	<b>173.7</b>	<b>178.3</b>
Membrane/(L12)	174.3	178.2	174.4	178.4
<b>Membrane + D-malt/(L12)</b>	<b>174.5</b>	<b>178.3</b>	<b>174.4</b>	<b>178.2</b>
<b>Membrane + D-malt/(F11)</b>	<b>172.9</b>	<b>177.6</b>	<b>172.9</b>	<b>178.2</b>
<b>Membrane + D-malt/(I4)</b>	<b>174.3</b>	<b>177.9</b>	<b>174.3</b>	<b>177.9</b>

Most spectra have well-resolved  $\beta$  and  $\alpha$  peaks and the peak shifts were measured by visual inspection. For the more poorly resolved membrane/(F8) and membrane/(L12) samples, the peak shifts were measured by deconvolution.

<sup>a</sup> Each  $S_{0f}$  shift is the average for all  $\tau$  values. For any particular sample, the standard deviation of the shift distribution is 0.1–0.2 ppm.

L12 samples for which  $(\Delta S/S_0)_f^\beta \leq 0.1$  for  $\tau \geq 32$  ms. The I4, F11, and L12 data correlate to  $r_{ij} > 10$  Å and are therefore inconsistent with a significant population of parallel strand arrangement for which  $r_{ij} \approx 5$  Å for all residues (Fig. 4a). The I4, F11, and L12 data are consistent with a significant population of antiparallel arrangement for which  $r_{ij} > 10$  Å for these residues (Fig. 4b). For this antiparallel arrangement with adjacent strand crossing near L7 and F8,  $r_{ij} \approx 5$  Å for the L7 and F8 samples. For 100% population of this antiparallel arrangement and  $r_{ij} \approx 5$  Å, the  $(\Delta S/S_0)_f^{\text{antipar}}$  will build up to  $\sim 0.8$  for  $\tau \geq 32$  ms [25,46]. However, for the L7 and F8 samples,  $(\Delta S/S_0)_f^\beta \approx 0.35$  suggesting that only  $\sim 0.4$  fraction of the FP-Hairpin molecules have these particular antiparallel registries. This result is consistent with the experimentally derived  $\sim 0.4$  fraction of these registries in membrane-associated HFP. It is likely that the remaining  $\sim 0.6$  fraction for FP-Hairpin and HFP are in antiparallel registries with adjacent strand crossing at residues other than L7 and F8.

In summary, observation of a strong dependence of  $(\Delta S/S_0)_f^\beta$  on *lab* site supports a significant fraction of antiparallel (Fig. 4b) but not parallel (Fig. 4a) structure for FP-Hairpin. The antiparallel structure of membrane-associated FP-Hairpin appears similar to that of membrane-associated HFP and supports the hypothesis of the FP as an autonomous  $\beta$  sheet folding unit in membranes.

### Effects of *na* $^{13}\text{C}$ nuclei

The analysis of the previous section is based on the reasonable approximations that the  $\beta$   $^{13}\text{C}$  signal is dominated by *lab* nuclei and that the  $(\Delta S/S_0)_f^\beta$  build-ups are dominated by *lab*  $^{13}\text{C}$ –*lab*  $^{13}\text{C}$  couplings. A more accurate analysis requires calculation of the effects of *na*  $^{13}\text{C}$  nuclei including (1) the contribution of these nuclei to the  $\beta$   $^{13}\text{C}$  signal and (2)  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings.

Effect (1) is calculated based on  $S_{0f}^{\text{na}}$  and  $S_{0f}^{\text{lab}}$  contributions to a total normalized  $S_{0f}^\beta$  intensity, that is,  $S_{0f}^\beta = S_{0f}^{\text{na}} + S_{0f}^{\text{lab}} = 1$ . The  $S_{0f}^{\text{na}}$  and  $S_{0f}^{\text{lab}}$  values are calculated using  $S_{0f}^{\text{na}} = (S_1/S_0)_r^\beta \equiv A$  and  $S_{0f}^{\text{lab}} = 1 - A$  where  $S_{0f}^\beta$  and  $S_{1f}^\beta$  are the experimental intensities in the  $\beta$  region of the REDOR  $S_{0f}$  and  $S_{1f}$  spectra. The  $A$  value for each sample is presented in Supplementary Information. This approach to calculation of  $S_{0f}^{\text{na}}$  and  $S_{0f}^{\text{lab}}$  is based on the reasonable approximations  $(S_1/S_0)_r^{\text{na}} = 1$  and  $(S_1/S_0)_r^{\text{lab}} = 0$  for samples with directly bonded labeled  $^{13}\text{C}/^{15}\text{N}$  spin pairs [34,35]. The same 1.0-ppm shift interval is used to calculate the  $S_{0f}^\beta$ ,  $S_{1f}^\beta$ ,  $S_{0f}^\alpha$ , and  $S_{1f}^\alpha$  intensities.

The contribution of  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings to the  $(\Delta S/S_0)_f^\beta$  data [effect (2) mentioned above] is assessed using the  $(\Delta S/S_0)_f^\alpha$  build-ups of the  $^{13}\text{C}$   $\alpha$  signals (Fig. 6). These  $(\Delta S/S_0)_f^\alpha$  build-ups are dominated by *na* couplings as evidenced by (1)  $(\Delta S/S_0)_f^\alpha \leq 0.1$  and (2)

$(\Delta S/S_0)_f^\alpha$  that are approximately independent of both the identity of the *lab* residue and the relative contributions of the *na* and *lab*  $^{13}\text{C}$  nuclei to the  $\alpha$  signal (Fig. 3).

For each  $\tau$ , the average  $(\Delta S/S_0)_f^\alpha$  of the membrane + D-malt samples serve as consensus values of  $(\Delta S/S_0)_f^\alpha$  due to  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings (red circles in Fig. 6f). There are similar average  $(\Delta S/S_0)_f^\alpha$  (to within 0.01) for the membrane samples (black circles in Fig. 6f), and the  $(\Delta S/S_0)_f^\alpha$  of each sample typically agrees within error with these averages. Because there are smaller uncertainties in the  $(\Delta S/S_0)_f^\alpha$  of the membrane + D-malt samples, the average  $(\Delta S/S_0)_f^\alpha(\tau) \equiv 1 - \gamma^{\text{na}}(\tau)$  of these samples are used to describe  $(\Delta S/S_0)_f^\alpha$  due to  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings. Table 3 lists these  $1 - \gamma^{\text{na}}(\tau)$  values along with the corollary  $\gamma^{\text{na}}(\tau) \equiv (S_1/S_0)_f^\alpha(\tau)$ . In the remainder of the paper, the notation is typically simplified by not explicitly writing “( $\tau$ )”. A model for  $1 - \gamma^{\text{na}}$  is described in Supplementary Information and yields calculated values in the 0.03–0.08 range that are consistent with the experimental values in Table 3.

The  $S_{0f}^\beta = S_{0f}^{\text{na}} + S_{0f}^{\text{lab}}$  and the  $S_{1f}^\beta = S_{1f}^{\text{na}} + S_{1f}^{\text{lab}}$ . The attenuation of the  $S_{1f}^{\text{na}}$   $\beta$   $^{13}\text{C}$  signal relative to the  $S_{0f}^{\text{na}}$   $\beta$   $^{13}\text{C}$  signal depends on *na*  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings, whereas attenuation of the  $S_{1f}^{\text{lab}}$   $\beta$   $^{13}\text{C}$  signal relative to  $S_{0f}^{\text{lab}}$   $\beta$   $^{13}\text{C}$  signal depends on *lab*  $^{13}\text{C}$ –*na*  $^{13}\text{C}$  couplings and *lab*  $^{13}\text{C}$ –*lab*  $^{13}\text{C}$  couplings. These *lab*  $^{13}\text{C}$ –*lab*  $^{13}\text{C}$  couplings are described by the parameter  $\gamma^{\text{lab}} = \gamma^{\text{lab}}(\tau)$ , that is, the  $(S_1/S_0)_f$  ratio for *lab*  $^{13}\text{C}$  nuclei coupled to other *lab*  $^{13}\text{C}$  nuclei in the  $\beta$  sheet. The variation of  $\gamma^{\text{lab}}$  [or equivalently  $1 - \gamma^{\text{lab}} \equiv (\Delta S/S_0)_f^{\text{lab}}$ ] with *lab* residue provides detailed information about the populations of parallel and antiparallel  $\beta$  strand arrangements (Fig. 4).

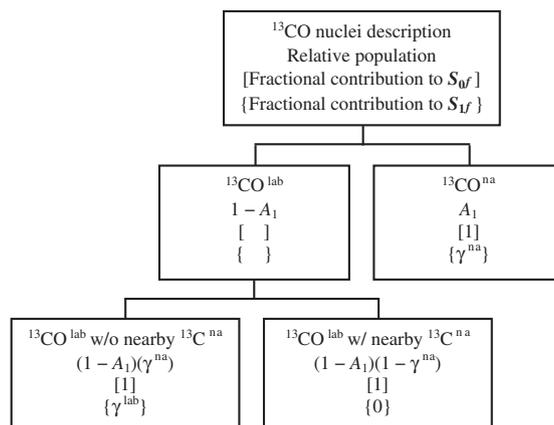
The  $\gamma^{\text{lab}}$  for each sample are calculated using a model outlined in Fig. 7. Aspects of the model include (1) quantitative separation of the  $S_{0f}^\beta$  signal into  $S_{0f}^{\text{na}}$  and  $S_{0f}^{\text{lab}}$  fractions using  $(S_1/S_0)_r^\beta \equiv A$  and  $1 - A$  as described previously, (2) further separation of  $S_{0f}^{\text{lab}}$  into fractions for which the *lab*  $^{13}\text{C}$  nuclei are respectively close and far from *na*  $^{13}\text{C}$  nuclei, and (3) calculation of  $\gamma^{\text{lab}}$  using (1) and (2) and the experimentally determined  $\gamma^{\text{na}}$  and  $(\Delta S/S_0)_f^\beta$ . For

**Table 3.**  $\gamma^{\text{na}}$  of FP-Hairpin

$\tau$ (ms)	$\gamma^{\text{na}}$	$1 - \gamma^{\text{na}}$
8	0.96	0.04
16	0.93	0.07
24	0.91	0.09
32	0.92	0.08
40	0.94	0.06

$\gamma^{\text{na}} \equiv (S_1/S_0)_f$  ratio from attenuation due to couplings to nearby  $^{13}\text{C}$  nuclei. This ratio is the average of the experimental  $(S_1/S_0)_f^\alpha$  ratios of the membrane + D-malt samples.

$1 - \gamma^{\text{na}} \equiv (\Delta S/S_0)_f$  ratio from attenuation due to couplings to nearby  $^{13}\text{C}$  nuclei.



**Fig. 7.** Flow chart for the calculation of  $\gamma^{lab}$  that reflects intermolecular *lab*  $^{13}\text{CO}$ –*lab*  $^{13}\text{CO}$  dipolar couplings and distances.

the same labeling, there are similar  $(\Delta S/S_0)_f^\beta$  build-ups for membrane and membrane + D-malt samples and the  $\gamma^{lab}$  values are calculated for the latter data that have larger and narrower signals.

This model results in an equation for the  $S_{0f}^\beta$  signal:

$$S_{0f}^\beta = 1 = A + [(1-A) \times (1-\gamma^{na})] + [(1-A) \times \gamma^{na}] \quad (1)$$

The left-most  $A$  term is the signal of  $na$   $^{13}\text{CO}$  nuclei and the  $[(1-A) \times (1-\gamma^{na})]$  and  $[(1-A) \times \gamma^{na}]$  terms are the signals of *lab*  $^{13}\text{CO}$  nuclei close to and far from  $na$   $^{13}\text{C}$  nuclei, respectively. The  $\gamma = (S_1/S_0)_f$  for each of these terms are  $\gamma^{na}$ , 0, and  $\gamma^{lab}$ , respectively. The  $S_{1f}^\beta$  signal:

$$S_{1f}^\beta = [A \times \gamma^{na}] + [(1-A) \times \gamma^{na} \times \gamma^{lab}] \quad (2)$$

The  $S_{0f}^\beta$  and  $\gamma^{lab}$  depend on  $\tau$  and labeled  $^{13}\text{CO}$  site. The quantity  $1 - \gamma^{lab}$  is the  $(\Delta S/S_0)_f$  for *lab*  $\beta$  sheet  $^{13}\text{CO}$  nuclei close to other *lab*  $^{13}\text{CO}$  nuclei and not close to  $na$   $^{13}\text{C}$  nuclei. As noted above,  $1 - \gamma^{lab}$  is diagnostic of the FP antiparallel  $\beta$  sheet strand registry distribution. Eqs. (1) and (2) are algebraically combined:

$$1 - \gamma^{lab} = \left\{ 1 - \left[ (S_1/S_0)_f^\beta / \gamma^{na} \right] \right\} / (1 - A_1) \quad (3)$$

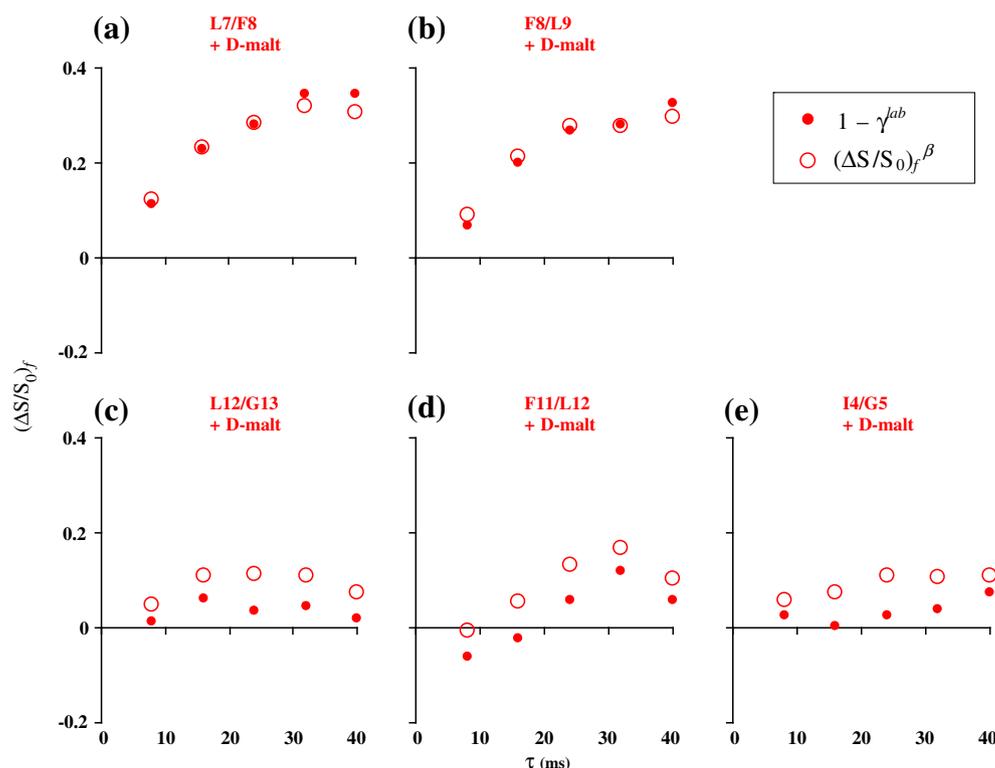
Figure 8 displays the  $(\Delta S/S_0)_f^\beta$  (open circles) and the  $1 - \gamma^{lab}$  (filled circles) versus  $\tau$  for the differently labeled membrane + D-malt samples. The  $1 - \gamma^{lab}$  of the (a) L7 and (b) F8 samples reach  $\sim 0.35$  for  $\tau \geq 32$  ms, and there are only small differences with the experimental  $(\Delta S/S_0)_f^\beta$ ; that is, consideration of  $na$   $^{13}\text{C}$  effects has little impact on the analysis. The  $1 - \gamma^{lab}$  of the (c) L12, (d) F11, and (e) I4 samples are close to 0 and support negligible in-register parallel  $\beta$  sheet FP structure in FP-Hairpin. Consideration of  $na$   $^{13}\text{C}$  effects for these samples results in reduced  $1 - \gamma^{lab}$  relative to  $(\Delta S/S_0)_f^\beta$ . The  $1 - \gamma^{lab} \approx 0.35$  values at  $\tau \geq 32$  ms for the L7 and F8 samples support a significant antiparallel  $\beta$

sheet population ( $\sim 0.4$  fraction) with adjacent strand crossing near L7 and F8. This antiparallel  $\beta$  sheet structure is similar to that reported for membrane-associated HFP that contains the 23 N-terminal residues of gp41. These findings suggest that membrane-associated antiparallel  $\beta$  sheet FP oligomers are a stable structure either with or without the rest of the ectodomain. To our knowledge, this is the first higher-resolution study of the tertiary structure of the FP region of any larger gp41 ectodomain construct.

FP-Hairpin is a model for the final SHB gp41 state during HIV/host cell membrane fusion. The SHB is a molecular trimer with the three FPs extending in parallel from the SHB. Our detection of FPs with oligomeric antiparallel  $\beta$  sheet structure in FP-Hairpin indicates interleaving of FP strands between two or more trimers (Fig. 9). The membrane location of the FP in Fig. 9 is based on earlier SSNMR data for HFP [21]. The SHB is shown away from the membrane because both the SHB and the membrane are negatively charged at pH 7 and there may be considerable motion of the SHB [18,30].

### REDOR and fpCTDQBU spectra and analysis of FP34 and N70 in membranes

FP34 is the 34 N-terminal residues of gp41 and is an FP model while N70 is the 70 N-terminal residues and is an *N*-PHI model (Fig. 2). Both constructs bind anionic membranes and induce vesicle fusion at physiologic pH [15,32]. N70 is significantly more fusogenic than FP34, which may be due to formation of oligomers of N70 with parallel helical coiled coils in the *N*-helix region and consequent clusters of FPs similar in the model PHI structure. Figure 10 displays REDOR spectra and fpCTDQBU  $(\Delta S/S_0)_f^\beta$  build-ups of L7- $^{13}\text{CO}$ /F8- $^{15}\text{N}$  labeled FP34 and N70 in membranes. The  $^{13}\text{CO}$   $\Delta S_r$  spectrum of the FP34 sample (Fig. 10a-1, inset) has a peak shift of 174 ppm that is consistent with predominant  $\beta$  conformation at L7. The

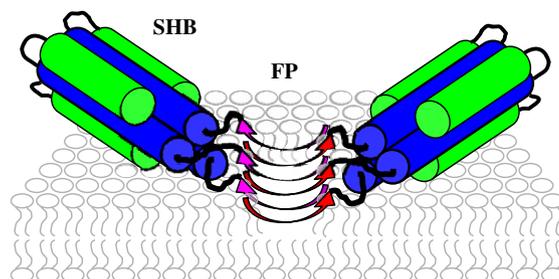


**Fig. 8.**  $(\Delta S/S_0)_f^\beta$  (open circles) and  $1 - \gamma^{lab}$  (filled circles) versus dephasing time for membrane + D-malt samples with FP-Hairpin  $^{13}\text{C}$  labeled at (a) L7, (b) F8, (c) L12, (d) F11, and (e) I4.

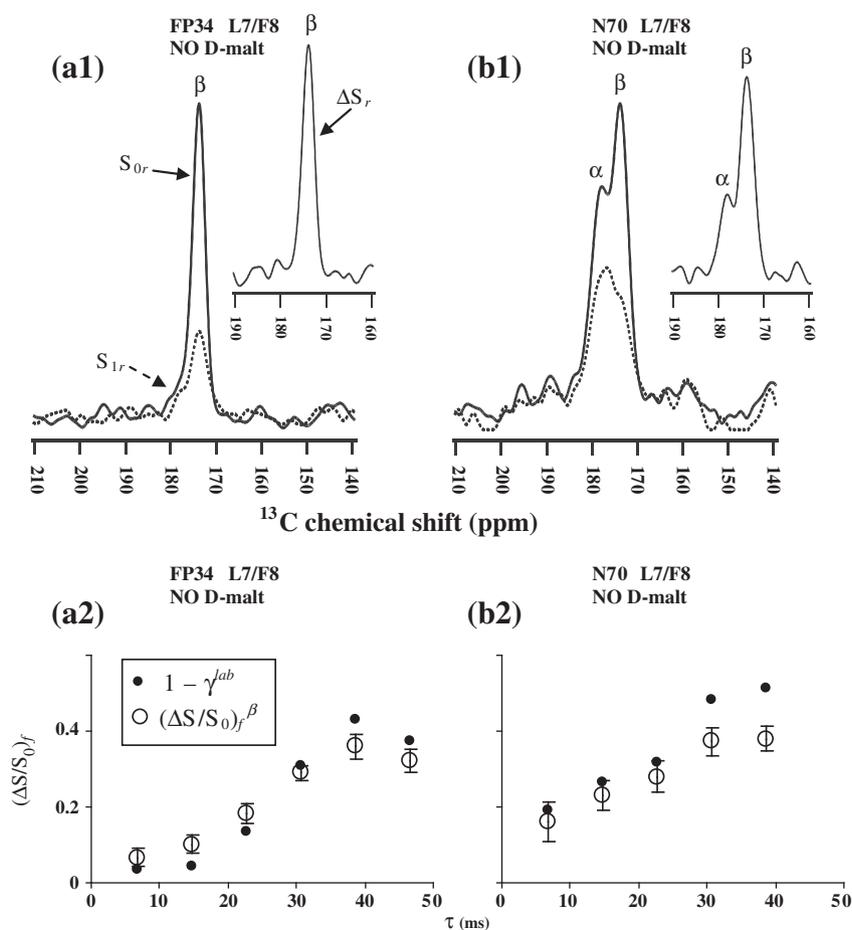
$^{13}\text{C}$   $\Delta S_r$  spectrum of the N70 sample (Fig. 10b-1, inset) has two peaks with shifts of 174 and 178 ppm that are respectively assigned to populations of N70 molecules with  $\beta$  and  $\alpha$  conformations at L7. Figure 10a-2 and b-2 displays plots of  $(\Delta S/S_0)_f^\beta$  (open circles) and  $1 - \gamma^{lab}$  (filled circles) for the two samples. Determination of  $1 - \gamma^{lab}$  is performed as described for FP-Hairpin. The analysis uses the  $\gamma^{na}$  from Table 3 with  $\gamma^{na} = 0.94$  for  $\tau \geq 38$  ms. The fpCTDQBU spectra are shown in Supplementary Information. The  $(\Delta S/S_0)_f^\beta$  build-ups of the L7-labeled FP34, N70, and FP-Hairpin samples are all comparable with values of  $\sim 0.35$  for  $\tau \geq 32$  ms. The  $1 - \gamma^{lab}$  are also

comparable and are consistent with  $\sim 0.4$  fraction of adjacent strands crossing near L7 in the  $\beta$  sheet as has also been observed for HFP. Earlier HFP data and FP-Hairpin data from the present paper are consistent with antiparallel rather than parallel arrangement of adjacent FP strands in the  $\beta$  sheet.

We note that different earlier infrared studies of FP34 and N70 in membranes had been interpreted to support in-register parallel  $\beta$  sheet structure of the FP [48]. This interpretation is based on ascribing spectral changes to intermolecular rather than intramolecular electric dipole–electric dipole couplings. However, the peptides contained  $^{13}\text{C}$  labels



**Fig. 9.** Schematic of membrane-associated FP-Hairpin at pH 7. There is an SHB and antiparallel  $\beta$  sheet FP structure so FP-Hairpin minimally oligomerizes as a dimer of trimers. The FP membrane location is chosen to be the same as the 23-residue HFP that lacks the SHB. The FP region binds negatively charged membranes at pH 7, whereas the negatively charged SHB region may not bind these membranes.



**Fig. 10.** SSNMR spectra of membrane samples containing (a-1 and a-2) FP34 and (b-1 and b-2) N70 with a L7 <sup>13</sup>CO/F8 <sup>15</sup>N spin pair. (a-1) and (b-1) display REDOR total  $S_{0r}$ ,  $na S_{1r}$ , and  $lab \Delta S_r$  spectra as continuous, dotted, and inset lines, respectively. Each  $S_{0r}$  and  $S_{1r}$  spectrum is the sum of (a-1) 29,136 and (b-1) 76,832 scans. Spectra are processed with 200 Hz Gaussian line broadening and baseline correction. (a-2) and (b-2) display fpCTDQBU  $(\Delta S/S_0)_f^\beta$  (open circles) and  $1 - \gamma^{lab}$  (filled circles) versus dephasing time. Each  $S_{0f}$  and  $S_{1f}$  spectrum is the sum of (a-2) 20,000 and (b-2) 16,000 scans.

at sequential residues and the large intramolecular couplings could not be distinguished from intermolecular couplings.

In our view, the site-specific <sup>13</sup>CO labeling of the present study was critical to provide unambiguous information about the FP regions of these larger membrane-associated gp41 constructs. Relative to uniform <sup>13</sup>C labeling, there is spectral simplification with the specific labeling so that the peaks can be unambiguously assigned [49]. This simplification is important because of the ~3 ppm <sup>13</sup>CO linewidths (Fig. 3) and the structural heterogeneity in the FP region including a mixture of β and α populations and a distribution of antiparallel strand registries.

## Conclusions

SSNMR is applied to probe FP conformation and intermolecular FP β sheet structure of N70 and

FP-Hairpin gp41 constructs that respectively model the putative early PHI and final SHB gp41 states during fusion. The shorter FP-only FP34 construct is also studied. The samples are prepared with membranes. FP-Hairpin is fusion inactive at physiologic pH, whereas N70 and FP34 are highly and moderately active, respectively. There are typically positive correlations between membrane cholesterol content and (1) β population and (2) rate and extent of vesicle fusion [17,29,32]. The β and α structures have also been observed for the membrane-associated structures of FPs of other viral fusion proteins [28,34,41–43]. Our main findings and interpretations are summarized.

- (1) FP-Hairpin contains *N*- and *C*-helices that are *C*-terminal of the FP and are likely part of an SHB structure. There are two populations of FP-Hairpin with either β sheet or α helical FP structure.

- (2) Relative to initial solubilization of FP-Hairpin in Buffer, solubilization in Buffer + D-malt results in a sample with increased population with  $\beta$  FP. D-malt is expected to reduce aggregation of FP-Hairpin; thus, the  $\beta$  and  $\alpha$  FP populations may respectively correlate with smaller and larger FP-Hairpin oligomers/aggregates in membranes.
- (3) The FP-Hairpin  $^{13}\text{C}$   $\alpha$  peaks are dominated by  $na$  signals of the  $N$ - and  $C$ -helices. The fpCTDQBU intensity ratios  $(S_1/S_0)_f^\alpha = \gamma^{na} = \gamma^{na}(\tau)$  are consistent with the dominance of couplings to  $na$   $^{13}\text{C}$  nuclei. The samples have a directly bonded  $^{13}\text{C}/^{15}\text{N}$  spin pair and the REDOR intensity ratio in the  $\beta$  region,  $A = (S_1/S_0)_f^\beta$ , and  $1 - A$  are the fractional  $na$  and  $lab$  contributions to  $S_{0f}^\beta$ , respectively. The  $A$ ,  $\gamma^{na}$ , and  $(S_1/S_0)_f^\beta$  values are used to calculate  $\gamma^{lab} = \gamma^{lab}(\tau)$ , the  $(S_1/S_0)_f^\beta$  ratio of  $lab$  FP  $\beta$   $^{13}\text{C}$  nuclei coupled to other  $lab$  FP  $\beta$   $^{13}\text{C}$  nuclei. The variation of  $\gamma^{lab}$  values with  $lab$  site supports negligible FP population with in-register parallel  $\beta$  sheet structure and a significant population ( $> 1/3$  fraction) with antiparallel structure and adjacent strand crossing near L7 and F8 in the most apolar part of the FP. The remaining population is likely in antiparallel registries with adjacent strand crossing at residues other than L7 and F8. There may be interleaving of FP strands between two FP-Hairpin trimers (Fig. 9).
- (4) The antiparallel  $\beta$  sheet FP structural model for FP-Hairpin in membranes is similar to the FP structure of membrane-associated HFP that only contains the 23 N-terminal residues of the gp41 protein. The FP sequence in membranes may therefore be an autonomous antiparallel  $\beta$  sheet folding unit.
- (5) The L7  $^{13}\text{C}$ -labeled samples containing either FP-Hairpin, N70, or FP34 all have comparable  $(\Delta S/S_0)_f^\beta$  build-ups that indicate a significant population with strands crossing near L7. Like FP-Hairpin that models the final SHB gp41 state, FP34 and N70 (which models the intermediate PHI state) likely have antiparallel rather than parallel  $\beta$  sheet structure.
- (6) N70 induces rapid vesicle fusion at pH 7, whereas FP-Hairpin is fusion inactive. This functional difference might be explained by (1) deeper FP membrane insertion of N70 with consequent greater membrane perturbation

and/or (2) a somewhat different distribution of antiparallel FP registries for N70 that is correlated with higher hydrophobicity of the  $\beta$  sheet and/or (3) differential membrane perturbation by the  $N$ -helix of N70 versus the SHB of FP-Hairpin. Earlier studies on different HFP constructs showed that there is a strong positive correlation between depth of FP membrane insertion and vesicle fusion rate [21].

## Experimental Procedures

### Materials

Boc and Fmoc amino acids, Boc-MBHA resin, and Fmoc-rink amide MBHA resin were purchased from Novabiochem (San Diego, CA). S-Trityl- $\beta$ -mercapto-propionic acid was purchased from Peptides International (Louisville, KY). Di-*t*-butyl-dicarbonate, tris(2-carboxyethyl) phosphine hydrochloride, 4-mercaptophenylacetic acid, 2-mercaptoethanesulfonic acid sodium salt, and  $N$ -(2-hydroxyethyl) piperazine- $N'$ -2-ethanesulfonic acid (Hepes) were purchased from Sigma. 1,2-Di-*O*-tetradecyl-*sn*-glycero-3-phosphocholine (DTPC), 1,2-di-*O*-tetradecyl-*sn*-glycero-3-phospho-*rac*-(1-glycerol) sodium salt (DTPG), and cholesterol were purchased from Avanti Polar Lipids (Alabaster, AL). Labeled amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA) and were Boc protected using literature methods [50]. The micro bicinchoninic acid (BCA) protein assay was obtained from Pierce (Rockford, IL). "Buffer" refers to 10 or 20 mM formate buffer and 200  $\mu\text{M}$  tris(2-carboxyethyl) phosphine hydrochloride (pH 3).

### Protein preparation

The sequences (Fig. 1) are from the HXB2 laboratory strain of the HIV-1 Envelope protein with names and residue numbering as follows: (a) FP34(linker), 512-545-(thioester); (b) FP23(linker), 512-534(S534A)-(thioester); (c) FP34, 512-545; (d) N36, 546-581(S546C); and (e) N47(L6)C39 also known as "Hairpin", 535(M535C)-581 + non-native SGGRRG loop + 628-666. Chemical syntheses of (a)–(d) have been previously described and usually included a single directly bonded  $^{13}\text{C}/^{15}\text{N}$  spin pair [51]. Hairpin was produced recombinantly in bacteria with subsequent dialysis of the soluble lysate against 0.05% trifluoroacetic acid + 150  $\mu\text{M}$  dithiothreitol, and then filtration and concentration [15,52]. All proteins were purified by reverse-phase HPLC and quantitated by (a)–(d) BCA assay or (e)  $A_{280}$ . Mass spectra were consistent with ~95% protein purity.

### Native chemical ligation

N70 was prepared by ligating FP34(linker) with N36(S546C) as described previously [51]. FP-Hairpin was prepared by ligating FP23(linker) with purified Hairpin at ambient temperature in a solution containing 8 M guanidine hydrochloride and either 30 mM 4-mercaptophenylacetic acid or 30 mM 2-mercaptoethanesulfonic acid, sodium salt [15]. N70 and FP-Hairpin were purified by reverse-phase HPLC and eluted as single peaks and were respectively quantitated using the BCA assay and  $A_{280}$ . N70 was lyophilized and FP-Hairpin was dialyzed against Buffer at 4 °C.

### Lipid vesicle preparation

Large unilamellar vesicles were prepared in 25 mM Hepes (pH 7.4) buffer by extrusion through 100-nm-diameter pores. The typical vesicle composition was DTPC:DTPG:cholesterol at 1.6:0.4:1.0 mM, which reflects dominant choline headgroup, significant negatively charged lipid, and approximate lipid:cholesterol ratio found in viral and host cell plasma membranes [22]. The  $na$   $^{13}\text{C}$  signal was reduced with use of DTPC and DTPG lipids that lack CO nuclei. For membrane-associated HFP, similar FP structure was observed with either ether- or ester-linked lipids [25]. FP-Hairpin binds to membranes with this headgroup and cholesterol composition [18].

### NMR sample preparation

The protein solution was added dropwise to the vesicle suspension that maximized membrane binding of the protein and minimized protein aggregation. The mixture was stirred over the addition time of ~1 h and the pH was maintained above 7 with addition of pH 7.4 buffer. The initial protein solution either contained Buffer with pH  $\approx$  3 and [protein] = 40  $\mu\text{M}$  (membrane samples) or [protein] = 80  $\mu\text{M}$  and [D-malt] = 9 mM (membrane + D-malt samples). A typical mixture contained [lipid + cholesterol]  $\approx$  1100  $\mu\text{M}$  and [protein]  $\approx$  25  $\mu\text{M}$  with ~0.7  $\mu\text{mol}$  protein. Supplementary Information provides the specific molar quantity protein and lipid:protein mole ratio for each sample. Proteoliposomes settled during overnight incubation at 4 °C with subsequent centrifugation first at 4000g and then at either 16,000 or 100,000g. The pellet was transferred to a 4-mm-diameter magic angle spinning (MAS) rotor.

### General SSNMR

Experiments were performed on a 9.4-T (400-MHz) spectrometer (Agilent Infinity Plus, Palo Alto, CA) using a MAS probe in triple resonance  $^1\text{H}/^{13}\text{C}/^{15}\text{N}$

configuration for REDOR experiments and double resonance  $^1\text{H}/^{13}\text{C}$  configuration for fpCTDQBU experiments. The  $^{13}\text{C}$  shifts were externally referenced to the methylene resonance of adamantane at 40.5 ppm that allows direct comparison with database  $^{13}\text{C}$  shifts [53]. Samples were cooled with nitrogen gas at -50 °C because, relative to ambient temperature, signal-to-noise was improved without changing FP structure [54]. The  $^{13}\text{C}$  channel was tuned to 100.8 MHz, the  $^1\text{H}$  channel was tuned to 400.8 MHz, and the  $^{15}\text{N}$  channel was tuned to 40.6 MHz. The  $^1\text{H}$  and  $^{13}\text{C}$  rf fields were optimized with polycrystalline samples containing either *N*-acetyl-leucine (NAL) or GFF tripeptide. NAL was crystallized by slow evaporation of a solution containing doubly  $^{13}\text{C}$ -labeled and unlabeled molecules in a 1:9 ratio and GFF tripeptide was crystallized from doubly  $^{13}\text{C}$ -labeled (G1 + F3) and unlabeled molecules in a 1:50 ratio [25,46]. The  $^{15}\text{N}$   $\pi$  pulse was optimized using the lyophilized helical "I4" peptide with A9  $^{13}\text{C}$  and A13  $^{15}\text{N}$  labels [25].

### REDOR SSNMR

The REDOR pulse sequence begins with generation of  $^{13}\text{C}$  transverse magnetization followed by a dephasing period and then  $^{13}\text{C}$  detection. During the 2-ms dephasing period of both the  $S_{0r}$  and  $S_{1r}$  acquisitions, there was a  $^{13}\text{C}$   $\pi$  pulse at the end of each rotor cycle except the last cycle. For the  $S_{1r}$  acquisition, a  $^{15}\text{N}$   $\pi$  pulse was included at the midpoint of each cycle, and the two  $\pi$  pulses per cycle resulted in near-complete attenuation of the signal of directly bonded  $^{13}\text{C}/^{15}\text{N}$  spin pairs. The  $S_{0r}$ ,  $S_{1r}$  and  $\Delta S_r = S_{0r} - S_{1r}$  signals respectively correspond to the total,  $na$ , and  $lab$   $^{13}\text{C}$  signals [34]. The  $^{13}\text{C}$  transmitter shift (153 ppm), the MAS frequency (8.0 kHz), and pulse sequence times included the following: cross-polarization (CP), 1.6 ms; dephasing, 2.0 ms; and recycle delay, 1.0 s. Typical rf fields were as follows:  $^{13}\text{C}$   $\pi$ , 54 kHz;  $^{13}\text{C}$  CP ramp, 58–68 kHz;  $^{15}\text{N}$   $\pi$ , 43 kHz;  $^1\text{H}$   $\pi/2$  and CP, 50 kHz; and  $^1\text{H}$  decoupling, 90 kHz. XY-8 phase cycling was applied to both  $^{15}\text{N}$  and  $^{13}\text{C}$   $\pi$  pulses except for the final  $^{13}\text{C}$   $\pi$  pulse. Pulses were calibrated as previously described [25].

### fpCTDQBU SSNMR

The fpCTDQBU pulse sequence is sequential generation of  $^{13}\text{C}$  transverse magnetization, a constant-time period of duration  $CT$  split between a variable dephasing period of duration  $\tau$  and a variable echo period of duration  $CT - \tau$ , and  $^{13}\text{C}$  detection. One finite  $^{13}\text{C}$   $\pi$  pulse per rotor cycle is applied during the  $CT$  period, that is, the fpRFDR sequence, and there is a pair of back-to-back  $^{13}\text{C}$   $\pi/2$  pulses at the midpoints of both the dephasing and the echo periods. For  $S_{0r}$ , the phases of the  $\pi/2$  pulses of each pair are

offset by 90° and resulted in refocusing of  $^{13}\text{CO}-^{13}\text{C}$  dipolar evolution. For  $S_{1f}$ , the 90° phase offset is retained for the echo period pair, whereas for the dephasing period pair, either 0° or 180° offset is used and results in dipolar evolution during this period and attenuation of the  $^{13}\text{CO}$  signal intensity. Because  $CT$  is the same for all  $\tau$ , couplings other than  $^{13}\text{CO}-^{13}\text{C}$  and relaxation could be neglected and the build-up of  $(S_{0f} - S_{1f})/S_{0f} = (\Delta S/S_0)_f$  with  $\tau$  is attributed to  $^{13}\text{CO}-^{13}\text{C}$  couplings. The fpCTDQBU pulses were calibrated as previously described, and the experiment was verified by good agreement between the  $(\Delta S/S_0)_f$  versus  $\tau$  for NAL and GFF and the  $(\Delta S/S_0)_f$  calculated based on the known intramolecular  $lab$   $^{13}\text{CO}-lab$   $^{13}\text{CO}$  distances of 3.1 and 5.4 Å, respectively [25,46]. Couplings due to intermolecular  $lab$   $^{13}\text{CO}-lab$   $^{13}\text{CO}$  distances could be neglected because of dilution with unlabeled molecules.

Typical fpCTDQBU experimental parameters included the following: MAS frequency, 12.0 kHz;  $^{13}\text{C}$  transmitter shift, 163 ppm; CP, 2.0 ms;  $CT$ , 41.33 ms;  $8\text{ ms} \leq \tau \leq 40\text{ ms}$ ; recycle delay, 1.5 s; 50–60 kHz ramped  $^{13}\text{C}$  and constant 56 kHz  $^1\text{H}$  fields during CP; 22 kHz  $^{13}\text{C}$   $\pi$  pulse and 50 kHz  $^{13}\text{C}$   $\pi/2$  pulse fields; and 90 and 60 kHz  $^1\text{H}$  decoupling fields during the  $CT$  and acquisition periods, respectively. The  $(\Delta S/S_0)_f^\beta$  and  $(\Delta S/S_0)_f^\alpha$  were determined from spectral integrations in 1.0 ppm windows centered at the peak  $\beta$  or  $\alpha$  shifts. The  $\sigma$  parameter was the root-mean-square deviation of the intensities of 24 separate 1.0-ppm windows in noise regions of the  $S_{0f}$  and  $S_{1f}$  spectra. The uncertainty ( $\sigma^{\text{exp}}$ ) in  $(\Delta S/S_0)_f^{\text{exp}}$  was calculated:

$$\sigma^{\text{exp}} = \frac{\sigma}{S_{0f}} \sqrt{1 + \frac{S_{1f}^2}{S_{0f}^2}} = \frac{\sigma S_{1f}}{S_{0f}} \sqrt{\frac{1}{S_{1f}^2} + \frac{1}{S_{0f}^2}} \quad (4)$$

## Acknowledgements

The Michigan State University Mass Spectrometry and NMR facilities are acknowledged. The work was supported by National Institutes of Health awards R01AI047153 to D.P.W. and F32AI080136 to K.S. We acknowledge Yechiel Shai for tBoc peptide synthesis and cleavage facilities that were partially supported by the Israel Science Foundation.

## Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2013.11.010>.

Received 21 July 2013;

Received in revised form 18 October 2013;

Accepted 11 November 2013

Available online 16 November 2013

## Keywords:

HIV;

gp41;

structure;

fusion peptide;

NMR

## Abbreviations used:

BCA, bicinechonic acid; CP, cross-polarization; DTPC, 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine; DTPG, 1,2-di-O-tetradecyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt; FP, fusion peptide; fpCTDQBU, finite-pulse constant-time double-quantum build-up; MAS, magic angle spinning; PHI, pre-hairpin intermediate; REDOR, rotational-echo double-resonance; SHB, six-helix bundle; SSNMR, solid-state NMR.

## References

- [1] White JM, Delos SE, Brecher M, Schornberg K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* 2008;43:189–219.
- [2] Mao YD, Wang LP, Gu C, Herschhorn A, Xiang SH, Haim H, et al. Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer. *Nat Struct Mol Biol* 2012;19:893–9.
- [3] Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* 2008;455:109–13.
- [4] Tran EEH, Borgnia MJ, Kuybeda O, Schauder DM, Bartesaghi A, Frank GA, et al. Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. *PLoS Pathog* 2012;8:e1002797.
- [5] Tan K, Liu J, Wang J, Shen S, Lu M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci USA* 1997;94:12303–8.
- [6] Caffrey M, Cai M, Kaufman J, Stahl SJ, Wingfield PT, Covell DG, et al. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J* 1998;17:4572–84.
- [7] Yang ZN, Mueser TC, Kaufman J, Stahl SJ, Wingfield PT, Hyde CC. The crystal structure of the SIV gp41 ectodomain at 1.47 Å resolution. *J Struct Biol* 1999;126:131–44.
- [8] Sackett K, Wexler-Cohen Y, Shai Y. Characterization of the HIV N-terminal fusion peptide-containing region in context of key gp41 fusion conformations. *J Biol Chem* 2006;281:21755–62.
- [9] Freed EO, Delwart EL, Buchschacher GL, Panganiban AT. A mutation in the human immunodeficiency virus type 1 transmembrane glycoprotein gp41 dominantly interferes with fusion and infectivity. *Proc Natl Acad Sci USA* 1992;89:70–4.
- [10] Sougrat R, Bartesaghi A, Lifson JD, Bennett AE, Bess JW, Zabransky DJ, et al. Electron tomography of the contact between T cells and SIV/HIV-1: implications for viral entry. *PLoS Pathog* 2007;3:571–81.

- [11] Magnus C, Rusert P, Bonhoeffer S, Trkola A, Regoes RR. Estimating the stoichiometry of human immunodeficiency virus entry. *J Virol* 2009;83:1523–31.
- [12] Blumenthal R, Durell S, Viard M. HIV entry and envelope glycoprotein-mediated fusion. *J Biol Chem* 2012;287:40841–9.
- [13] Shu W, Liu J, Ji H, Radigen L, Jiang SB, Lu M. Helical interactions in the HIV-1 gp41 core reveal structural basis for the inhibitory activity of gp41 peptides. *Biochemistry* 2000;39:1634–42.
- [14] Markosyan RM, Cohen FS, Melikyan GB. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol Biol Cell* 2003;14:926–38.
- [15] Sackett K, Nethercott MJ, Shai Y, Weliky DP. Hairpin folding of HIV gp41 abrogates lipid mixing function at physiologic pH and inhibits lipid mixing by exposed gp41 constructs. *Biochemistry* 2009;48:2714–22.
- [16] Durell SR, Martin I, Ruyschaert JM, Shai Y, Blumenthal R. What studies of fusion peptides tell us about viral envelope glycoprotein-mediated membrane fusion. *Mol Membr Biol* 1997;14:97–112.
- [17] Yang R, Prorok M, Castellino FJ, Weliky DP. 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* 2004;126:14722–3.
- [18] Sackett K, TerBush A, Weliky DP. 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* 2011;40:489–502.
- [19] Pereira FB, Goni FM, Muga A, Nieva JL. Permeabilization and fusion of uncharged lipid vesicles induced by the HIV-1 fusion peptide adopting an extended conformation: dose and sequence effects. *Biophys J* 1997;73:1977–86.
- [20] Gabrys CM, Qiang W, Sun Y, Xie L, Schmick SD, Weliky DP. Solid-state nuclear magnetic resonance measurements of HIV fusion peptide  $^{13}\text{C}$  to lipid  $^{31}\text{P}$  proximities support similar partially inserted membrane locations of the  $\alpha$  helical and  $\beta$  sheet peptide structures. *J Phys Chem A* 2013;117:9848–59.
- [21] Qiang W, Sun Y, Weliky DP. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proc Natl Acad Sci USA* 2009;106:15314–9.
- [22] Brugger B, Glass B, Haberkant P, Leibrecht I, Wieland FT, Krasslich HG. The HIV lipidome: a raft with an unusual composition. *Proc Natl Acad Sci USA* 2006;103:2641–6.
- [23] Qiang W, Bodner ML, Weliky DP. 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* 2008;130:5459–71.
- [24] Schmick SD, Weliky DP. Major antiparallel and minor parallel beta sheet populations detected in the membrane-associated human immunodeficiency virus fusion peptide. *Biochemistry* 2010;49:10623–35.
- [25] Zheng Z, Yang R, Bodner ML, Weliky DP. Conformational flexibility and strand arrangements of the membrane-associated HIV fusion peptide trimer probed by solid-state NMR spectroscopy. *Biochemistry* 2006;45:12960–75.
- [26] Qiang W, Weliky DP. HIV fusion peptide and its cross-linked oligomers: efficient syntheses, significance of the trimer in fusion activity, correlation of  $\beta$  strand conformation with membrane cholesterol, and proximity to lipid headgroups. *Biochemistry* 2009;48:289–301.
- [27] Nieva JL, Agirre A. Are fusion peptides a good model to study viral cell fusion? *Biochim Biophys Acta* 2003;1614:104–15.
- [28] Epand RM. Fusion peptides and the mechanism of viral fusion. *Biochim Biophys Acta* 2003;1614:116–21.
- [29] Lai AL, Moorthy AE, Li YL, Tamm LK. Fusion activity of HIV gp41 fusion domain is related to its secondary structure and depth of membrane insertion in a cholesterol-dependent fashion. *J Mol Biol* 2012;418:3–15.
- [30] Lakomek NA, Kaufman JD, Stahl SJ, Louis JM, Grishaev A, Wingfield PT, et al. Internal dynamics of the homotrimeric HIV-1 viral coat protein gp41 on multiple time scales. *Angew Chem Int Ed Engl* 2013;52:3911–5.
- [31] Grasnick D, Sternberg U, Strandberg E, Wadhvani P, Ulrich AS. Irregular structure of the HIV fusion peptide in membranes demonstrated by solid-state NMR and MD simulations. *Eur Biophys J* 2011;40:529–43.
- [32] Sackett K, Nethercott MJ, Epand RF, Epand RM, Kindra DR, Shai Y, et al. 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* 2010;397:301–15.
- [33] Gullion T, Schaefer J. Rotational-echo double-resonance NMR. *J Magn Reson* 1989;81:196–200.
- [34] Yang J, Parkanzky PD, Bodner ML, Duskin CG, Weliky DP. Application of REDOR subtraction for filtered MAS observation of labeled backbone carbons of membrane-bound fusion peptides. *J Magn Reson* 2002;159:101–10.
- [35] Gullion T. Introduction to rotational-echo, double-resonance NMR. *Concepts Magn Reson* 1998;10:277–89.
- [36] Murphy OJ, Kovacs FA, Sicard EL, Thompson LK. Site-directed solid-state NMR measurement of a ligand-induced conformational change in the serine bacterial chemoreceptor. *Biochemistry* 2001;40:1358–66.
- [37] Zhang HY, Neal S, Wishart DS. RefDB: a database of uniformly referenced protein chemical shifts. *J Biomol NMR* 2003;25:173–95.
- [38] Vogel EP, Curtis-Fisk J, Young KM, Weliky DP. 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. *Biochemistry* 2011;50:10013–26.
- [39] Yang J, Gabrys CM, Weliky DP. Solid-state nuclear magnetic resonance evidence for an extended beta strand conformation of the membrane-bound HIV-1 fusion peptide. *Biochemistry* 2001;40:8126–37.
- [40] Yang J, Prorok M, Castellino FJ, Weliky DP. Oligomeric  $\beta$ -structure of the membrane-bound HIV-1 fusion peptide formed from soluble monomers. *Biophys J* 2004;87:1951–63.
- [41] Wasniewski CM, Parkanzky PD, Bodner ML, Weliky DP. Solid-state nuclear magnetic resonance studies of HIV and influenza fusion peptide orientations in membrane bilayers using stacked glass plate samples. *Chem Phys Lipids* 2004;132:89–100.
- [42] Pan JH, Lai CB, Scott WRP, Straus SK. Synthetic fusion peptides of tick-borne encephalitis virus as models for membrane fusion. *Biochemistry* 2010;49:287–96.
- [43] Yao HW, Hong M. Membrane-dependent conformation, dynamics, and lipid interactions of the fusion peptide of the paramyxovirus PIV5 from solid-state NMR. *J Mol Biol* 2013;425:563–76.
- [44] Gullion T, Vega S. A simple magic angle spinning NMR experiment for the dephasing of rotational echoes of dipolar coupled homonuclear spin pairs. *Chem Phys Lett* 1992;194:423–8.

- [45] Bennett AE, Weliky DP, Tycko R. Quantitative conformational measurements in solid state NMR by constant-time homonuclear dipolar recoupling. *J Am Chem Soc* 1998;120:4897–8.
- [46] Zheng Z, Qiang W, Weliky DP. Investigation of finite-pulse radiofrequency-driven recoupling methods for measurement of intercarbonyl distances in polycrystalline and membrane-associated HIV fusion peptide samples. *Magn Reson Chem* 2007;245:S247–60.
- [47] Ishii Y.  $^{13}\text{C}$ - $^{13}\text{C}$  dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J Chem Phys* 2001;114:8473–83.
- [48] Sackett K, Shai Y. The HIV fusion peptide adopts intermolecular parallel  $\beta$ -sheet structure in membranes when stabilized by the adjacent N-terminal heptad repeat: a  $^{13}\text{C}$  FTIR study. *J Mol Biol* 2005;350:790–805.
- [49] Bodner ML, Gabrys CM, Struppe JO, Weliky DP.  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  correlation spectroscopy of membrane-associated and uniformly labeled HIV and influenza fusion peptides: amino acid-type assignments and evidence for multiple conformations. *J Chem Phys* 2008;128:052319.
- [50] Harris RB, Wilson IB. Tert-butyl aminocarbonate (tert-butyloxycarbonyloxyamine)—a new acylating reagent for amines. *Int J Pept Protein Res* 1984;23:55–60.
- [51] Sackett K, Shai Y. The HIV-1 gp41 N-terminal heptad repeat plays an essential role in membrane fusion. *Biochemistry* 2002;41:4678–85.
- [52] Curtis-Fisk J, Spencer RM, Weliky DP. Isotopically labeled expression in *E. coli*, purification, and refolding of the full ectodomain of the influenza virus membrane fusion protein. *Protein Expression Purif* 2008;61:212–9.
- [53] Morcombe CR, Zilm KW. Chemical shift referencing in MAS solid state NMR. *J Magn Reson* 2003;162:479–86.
- [54] Bodner ML, Gabrys CM, Parkanzky PD, Yang J, Duskin CA, Weliky DP. Temperature dependence and resonance assignment of  $^{13}\text{C}$  NMR spectra of selectively and uniformly labeled fusion peptides associated with membranes. *Magn Reson Chem* 2004;42:187–94.

Supplementary Information for “Solid-State NMR Spectroscopy of the HIV gp41 Membrane Fusion Protein Supports Intermolecular Antiparallel  $\beta$  Sheet Fusion Peptide Structure in the Final Six-Helix Bundle State”

**Model of  $(\Delta S/S_0)_f$  from couplings to  $na$   $^{13}\text{C}$  nuclei**

Table III lists the  $(S_1/S_0)_f^\alpha$  of FP-Hairpin samples which are used to estimate  $\gamma^{na} \equiv (S_1/S_0)^f$  due to couplings to  $na$   $^{13}\text{C}$  nuclei. An earlier model is applied to understand these experimental  $\gamma^{na}$  or equivalently the  $(\Delta S/S_0)_f^\alpha = 1 - \gamma^{na}$ , cf. Z. Zheng et. al. *Biochemistry*, (2006) **45**, 12960-12975 and Z. Zheng et. al. *Magn. Reson. Chem.* (2007) **45**, S247-S260. Consider a sphere of radius  $r$  around a  $^{13}\text{CO}$  nucleus where  $(S_1/S_0)_f \approx 0$  if there is a  $na$   $^{13}\text{C}$  nucleus within this sphere. The value of  $r$  that satisfies this criterion is estimated from density operator-based simulations of  $(S_1/S_0)_f$  for a single  $^{13}\text{CO}/^{13}\text{C}$  spin pair as a function of internuclear distance. The number of C nuclei within this sphere ( $\equiv p$ ) is estimated from the geometric arrangement of C nuclei around  $^{13}\text{CO}$  nucleus. Given the  $na$   $^{13}\text{C}$  probability = 0.011, the model estimates that  $1 - \gamma^{na} \approx 0.011 \times p$ . Both  $r$  and  $p$  will increase with dephasing time. Previous calculations are consistent with  $p$  values ranging between  $\sim 3$  for  $\tau = 8$  ms to  $\sim 7$  for  $\tau \geq 32$  ms ( $r \approx 5$  Å). These values correspond to a  $1 - \gamma^{na}$  range between 0.03 and 0.08 which matches well with the  $(\Delta S/S_0)_f^\alpha$  listed in Table III.

**Table S1.** Values of  $(S_1/S_0)_r^\beta \equiv A$  <sup>a</sup>

Sample ( <sup>13</sup> CO/ <sup>15</sup> N labels) <sup>b</sup>	A
FP-Hairpin (L7/F8)	0.28
FP-Hairpin (F8/L9)	0.24
FP-Hairpin (L12/G13)	0.32
FP-Hairpin (F11/L12)	0.24
FP-Hairpin (I4/G5)	0.36
FP34 (L7/F8)	0.27
N70 (L7/F8)	0.33

<sup>a</sup>  $S_{0r}$  and  $S_{1r}$  intensities were determined using 1.0 integration windows centered at the peak  $\beta$  shift.

<sup>b</sup> Black and red typefaces respectively refer to initial protein solubilization in Buffer and in Buffer + D-malt.

**Table S2.** Sample characteristics

Sample ( <sup>13</sup> CO/ <sup>15</sup> N labels) <sup>a</sup>	(L+C):P <sup>b,c</sup>	$\mu$ mole protein <sup>c</sup>
FP-Hairpin (L7/F8)	45:1	0.60
FP-Hairpin (L7/F8)	38:1	0.84
FP-Hairpin (F8/L9)	30:1	1.10
FP-Hairpin (F8/L9)	38:1	0.99
FP-Hairpin (L12/G13)	30:1	1.00
FP-Hairpin (L12/G13)	48:1	0.53
FP-Hairpin (F11/L12)	42:1	0.72
FP-Hairpin (I4/G5)	38:1	0.46
FP34 (L7/F8)	38:1	0.68
N70 (L7/F8)	45:1	0.53

<sup>a</sup> Black and red typefaces respectively refer to initial protein solubilization in Buffer and in Buffer + D-malt.

<sup>b</sup> L, C, and P respectively refer to lipid, cholesterol, and protein.

<sup>c</sup> Quantities in initial protein solution and vesicle suspension.

### FP-Hairpin $\Delta S_r$ spectra with deconvolution

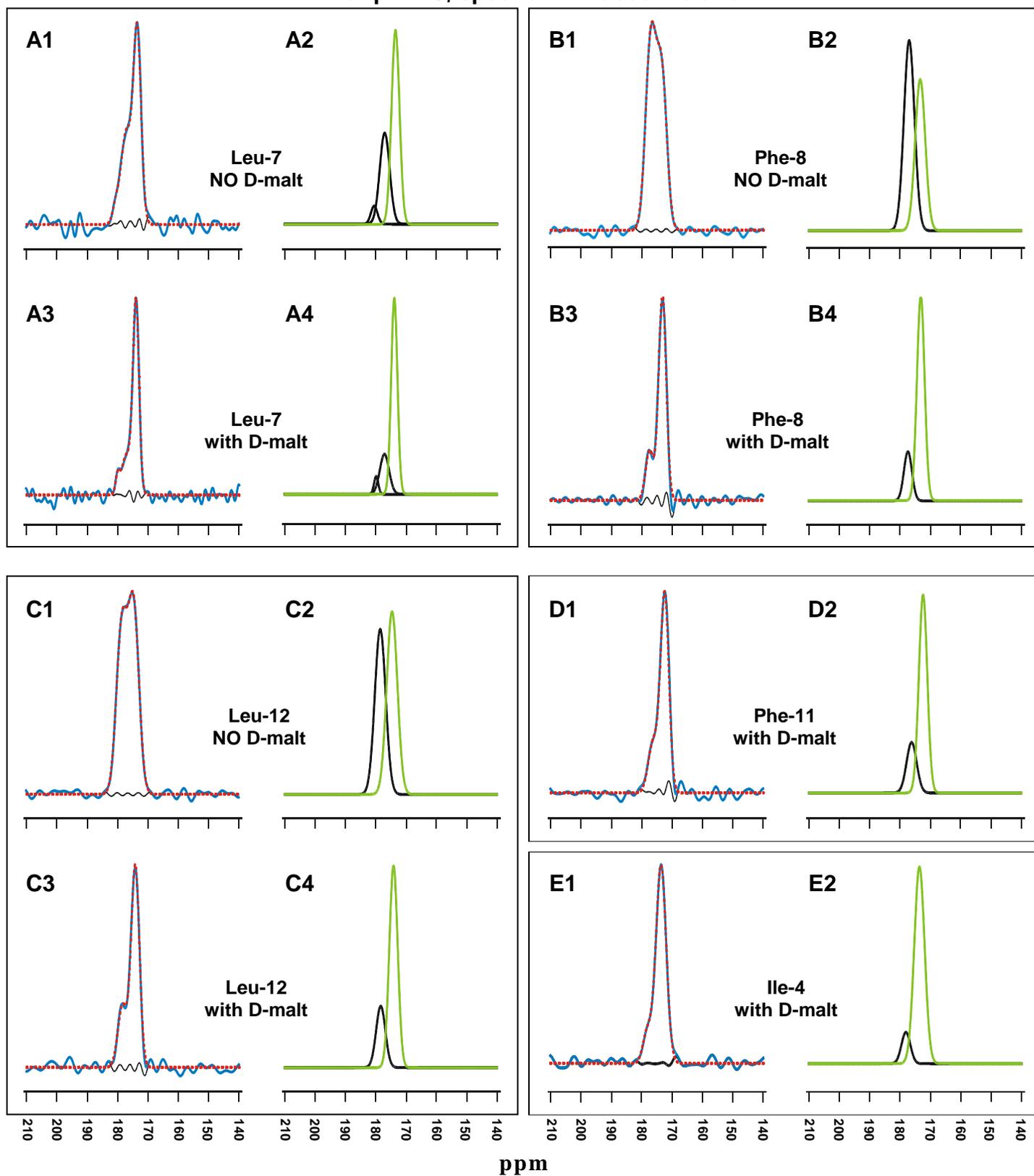
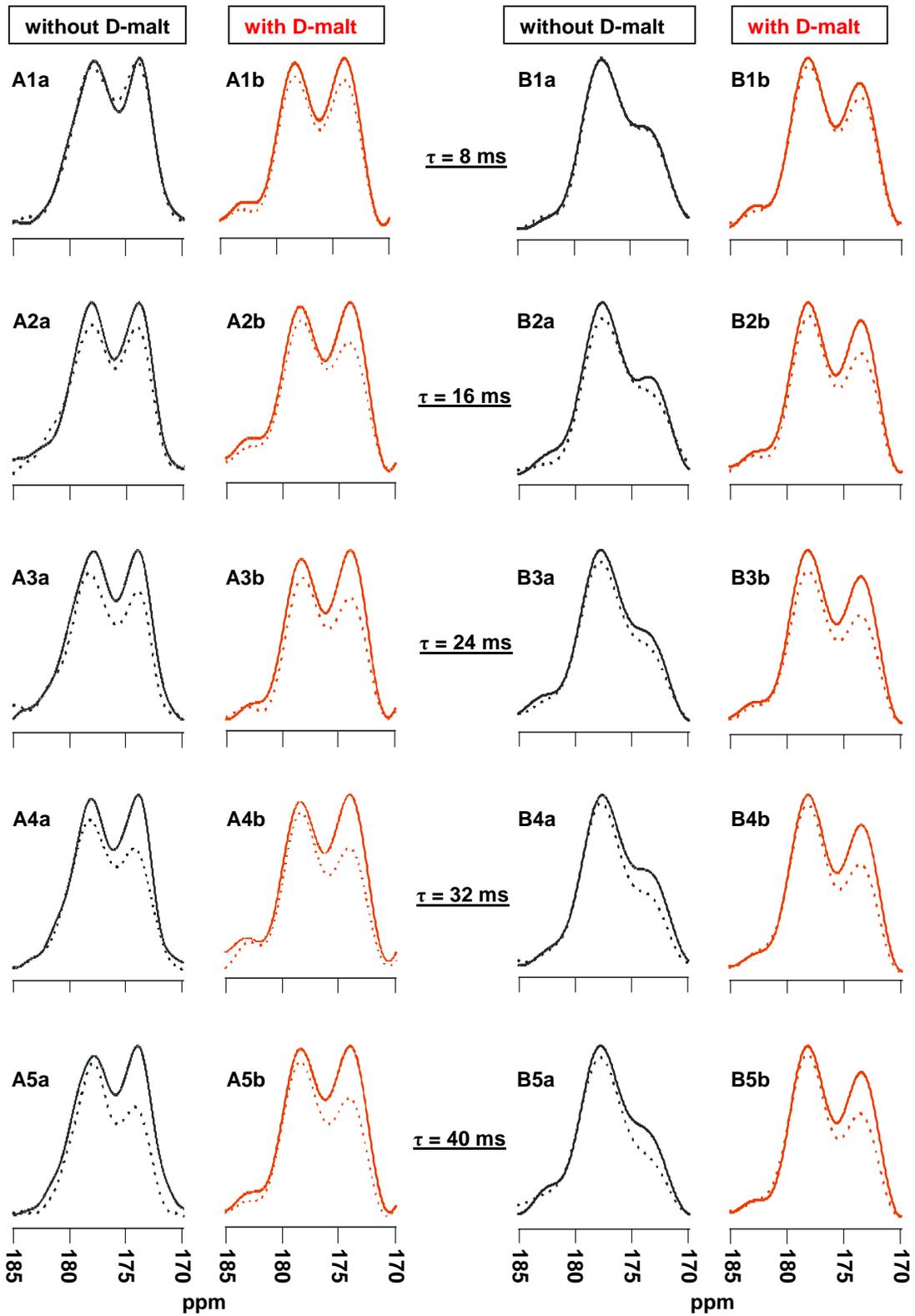
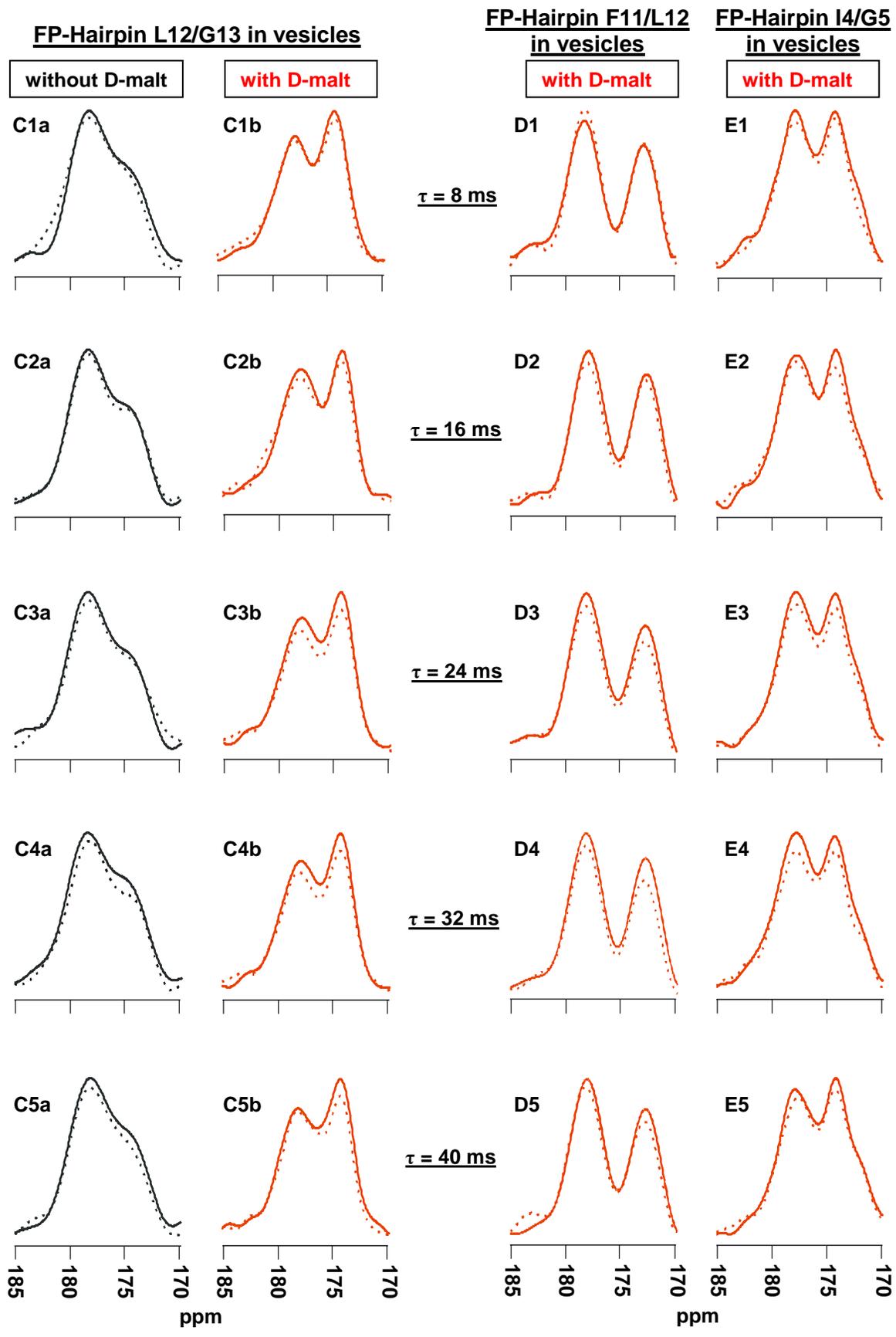


Fig. S1. Deconvolution of REDOR  $\Delta S_r$  spectra of FP-Hairpin samples in membranes (fig. 3) with a single backbone  $^{13}\text{CO}/^{15}\text{N}$ -amide labeled spin pair at (A) L7/F8, (B) F8/L9, (C) L12/G13, (D), F11/L12, and (E) I4/G5. Panels A1, B1, and C1 display spectra of samples with initial FP-Hairpin solubilization in Buffer, and panels A3, B3, C3, D1 and E1 display spectra of samples with initial FP-Hairpin solubilization in Buffer + D-malt. For each of these panels, the experimental  $\Delta S_r$ , deconvolution sum, and difference spectra are respectively displayed as solid blue, broken red, and black lines. Panels A2, A4, B2, B4, C2, C4, D2, and E2 respectively display the individual deconvolution Gaussian lineshapes of the A1, A3, B1, B3, C1, C3, D1, and E1 spectra. The green and black lineshapes are respectively assigned to  $\beta$  sheet and  $\alpha$  helical conformation and Table I lists the peak shift, linewidth, and fractional integrated intensity of each of the lineshapes.<sup>44</sup> Good fitting of the L7/F8 spectra required two  $\alpha$  helical lineshapes.

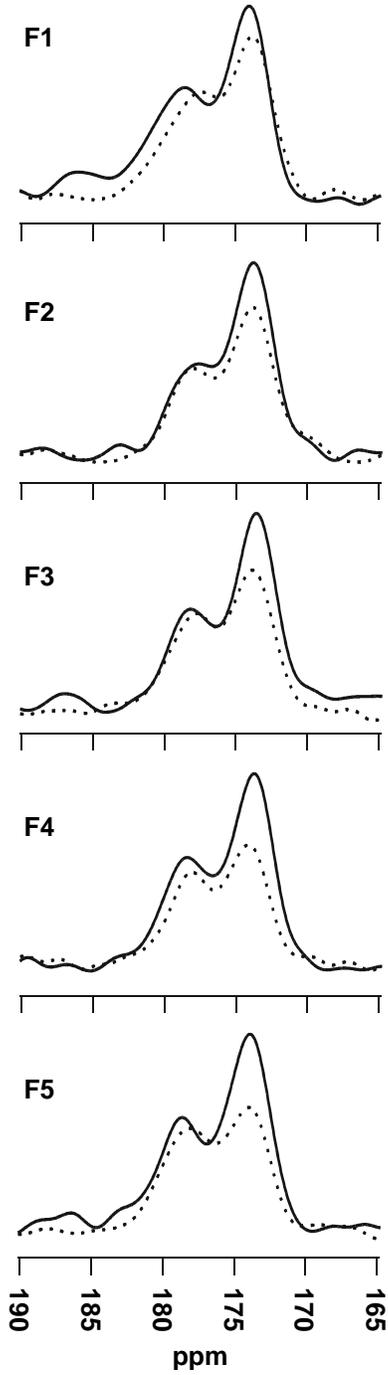
**FP-Hairpin L7/F8 in vesicles**

**FP-Hairpin F8/L9 in vesicles**





**N70 L7/F8 in vesicles**



$\tau = 6.7$  ms

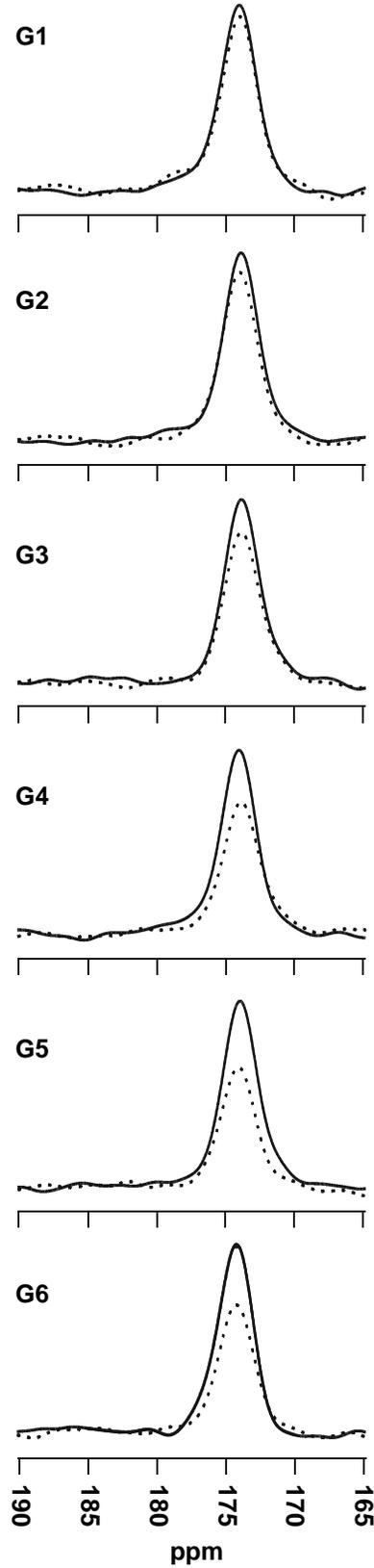
$\tau = 14.7$  ms

$\tau = 22.7$  ms

$\tau = 30.7$  ms

$\tau = 38.7$  ms

**FP34 L7/F8 in vesicles**



$\tau = 46.7$  ms

Fig. S2. fpCTDQBU  $^{13}\text{C}$ O spectra for (A-E) FP-Hairpin, (F) N70, and (G) FP34 samples. Solid and dashed traces are respectively  $S_{0f}$  and  $S_{1f}$  spectra. The proteins had  $^{13}\text{C}/^{15}\text{N}$  labels at (A, F, G) L7/F8, (B) F8/L9, (C) L12/G13, (D) F11/L12, and (E) I4/G5. Black and red spectra respectively correspond to samples with initial protein solubilization in Buffer or Buffer + D-malt. Processing of spectra Aa, Cb, and F + G included Gaussian line broadening of 150, 100, and 200 Hz. Processing of all spectra included polynomial baseline correction.