



Published in final edited form as:

*Biochemistry*. 2010 December 21; 49(50): 10623–10635. doi:10.1021/bi101389r.

## Major Antiparallel and Minor Parallel $\beta$ Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide<sup>†</sup>

Scott D. Schmick and David P. Weliky<sup>\*,†</sup>

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

### Abstract

The HIV gp41 protein catalyzes fusion between viral and host cell membranes and its apolar N-terminal region or “fusion peptide” binds to the host cell membrane and plays a key role in fusion. “HFP” is a construct containing the fusion peptide sequence, induces membrane vesicle fusion, and is an important fusion model system. Earlier solid-state NMR (SSNMR) studies showed that when HFP is associated with membranes with ~30 mol% cholesterol, the first sixteen residues have predominant  $\beta$  strand secondary structure and a fraction of the strands form antiparallel  $\beta$  sheet structure with residue 16 $\rightarrow$ 1/1 $\rightarrow$ 16 or 17 $\rightarrow$ 1/1 $\rightarrow$ 17 registries of adjacent strands. In some contrast, other SSNMR and infrared studies have been interpreted to support a large fraction of approximately in-register parallel registry of adjacent strands. However, the samples had extensive isotopic labeling and other structural models were also consistent with the data. The present SSNMR study uses sparse labeling schemes that reduce ambiguity in the determination of the fraction of HFP molecules with parallel  $\beta$  registry. Quantitative analysis of the data shows that the parallel fraction is at most 0.15 with a much greater fraction of antiparallel 16 $\rightarrow$ 1/1 $\rightarrow$ 16 and 17 $\rightarrow$ 1/1 $\rightarrow$ 17 registries. These data strongly support a model of HFP-induced vesicle fusion caused by antiparallel rather than parallel registries and provide insight into the arrangement of gp41 molecules during HIV/host cell fusion. This study is an example of quantitative determination of a complex structural distribution by SSNMR including experimentally-validated inclusion of natural abundance contributions to the SSNMR data.

### Keywords

HIV; fusion; peptide; parallel; antiparallel; solid-state NMR

AIDS is caused by membrane-enveloped human immunodeficiency virus (HIV) which infects host cells by fusion, i.e. joining of viral and host cell membranes (1). Fusion is facilitated by gp41 which is an integral HIV membrane protein. The N-terminal ~175-residue gp41 ectodomain lies outside the virus and X-ray crystal and liquid-state nuclear magnetic resonance (LSNMR) structures of soluble regions of the ectodomain have shown molecular trimers (1–5). These structures lacked the ~25-residue N-terminal “fusion peptide” region which plays a key role in fusion and infection as evidenced by inhibition of fusion and infection when gp41 had mutations within the fusion peptide region (1, 6, 7). Peptides with the fusion peptide sequence are denoted HFPs and have been studied as model

<sup>†</sup>This work was supported by NIH award AI47153 to D.P.W.

<sup>\*</sup>Corresponding author. Tel: 517-355-9715. Fax: 517-353-1793. weliky@chemistry.msu.edu.

#### SUPPORTING INFORMATION AVAILABLE

Table of  $(\Delta S/S_0)^{exp}$  values and schematics of spin geometries and associated  $\gamma_{1uv}^{lab}(\tau)$  or  $\gamma_{1112uv}^{lab}(\tau)$  calculated by SIMPSON simulation. This information is available free of charge via the Internet at <http://pubs.acs.org>.

fusion systems because they induce vesicle fusion and because their mutation-fusion activity relationships are similar to those of *in vivo* fusion and infection (1, 8–10).

The HFP structure-function literature includes NMR data showing random coil structure for HFP in aqueous solution (11, 12). Solid-state nuclear magnetic resonance (SSNMR) has shown predominant  $\beta$  sheet structure for residues 1–16 of membrane-associated HFP where the membranes contained ~30 mol% cholesterol which is comparable to the mol% cholesterol of membranes of HIV and host cells of HIV (13–16). A fluorescence and infrared (IR) study reported the time-resolved courses of HFP structural changes and the intervesicle lipid mixing function following addition of a HFP solution to a membrane vesicle solution (17). The experimental rates ( $R_s$ ) were ordered  $R_{HFP\ membrane\ binding} > R_{HFP\ \beta\ sheet\ formation} > R_{lipid\ mixing}$  and were consistent with the sequence: (1) random coil HFPs bind to a membrane vesicle; (2) HFP structure changes to oligomeric  $\beta$  sheet; and finally (3) vesicle fusion.

The biological relevance of HFP oligomers is further supported by the molecular trimer structure of soluble regions of the gp41 ectodomain (3–5). The region between residues T25 and G85 of each molecule was a continuous helix and the helices of the different molecules formed a parallel coiled-coil. The fusion peptide region was not included in the protein constructs for these structures but would be N-terminal of residue T25. A C-terminally cross-linked HFP trimer (HFPtr) was therefore synthesized to mimic the close proximity of the three T25 residues in the coiled-coil. Relative to HFP monomer, HFPtr induced membrane vesicle fusion with ~40-fold faster rate which supported the functional significance of the trimer (18). Although both the monomer and trimer formed  $\beta$  sheet oligomers in membranes with cholesterol, HFPtr is more deeply inserted which correlates with greater membrane perturbation and a reduction of the vesicle fusion activation energy (19). The *in vivo* importance of fusion peptide oligomers was also demonstrated by dominant inhibition of fusion and infection in viruses and cells for which a small fraction of the gp41 had the V2E point mutation in the fusion peptide region (7, 20). Analyses of these data supported the involvement of multiple gp41 trimers and fusion peptides in fusion (21). Electron micrographs of virus-cell contacts have also been interpreted to show multiple gp41 trimers at the contact site (22). Functional importance of fusion peptide trimers has also been demonstrated for fusion peptides of other viruses (23, 24).

Because of the aforementioned functional significance of HIV fusion peptide oligomers, there has been effort to elucidate the distribution of structures of membrane-associated HFP oligomers. SSNMR has played a key role in this effort in particular for samples prepared in a manner similar to that of fusion assays with addition of an aqueous fusion peptide solution to a membrane vesicle solution (14). Appendage of a C-terminal lysine tag to HFP greatly reduced HFP aggregation in aqueous solution and allowed separation of pelleted fused vesicles with bound HFP from unbound HFP in the supernatant (12, 18, 25). HFP/lipid binding was supported by SSNMR detection of a HFP A1  $^{13}\text{C}$ (carbonyl)-lipid  $^{31}\text{P}$  distance of ~5 Å (19). For membrane-associated HFP, the  $^{13}\text{C}$  chemical shifts derived from an unambiguous assignment were consistent with a fully extended  $\beta$  strand conformation for residues between A1 and G16 (15). Detection of intermolecular  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{13}\text{C}$ - $^{15}\text{N}$  distances of ~5 Å supported  $\beta$  sheet oligomer/aggregate structure and the A1  $^{13}\text{C}$ -lipid  $^{31}\text{P}$  contact and other data suggest that the number of molecules in the oligomer is small (15, 19, 26).

The present work focuses on quantitative determination of populations of specific  $\beta$  sheet registries. The clearest information to-date on this topic has been a SSNMR experiment on membrane-associated HFP with an A14  $^{13}\text{C}$  label and a G3  $^{15}\text{N}$  label whose separation ( $r_{CN}$ ) was >20 Å along a single  $\beta$  strand (15). SSNMR can detect labeled  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar

coupling ( $d_{CN}$ ) where  $d_{CN} = 3109/r_{CN}^3$  with  $d$  in Hz and  $r$  in Å. The minimum detectable  $d_{CN} \approx 10$  Hz correlates with  $r_{CN} \approx 7$  Å so that detectable  $d_{CN}$  in this sample were necessarily ascribed to intermolecular rather than intramolecular  $^{13}\text{C}$ - $^{15}\text{N}$  proximity. SSNMR detection of  $d > 30$  Hz strongly supported a significant fraction of molecules with intermolecular A14-G3 hydrogen bonding and labeled  $r_{CN}$  of 4.1 and 5.5 Å, i.e. 16→1/1→16 antiparallel β sheet registry. Fig. 1c displays this registry with isotopic labeling from the present study and not the earlier study. Detection of similarly large  $d$  in an A14  $^{13}\text{C}$ O/I4  $^{15}\text{N}$  HFP sample supported a fraction of 17→1/1→17 antiparallel registry.

At most half of the membrane-associated HFP molecules were in the 16→1/1→16 or 17→1/1→17 registries, i.e. a large fraction of the molecules were in registries not detected in either the A14  $^{13}\text{C}$ O/G3  $^{15}\text{N}$  or A14  $^{13}\text{C}$ O/I4  $^{15}\text{N}$  labeled samples. Because of the close proximity of the T25 residues of the three molecules of the gp41 trimer, a reasonable hypothesis for a populated HFP registry is in-register parallel β sheet, i.e. 1→17/1→17 in Fig. 1c. An earlier SSNMR study attempted to test this hypothesis using samples each containing an equimolar mixture of two labeled HFPs, one with three sequential backbone  $^{13}\text{C}$ O labels and the other with three sequential backbone  $^{15}\text{N}$  labels (27). Detection of an average  $d_{CN} > 10$  Hz for a G5-L7  $^{13}\text{C}$ O/G5-L7  $^{15}\text{N}$  sample and a F11-G13  $^{13}\text{C}$ O/F11-G13  $^{15}\text{N}$  sample were consistent with a fraction of in-register parallel HFP molecules. However, because the samples were extensively labeled, the data were also consistent with other parallel or antiparallel registries. In addition, the data reflected an average of many intermolecular  $d_{CN}$ s so it was not possible to determine the fraction of molecules with a particular registry. There have also been efforts to detect in-register parallel structure using SSNMR measurement of intermolecular  $^{13}\text{C}$ - $^{13}\text{C}$  dipolar couplings ( $d_{CC}$ s) where  $d_{CC} = 7710/r_{CC}^3$  with  $d_{CC}$  in Hz and  $r_{CC}$  in Å. For HFP with a single  $^{13}\text{C}$ O label and in-register parallel structure, the labeled interstrand  $r_{CC} \approx 5$  Å with  $d_{CC} \approx 70$  Hz (28, 29). These parameters will be independent of the residue that is  $^{13}\text{C}$ O labeled. For membrane-associated HFP with F8  $^{13}\text{C}$ O, a best-fit  $d_{CC} \approx 70$  Hz was detected whereas for membrane-associated HFPtr,  $d_{CC}$  depended on the position of the labeled  $^{13}\text{C}$ O residue with a range of 10–60 Hz (30, 31). This residue dependence argued against a major fraction of in-register parallel structure in HFPtr.

There was also an IR spectroscopy effort to distinguish between the 1→17/1→17 parallel and 16→1/1→16 antiparallel registries using samples that contained backbone  $^{13}\text{C}$ O labeling at either: (1) A1 to V3, G5 to I9; (2) F8 to G16; or (3) A1 to V3, G5 to G16 (32). The IR wavenumbers and intensities of different samples were interpreted to support a large fraction of parallel structure and little antiparallel structure. However, in our view, the extensive labeling of the IR samples precluded quantitation of specific registries and greater support for this argument is provided in the Discussion section.

The present paper reports a determination of the fraction of parallel structure in membrane-associated HFP oligomers. We were motivated to study this question because of: (1) the functional significance of HIV fusion peptide oligomers; and (2) the existing indefinite and conflicting data and interpretations relevant to this question. As part of this effort, we developed a model to quantify effects of natural abundance  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei on SSNMR measurements of  $d_{CN}$  and experimentally validated this model.

## MATERIALS AND METHODS

### Materials

9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids and Fmoc-Ala-Wang resin were purchased from Peptides International (Louisville, KY). Isotopically labeled amino acids were purchased from Cambridge Isotopes (Andover, MA) and were Fmoc-protected

using literature methods (33). Fig. 1a shows the labeled HFPs. The 23 N-terminal residues (AVGIGALFLGFLGAAGSTMGARS) are a consensus sequence of the gp41 fusion peptide, non-native W24 is an A<sub>280</sub> chromophore and the non-native lysines greatly reduced HFP aggregation in aqueous solution prior to membrane binding (12, 18). This ensured that membrane-associated  $\beta$  sheet oligomers/aggregates were formed after membrane binding.

HFP was manually synthesized and then cleaved from the resin for three hours in a solution of trifluoroacetic acid (TFA):water:anisole:thioanisole:ethanedithiol in a 90:5:2:2:2 volume ratio. After precipitation with cold diethyl ether, centrifugation, and dissolution of the pellet in water, crude HFP was purified by reversed-phase high-performance liquid chromatography with a semi-preparative C<sub>4</sub> column and a water-acetonitrile gradient containing 0.1% TFA. Acetonitrile and TFA were removed with nitrogen gas and water was then removed by lyophilization. HFP purity was >95% as determined by mass spectrometry. HFP amounts were quantified using A<sub>280</sub> of aqueous solutions of HFP with  $\epsilon = 5600 \text{ M}^{-1}\text{cm}^{-1}$ .

### SSNMR samples

As shown in Fig. 1b, each sample contained a <sup>13</sup>CO and a <sup>15</sup>N labeled peptide in a 1:2 mol ratio. Three samples contained membrane-associated HFP  $\beta$  sheet oligomers/aggregates that were each a statistical mixture of <sup>13</sup>CO and <sup>15</sup>N labeled HFPs. Detection of substantial  $d_{CN}$  by SSNMR indicated proximity of the labeled <sup>13</sup>CO and <sup>15</sup>N nuclei on adjacent strands and was used to estimate the fractional populations of specific registries as detailed below. As shown in Fig. 1c, the HFP-P sample was designed to detect parallel 1 $\rightarrow$ 17/1 $\rightarrow$ 17 and 2 $\rightarrow$ 17/1 $\rightarrow$ 16 registries, the HFP-A sample was designed to detect previously observed antiparallel 16 $\rightarrow$ 1/1 $\rightarrow$ 16 and 17 $\rightarrow$ 1/1 $\rightarrow$ 17 registries, and the HFP-AP sample was designed to detect both parallel 1 $\rightarrow$ 17/1 $\rightarrow$ 17 and 2 $\rightarrow$ 17/1 $\rightarrow$ 16 registries as well as antiparallel 16 $\rightarrow$ 1/1 $\rightarrow$ 16 and 17 $\rightarrow$ 1/1 $\rightarrow$ 17 registries.

In addition to the potential proximity of labeled <sup>13</sup>CO and <sup>15</sup>N nuclei, there will always be proximity between labeled <sup>13</sup>CO and some natural abundance (n.a.) <sup>15</sup>N nuclei as well as proximity between some n.a. <sup>13</sup>CO and labeled <sup>15</sup>N nuclei. These proximities will contribute to the  $d_{CN}$  detected in the SSNMR experiment and should be included in the data modeling. Quantitative understanding of these proximities required a negative control (HFP-NC) sample with: (1) the same relative fractions of labeled <sup>13</sup>CO, <sup>15</sup>N, and n.a. sites as the HFP-P, HFP-A, and HFP-AP samples; and (2) labeled <sup>13</sup>CO – labeled <sup>15</sup>N  $r_{CN}$  that are much greater than the REDOR detection limit of  $\sim 7 \text{ \AA}$ . One possibility was a sample made like HFP-P, HFP-A, and HFP-AP but with labels at sites that do not form intermolecular hydrogen bonds. This possibility was not pursued because the distribution of registries of membrane-associated HFP is not yet well-defined. Instead, the HFP-NC sample was a physical mixture of lyophilized HFP-F8 (5.0 mg) and HFP-A6L7 (10.0 mg) without any membrane. Each peptide was lyophilized separately and the two peptides were then mixed in the solid phase to form a uniform physical mixture. Water and membrane were not added to the physical mixture so that the labeled <sup>13</sup>COs and <sup>15</sup>Ns remained much farther apart than the  $7 \text{ \AA}$  REDOR detection limit. Although there were populations of  $\beta$  sheet as well as  $\alpha$  helical lyophilized peptides in the HFP-NC sample, each population yielded very similar ( $\Delta S/S_0$ ) – see Results section for further details.

For the HFP-P, HFP-A, and HFP-AP samples, the membrane composition was 1,2-di-*O*-tetradecyl-*sn*-glycero-3-phosphocholine (DTPC) lipid, 1,2-di-*O*-tetradecyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (DTPG) lipid, and cholesterol in a 8:2:5 mol ratio. This composition reflected the large amount of choline lipid and fractions of negatively charged lipid and cholesterol in membranes of host cells of HIV (16). Ether- rather than more physiologically abundant ester-linked lipids were used because the latter have two COs/

molecule that would contribute substantial n.a.  $^{13}\text{C}$ O signal. Bilayer phase is retained for ether-linked lipids with cholesterol and with HFP (34–36). In addition, membrane-associated HFP has predominant  $\beta$  sheet structure in either ester-linked lipid + cholesterol or ether-linked lipid + cholesterol compositions (30).

Samples were prepared by first dissolving DTPC (40  $\mu\text{mol}$ ), DTPG (10  $\mu\text{mol}$ ), and cholesterol (25  $\mu\text{mol}$ ) in chloroform and removing the chloroform with nitrogen gas and vacuum. The lipid film was suspended in 2 mL of 5 mM HEPES buffer at pH 7.0 with 0.01%  $\text{NaN}_3$  preservative. The suspension was homogenized with ten freeze-thaw cycles and large unilamellar vesicles were formed by extrusion through a 100 nm diameter polycarbonate filter (Avestin, Ottawa, ON). A separate solution was prepared with  $^{13}\text{C}$ O-labeled HFP (3.0 mg) and  $^{15}\text{N}$  labeled HFP (6.0 mg) in HEPES buffer (32 mL). The HFP solution was added drop-wise to the vesicle solution and the combined solution was gently stirred overnight. Ultracentrifugation at  $\sim 150000g$  for four hours pelleted membranes with bound HFP while unbound HFP remained in the supernatant (14). The pellet was lyophilized, transferred to the SSNMR rotor, and rehydrated with 30  $\mu\text{L}$   $\text{H}_2\text{O}$  (37). The validity of the lyophilization/rehydration approach was evidenced by peak  $^{13}\text{C}$ O chemical shifts that were within 0.6 ppm of those of samples that were not lyophilized (15).

### SSNMR experiments

Data were collected on a 9.4 T spectrometer (Varian Infinity Plus, Palo Alto, CA) using a triple resonance MAS probe equipped for 4.0 mm rotors and tuned to  $^{13}\text{C}$ ,  $^1\text{H}$ , and  $^{15}\text{N}$  nuclei at respective frequencies of 100.8, 400.8, and 40.6 MHz. The  $^{13}\text{C}$  chemical shift was externally referenced to the methylene resonance of adamantane at 40.5 ppm and the  $^{13}\text{C}$  transmitter was set to 153 ppm. The  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar coupling ( $d_{\text{CN}}$ ) was probed with the rotational-echo double-resonance (REDOR) experiment with typical parameters: (1) 52 kHz  $^1\text{H}$   $\pi/2$  pulse; (2) 2.2 ms cross-polarization with 74 kHz  $^1\text{H}$  field and 83–98 kHz ramped  $^{13}\text{C}$  field; (3) dephasing period of duration  $\tau$  for which the “ $S_0$ ” and “ $S_1$ ” acquisitions had 60 kHz  $^{13}\text{C}$   $\pi$  pulses at the end of each rotor cycle except the last cycle, and the  $S_1$  acquisitions additionally had 59 kHz  $^{15}\text{N}$   $\pi$  pulses in the middle of each rotor cycle; and (4)  $^{13}\text{C}$  detection (30, 38). The MAS frequency was 10 kHz, the recycle delay was 2 s, 88 kHz TPPM  $^1\text{H}$  decoupling was applied during the dephasing and detection periods, and XY-8 phase cycling was applied to the  $^{13}\text{C}$   $\pi$  pulses and to the  $^{15}\text{N}$   $\pi$  pulses (39, 40). Experiments were calibrated using a lyophilized helical peptide with sequence AEAAAKEAAAKEAAKA, N-terminal acetylation and C-terminal amidation, and A9  $^{13}\text{C}$ O and A13  $^{15}\text{N}$  labels (30, 41). The labeled  $r_{\text{CN}} \approx 4.1 \text{ \AA}$  corresponds to  $d_{\text{CN}} \approx 45 \text{ Hz}$ . Samples were typically cooled by nitrogen gas at  $-50 \text{ }^\circ\text{C}$  to enhance  $^{13}\text{C}$ O signal and reduce motional averaging of  $d_{\text{CN}}$  (42). The typical difference between  $^{13}\text{C}$  shift in cooled and uncooled membrane-associated HFP samples is  $\sim 0.5$  ppm and indicates little variation in secondary structure with temperature (26). For each sample, data were collected for  $\tau = 2.2, 8.2, 16.2, 24.2, 32.2, 40.2, \text{ and } 48.2 \text{ ms}$ .

The  $^{13}\text{C}$ O- $^{15}\text{N}$  dipolar coupling,  $d_{\text{CN}}$ , is detected by reduction of  $^{13}\text{C}$ O signal intensity in the  $S_1$  spectrum relative to the  $S_0$  spectrum with greater reduction at increased  $\tau$ . These intensities were also denoted  $S_1$  and  $S_0$  and were determined from integration over a shift range that encompassed most of the  $^{13}\text{C}$ O signal. A range of 8 ppm was used for HFP-NC spectra and a range of 5 ppm was used for HFP-P, HFP-A, and HFP-AP spectra. The normalized dephasing:

$$\left(\frac{\Delta S}{S_0}\right)^{\text{exp}} = \frac{S_0^{\text{exp}} - S_1^{\text{exp}}}{S_0^{\text{exp}}} = 1 - \frac{S_1^{\text{exp}}}{S_0^{\text{exp}}} \quad (1)$$



and its standard deviation:

$$\sigma^{exp} = \sqrt{\frac{(S_0^2 \times \sigma_{s_0}^2) + (S_1^2 \times \sigma_{s_1}^2)}{S_0^2}} \quad (2)$$

where  $\sigma S_0$  and  $\sigma S_1$  were the experimental root-mean-squared deviations of the spectral intensities derived from 12 regions of the spectrum that did not include spectral features (43).

Experimental dephasing of a membrane-associated HFP sample was modeled as a sum of  $S_0$  and a sum of  $S_1$  signals of different spin geometries of  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei where the geometries reflected statistical distributions of n.a.  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei as well as geometries of (1) 1→17/1→17 and 2→17/1→16 parallel adjacent strand registries; (2) 16→1/1→16 and 17→1/1→17 antiparallel registries; and (3) other “X” registries where all labeled  $r_{CN} > 7 \text{ \AA}$  and  $S_1 = S_0$ . The notation  $(\Delta S/S_0)^{sim}$  will be generally used for simulated  $(\Delta S/S_0)$  and can refer to a particular spin geometry or to the population weighted sum using calculations from different spin geometries. For the former case, the  $(S_1/S_0)^{sim} \equiv \gamma$  were calculated using the SIMPSON program with input parameters that included  $d_{CNS}$  as well as Euler angles in a fixed crystal frame for each  $^{13}\text{C}$ - $^{15}\text{N}$  vector and for the  $^{13}\text{C}$  chemical shift anisotropy (CSA) principal axis system (44, 45). These input parameters were calculated by the SIMMOL program using  $^{13}\text{C}$  and  $^{15}\text{N}$  coordinates from a region of a high-resolution crystal structure with the appropriate structural motif, e.g. parallel  $\beta$  sheet (46). Coordinates were obtained from the following Protein Data Bank (PDB) files: 1JK3, 1IGD, 1NKI, 2E4T, 1CEX, 1MNZ, and 2IWW. For each spin geometry,  $(S_1/S_0)^{sim}$  was the average of ten different SIMPSON calculations and each calculation was based on input parameters from a different set of atomic coordinates. The  $^{13}\text{C}$  CSA principal values of 247, 176, and 99 ppm were inputs to the SIMPSON calculations and  $^1\text{H}$ s and relaxation were not considered.

## RESULTS

### Chemical shift and conformational distributions

Fig. 2 displays REDOR  $S_0$  and  $S_1$   $^{13}\text{C}$  SSNMR spectra for  $\tau = 32.2$  ms. Each  $S_0$  spectrum has a ~50% contribution from the labeled  $^{13}\text{C}$  and ~50% contribution from n.a.  $^{13}\text{C}$ Os of the unlabeled residues. The full-width at half-maximum linewidths of the membrane-associated HFP samples in Fig. 2b–d are 3–4 ppm and indicate a distinct secondary structure. For the HFP-AP sample with F8  $^{13}\text{C}$  label, the peak  $^{13}\text{C}$  shift of 173 ppm is the same as was observed for F8  $^{13}\text{C}$  of HFP in known  $\beta$  strand conformation and is very different from the 178 ppm shift observed in  $\alpha$  helical conformation (15, 30). For the HFP-P and HFP-A samples with L12  $^{13}\text{C}$  label, the 174 ppm peak shift is also the same as  $\beta$  strand HFP and different from the 179 ppm shift of Leu in helical HFP (30, 47). Overall, the shifts and linewidths are consistent with the fully extended conformation that has been observed for the first sixteen residues of HFP associated with membranes with biologically relevant cholesterol content (15).

The linewidth of the lyophilized HFP-NC sample with F8  $^{13}\text{C}$  label is ~7 ppm and correlates with a broad distribution of secondary structures that is also evidenced by a 176 ppm peak  $^{13}\text{C}$  shift that is midway between typical Phe helical and  $\beta$  strand shifts (48).

### Qualitative analysis of REDOR data

Relative to the  $S_0$  signals, there is attenuation in the  $S_1$  signals of  $^{13}\text{C}$ Os within ~7  $\text{\AA}$  of  $^{15}\text{N}$ s and the associated  $\Delta S/S_0$  normalized dephasing increased with dephasing time, Fig. 2 and

Fig. 3. Because of the physical separation of the  $^{13}\text{CO}$  and  $^{15}\text{N}$  labeled HFPs in HFP-NC, the  $S_1$  attenuation and  $(\Delta S/S_0)^{exp}$  of this sample reflected F8  $^{13}\text{CO}$ -n.a.  $^{15}\text{N}$  and n.a.  $^{13}\text{CO}$ -A6,L7  $^{15}\text{N}$  proximities but not F8  $^{13}\text{CO}$ -A6,L7  $^{15}\text{N}$  proximity, Figs. 2a and 3a. There was similar  $S_1$  attenuation and  $(\Delta S/S_0)^{exp}$  in HFP-P, Figs. 2b and 3b, which demonstrated that there was little L12  $^{13}\text{CO}$ -G13,A14  $^{15}\text{N}$  proximity in HFP-P and only a small fraction of parallel  $1 \rightarrow 17/1 \rightarrow 17$  and  $2 \rightarrow 17/1 \rightarrow 16$  registries. There was much larger  $S_1$  attenuation and  $(\Delta S/S_0)^{exp}$  for the HFP-A sample, Figs. 2c and 3b, which indicated significant L12  $^{13}\text{CO}$ -G5,A6  $^{15}\text{N}$  proximity and therefore a substantial fraction of antiparallel  $16 \rightarrow 1/1 \rightarrow 16$  and  $17 \rightarrow 17/1 \rightarrow 17$  registries. Comparably large  $S_1$  attenuation and  $(\Delta S/S_0)^{exp}$  were observed for the HFP-AP sample, Figs. 2d and 3b. The similarity of the HFP-NC and HFP-P data and the similarity of the HFP-A and HFP-AP data were consistent with ascribing F8  $^{13}\text{CO}$ -L9,G10  $^{15}\text{N}$  proximity in HFP-AP to antiparallel  $16 \rightarrow 1/1 \rightarrow 16$  and  $17 \rightarrow 1/1 \rightarrow 17$  registries rather than parallel  $1 \rightarrow 17/1 \rightarrow 17$  and  $2 \rightarrow 17/1 \rightarrow 16$  registries. Detection of a substantial fraction of these antiparallel registries is consistent with earlier SSNMR data for sparsely labeled HFP (15). Detection of only a small fraction of parallel registries is a new result and disagrees with previous interpretations of SSNMR and IR data for samples with extensive labeling (27, 32). These new data highlight the importance of sparse labeling to reduce interpretational ambiguity for systems with a structural distribution like membrane-associated HFP.

### Natural abundance models

Quantitative analysis of the  $(\Delta S/S_0)^{exp}$  to yield the fraction of parallel and antiparallel HFP registries requires an accurate natural abundance dephasing (n.a.d.) model, i.e. a model that accounts for effects of labeled  $^{13}\text{CO}$ -n.a.  $^{15}\text{N}$  and n.a.  $^{13}\text{CO}$ -labeled  $^{15}\text{N}$  proximities. Both types of proximities were considered but for conciseness of presentation, the discussion in this paper focuses on labeled  $^{13}\text{CO}$ -n.a.  $^{15}\text{N}$ . One measure of validity of a n.a.d. model was agreement within experimental error between  $(\Delta S/S_0)^{exp}$  of HFP-NC and  $(\Delta S/S_0)^{sim}$  of the model. Consideration was first given to the HFP-F8 regions of HFP-NC including the spin geometries of one or two labeled  $^{13}\text{CO}$ s and one n.a.  $^{15}\text{N}$ . Geometries with two or more n.a.  $^{15}\text{N}$ s were not considered because the fractional isotopic abundance of  $^{15}\text{N}$  is only 0.0037. For each geometry, the SIMPSON program was used to calculate  $(S_1/S_0)^{sim}$  as a function of the dephasing time  $\tau$ . Only geometries with  $r_{CN} < 7 \text{ \AA}$  were considered because those with  $r_{CN} > 7 \text{ \AA}$  do not affect  $(S_1/S_0)^{sim}$  within our experimental signal-to-noise. We consider this a “long-range” n.a.d. model which is distinguished from a “short-range” model of earlier studies that only considered n.a. nuclei separated by one or two bonds from a labeled nucleus, i.e.  $r_{CN} < 3 \text{ \AA}$  (15, 30). The broad spectral linewidth of HFP-NC indicated both helical and  $\beta$  strand conformational populations and coordinates of spin geometries for both  $\alpha$  helical and  $\beta$  sheet structures were obtained from corresponding regions of high-resolution structure PDB files. For  $\alpha$  helical structure, the  $r_{CN} < 7 \text{ \AA}$  criterion resulted in geometries with a single labeled  $^{13}\text{CO}$  at residue  $i$  and a single n.a.  $^{15}\text{N}$  at a residue between  $i-3$  and  $i+5$ . These nine geometries are one aspect of the  $\alpha$  n.a.d. model.

Fig. 4 illustrates relevant labeled  $^{13}\text{CO}$ s and n.a.  $^{15}\text{N}$ s for antiparallel  $\beta$  sheet structure. The strands in panels a and c are “fully constrained” to a single registry with resultant six unique spin geometries. Three geometries had one labeled  $^{13}\text{CO}$  and one n.a.  $^{15}\text{N}$  within the same strand and three geometries had two labeled  $^{13}\text{CO}$ s on vicinal strands and one n.a.  $^{15}\text{N}$  in the intervening strand. In panels b and d, the strands have different registries so that the labeled  $^{13}\text{CO}$  in the top strand was  $>7 \text{ \AA}$  from the nine n.a.  $^{15}\text{N}$  sites of the  $^{13}\text{CO}$  of the third strand. The structure of panels b and d has nine unique spin geometries and is denoted a maximum  $\beta$  sheet n.a.d. (max  $\beta$  n.a.d.) model while the structure of panels a and c has six geometries and is denoted a minimum  $\beta$  sheet n.a.d. (min  $\beta$  n.a.d.) model. In either structure there are nine n.a.  $^{15}\text{N}$  sites within  $7 \text{ \AA}$  of each labeled  $^{13}\text{CO}$  but in the min n.a.d. model,

some sites (e.g. 4–6 in Fig. 4c) are “shared” between two  $^{13}\text{CO}$ s, i.e. within 7 Å of two  $^{13}\text{CO}$ s. This reduces the average number of n.a.  $^{15}\text{N}$  sites per  $^{13}\text{CO}$  and the overall n.a.d.

There are many indices and parameters in the quantitative modeling and descriptions and possible values for them are compiled in Table 1. For the  $\alpha$ , min  $\beta$ , or max  $\beta$  n.a.d. models, the average  $\gamma = (S_1/S_0)$  for the relevant spin geometries:

$$\gamma_{IN}^{na}(\tau) = (J)^{-1} \sum_{j=1}^J \left[ \frac{S_{lj}(\tau)}{S_{0j}(\tau)} \right]^{sim} \quad (3)$$

where “na” refers to natural abundance,  $l=0$  or 1, “N” refers to n.a.  $^{15}\text{N}$ ,  $j$  is the index of a particular spin geometry, and  $J$  is the number of unique spin geometries of the model ( $J=9$  for  $\alpha$  or max  $\beta$  n.a.d. and  $J=6$  for min  $\beta$  n.a.d.). The  $\gamma_{0N}^{na}(\tau) = 1$  for all  $\tau$  while  $\gamma_{1N}^{na}(\tau)$  were calculated using  $[S_{lj}(\tau)/S_{0j}(\tau)]^{sim}$  from the SIMPSON program and generally decreased with increasing  $\tau$ . After setting a total labeled  $^{13}\text{CO}$  population of 1.0 for HFP-F8 in HFP-NC, the relative population affected by n.a.  $^{15}\text{N}$  is  $J \times 0.0037$  while the remainder population,  $[1 - (J \times 0.0037)]$ , has  $S_1 = S_0$ . There is also an n.a.  $^{13}\text{CO}$  contribution from unlabeled residues in HFP-F8 with a relative population  $30 \times 0.011$  and with  $S_1 = S_0$ . Similar analysis for  $^{15}\text{N}$  labeled HFP-A6L7 in HFP-NC results in:

$$\gamma_{IC}^{na}(\tau) = (K)^{-1} \sum_{k=1}^K \left[ \frac{S_{lk}(\tau)}{S_{0k}(\tau)} \right]^{sim} \quad (4)$$

where  $l=0$  or 1, “C” refers to n.a.  $^{13}\text{CO}$ ,  $k$  is the index of a particular spin geometry,  $K$  is the number of unique n.a.  $^{13}\text{CO}$ -labeled  $^{15}\text{N}$  spin geometries of the model ( $K=10, 8,$  or  $12$ , respectively, for  $\alpha$ , min  $\beta$ , or max  $\beta$  n.a.d. models), and  $\gamma_{0C}^{na}(\tau) = 1$ . Accounting for the 1:2 ratio of HFP-F8:HFP-A6L7, the total n.a.  $^{13}\text{CO}$  population of HFP-A6L7 is  $2 \times 31 \times 0.011$  with population  $2K \times 0.011$  affected by labeled  $^{15}\text{N}$  and the remainder having  $S_1 = S_0$ . For HFP-NC in total:

$$S_{tot}^{sim}(\tau) = \{ [J \times 0.0037 \times \gamma_{IN}^{na}(\tau)] + [2K \times 0.011 \times \gamma_{IC}^{na}(\tau)] \} + \{ 2.0 - [J \times 0.0037] - [2K \times 0.011] \} \quad (5)$$

where  $l=0$  or 1 and the terms in the first braces are  $\tau$ - and  $l$ -dependent.

For each n.a.d. model ( $\alpha$ , min  $\beta$ , and max  $\beta$ ),  $(\Delta S/S_0)^{sim}$  for each  $\tau$  was calculated with Eq. 5 and statistical comparison was then made to  $(\Delta S/S_0)^{exp}$ :

$$\chi^2 = \sum_{m=1}^7 \left\{ \frac{\left( \frac{\Delta S}{S_0} \right)_m^{sim} - \left( \frac{\Delta S}{S_0} \right)_m^{exp}}{\sigma_m^{exp}} \right\}^2 \quad (6)$$

where  $m$  is the index for an experimental datum, i.e. a particular  $\tau$ . The respective  $\chi^2$  for the  $\alpha$ , min  $\beta$ , and max  $\beta$  n.a.d. models were 1.2, 3.8, and 2.0 which were all less than the number of degrees of fitting, 7, i.e. the number of data, 7, minus the number of independent fitting parameters, 0. The validity of the approach to n.a.d. calculation was supported by good fits for all models (43).

The broad  $^{13}\text{CO}$  linewidth of HFP-NC in Fig. 2a was consistent with two HFP populations, one with helical and one with  $\beta$  strand secondary structure. It was therefore reasonable to calculate  $(\Delta S/S_0)^{sim}$  for “mixtures” with contributions from multiple models:



$$S_{l, mix}^{sim}(\tau) = \sum_{t=1}^3 [f_t \times S_{lt}^{sim}(\tau)] \quad (7)$$

where  $l = 0$  or  $1$ ,  $t$  was the index that referred to the  $\alpha$ , min  $\beta$ , or max  $\beta$  model,  $f_t$  was fractional population with  $\sum f_t = 1$ , and each  $S_{lt}^{sim}(\tau)$  was calculated using Eq. 5. The HFP-NC distribution of  $^{13}\text{CO}$  shifts indicated  $f_\alpha \approx 0.5$  and  $f_{min\beta} + f_{max\beta} \approx 0.5$  but did not provide information about individual  $f_{min\beta}$  or  $f_{max\beta}$ . Fitting using  $f_\alpha = 0.5$  and  $f_{min\beta} = 0.5$ ,  $f_{max\beta} = 0.0$  yielded  $\chi^2 = 1.5$  while fitting using either  $f_\alpha = 0.5$ ,  $f_{min\beta} = 0.0$ ,  $f_{max\beta} = 0.5$  or  $f_\alpha = 0.5$ ,  $f_{min\beta} = 0.25$ ,  $f_{max\beta} = 0.25$  yielded  $\chi^2 = 1.2$  and  $(\Delta S/S_0)^{sim}$  in Fig. 3a were calculated with the latter distribution. The  $(\Delta S/S_0)^{sim}$  from all three conformational distributions fit well to the  $(\Delta S/S_0)^{exp}$  and these models are statistically similar. Together with previously described good fitting for different secondary structure models show that n.a.d. is accurately calculated with these models and only weakly dependent on secondary and tertiary structure. The key feature of all these well-fitting long-range models was consideration of the multiple n.a. sites within  $7 \text{ \AA}$  of a labeled site which led to continually increasing  $(\Delta S/S_0)$  with  $\tau$ , Fig. 3a. The  $(\Delta S/S_0)^{sim}$  were also calculated using a short-range model that only considered n.a. sites separated by one or two bonds from each labeled site. The  $(\Delta S/S_0)^{sim}$  were systematically less than the  $(\Delta S/S_0)^{exp}$  with resultant poor fit and  $\chi^2 = 29$ .

### Quantitative analysis of registry populations – fully constrained model

For membrane-associated HFP, there is a single distribution of registries which we model as fractions of: (1)  $1 \rightarrow 17/1 \rightarrow 17$  and  $2 \rightarrow 17/1 \rightarrow 16$  parallel registries; (2)  $16 \rightarrow 1/1 \rightarrow 16$  and  $17 \rightarrow 1/1 \rightarrow 17$  antiparallel registries; and (3)  $X$  registries not detected by any of our labeling schemes, Fig. 1c. Fraction 1 contributed to the  $(\Delta S/S_0)^{exp}$  of HFP-P, fraction 2 contributed to the  $(\Delta S/S_0)^{exp}$  of HFP-A, and fractions 1 and 2 contributed to  $(\Delta S/S_0)^{exp}$  of HFP-AP. The overall goal was best-fit determination of these fractions based on the  $(\Delta S/S_0)^{exp}$  of the three samples, Fig. 3b, and this analysis required calculation of the n.a.d. contribution to  $(\Delta S/S_0)^{exp}$ . Because a 1:2  $^{13}\text{CO}$ -HFP: $^{15}\text{N}$ -HFP ratio was used for all samples, this contribution was calculated using models developed for HFP-NC and resulted in a modified Eq. 5 appropriate for HFP-P, HFP-A, and HFP-AP:

$$S_{lu, tot}^{sim}(\tau) = \left\{ \left[ J \times 0.0037 \times \gamma_{ln}^{na}(\tau) \right] + \left[ 2K \times 0.011 \times \gamma_{lc}^{na}(\tau) \right] + \{1.0 - (2K \times 0.011)\} + \left\{ [1.0 - (J \times 0.0037)] \times \gamma_{luv}^{lab}(\tau) \right\} \right\} \quad (8)$$

$$= S_l^{na}(\tau) + S_{lu}^{lab}(\tau) \quad (9)$$

where  $S_l^{na}$  is the sum of the first two braced terms in Eq. 8 and  $S_{lu}^{lab}$  is the third braced term. Each membrane-associated sample is labeled by the index  $u$  where  $u = 1, 2$ , or  $3$  respectively refers to HFP-P, HFP-A, or HFP-AP, Table 1 and Fig. 1b. The first braced term in Eq. 8 corresponds to labeled and n.a.  $^{13}\text{CO}$  that experience n.a.d., the second braced term corresponds to n.a.  $^{13}\text{CO}$ s that do not experience n.a.d., and the third braced term corresponds to labeled  $^{13}\text{CO}$ s that do not experience n.a.d. but may experience dephasing from labeled  $^{15}\text{N}$ s. The secondary structure of membrane-associated HFP was predominantly  $\beta$  sheet, Fig. 2b–d, and the best estimates of the n.a.d. terms in the first braced term were taken to be the average of the max  $\beta$  and min  $\beta$  calculated values. In the second and third braced terms,  $K$  and  $J$  were estimated to be their respective average values of 10 and 7.5.  $S_{0u}^{lab}(\tau)$  was calculated using  $\gamma_{0tuv}^{lab}(\tau) = 1$  while  $\gamma_{1tuv}^{lab}(\tau)$  and therefore  $S_{1u}^{lab}(\tau)$  were first calculated with a “fully constrained” model, Fig. 4a,c, in which a  $\beta$  sheet region contained either: (1)  $1 \rightarrow 17/1 \rightarrow 17$  or  $2 \rightarrow 17/1 \rightarrow 16$  parallel registries; or (2)  $16 \rightarrow 1/16 \rightarrow 1$  or  $17 \rightarrow 1/1 \rightarrow 17$  antiparallel registries; or (3)  $X$  registries not directly detected by any of our labeling schemes, Fig. 1c. A sample was considered to be a mixture of the

three registry types each denoted by index  $t = 1, 2, \text{ or } 3$  and fractional population  $f_t$ , Table 1. The  $S_{1u}^{lab}(\tau)$  was calculated by modified Eq. 7:

$$S_{1u}^{lab}(\tau) = \sum_{t=1}^3 [f_t \times \gamma_{1tuv}^{lab}(\tau)] \times S_0^{lab}(\tau) \quad (10)$$

The  $\gamma_{1tuv}^{lab}(\tau)$  values depended on the labeled  $d_{CNs}$  and therefore  $r_{CNs}$  which in turn depended on registry type  $t$  and sample labeling  $u$ , Fig. 1c. For some combinations of  $t$  and  $u$ , all labeled  $r_{CN} > 7 \text{ \AA}$  with consequent  $d_{CN} \approx 0$  and  $\gamma_{1tuv}^{lab}(\tau) = 1$ . Specific examples are  $t = 1$  and  $u = 2$ ,  $t = 2$  and  $u = 1$ , and  $t = 3$  and  $u = 1, 2, \text{ or } 3$ . For other combinations of  $t$  and  $u$ ,  $\gamma_{1tuv}^{lab}(\tau)$  were determined from SIMPSON calculations and Fig. 5a–d displays schematic examples for  $t = 2, u = 2$  with numerical values of  $\gamma_{1tuv}^{lab}(\tau)$  and more details in the Supporting Information. Column a, b, c, or d corresponds to particular arrangements of  $^{13}\text{CO}$  and  $^{15}\text{N}$  labeled HFPs that are respectively denoted by the index  $v = 1, 2, 3, \text{ or } 4$ . For each  $v$ , the typical difference between the calculated  $\gamma_{122v}^{lab}(\tau)$  for the  $16 \rightarrow 1/1 \rightarrow 16$  or  $17 \rightarrow 1/1 \rightarrow 17$  registry was  $\approx 0.01$  and the final  $\gamma_{122v}^{lab}(\tau)$  were the average for the two registries. The antiparallel  $\gamma_{123v}^{lab}(\tau)$  of HFP-AP were analogously calculated and the parallel  $\gamma_{111v}^{lab}(\tau)$  of HFP-P and parallel  $\gamma_{113v}^{lab}(\tau)$  of HFP-AP were calculated using the  $1 \rightarrow 17/1 \rightarrow 17$  registry and had similar values to  $\gamma$ s calculated using the  $2 \rightarrow 17/1 \rightarrow 16$  registry. Fractional weightings  $w_v$  were based on the 1:2 ratio of  $^{13}\text{CO}$  HFP: $^{15}\text{N}$  HFP with  $w_1 = 1/9, w_2 = 2/9, w_3 = 2/9, \text{ and } w_4 = 4/9$ . A more complete version of Eq. 10:

$$S_{1u}^{lab}(\tau) = \sum_{t=1}^3 \left\{ f_t \times \sum_{v=1}^4 [w_v \times \gamma_{1tuv}^{lab}(\tau)] \right\} \times S_0^{lab}(\tau) \quad (11)$$

with indices and parameters summarized in Table 1.

The values of  $f_1, f_2, \text{ and } f_3$  were the same for the HFP-P, HFP-A, and HFP-AP samples with  $f_3 = 1 - f_1 - f_2$ . Best-fit values of  $f_1$  and  $f_2$  were obtained by calculating  $\chi^2(f_1, f_2)$  using an expression analogous to Eq. 6:

$$\chi^2(f_1, f_2) = \sum_{u=1}^3 \sum_{m=1}^7 \left\{ \frac{\left[ \frac{\Delta S(f_1, f_2)}{S_0} \right]_{um}^{sim} - \left[ \frac{\Delta S}{S_0} \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (12)$$

and then selecting the  $f_1$  and  $f_2$  values which corresponded to minimum  $\chi^2$ , i.e.  $\chi^2_{min}$ . In Eq. 12,  $m$  is the index for each  $\tau$ , and  $[\Delta S(f_1, f_2)/S_0]_{um}^{sim}$  was determined using Eqs. 8–11. For this fully constrained model, Fig. 6a displays a plot of  $\chi^2$  vs  $f_1$  and  $f_2$  with best-fit  $f_1 = 0.12$  and  $f_2 = 0.52$  and  $\chi^2_{min} = 11$ . The model was reasonable as evidenced by  $\chi^2_{min}$  which was smaller than the number of degrees of fitting, 19, i.e. the number of data, 21, minus the number of fitting parameters, 2. The  $f_1$  fractional parallel population, Fig. 1c, was small which was consistent with qualitative analysis of the data, Fig. 3b. The  $f_2$  antiparallel population was substantially larger and also consistent with Fig. 3b. The  $f_3 \approx 0.35$  indicated a substantial population of  $X$  registries not detected by the labeling of the three samples.

The above fitting was done using a long-range n.a.d. model that considered effects of n.a. sites within  $7 \text{ \AA}$  of each labeled nucleus. Fitting displayed in Fig. 6a was based on n.a.d. calculated from half min  $\beta$  and half max  $\beta$  sheet structure, Fig. 4, but the best-fit  $f_1$  and  $f_2$  and  $\chi^2_{min}$  were not sensitive to the structural composition of the long-range n.a.d. model. For example, fitting done using n.a.d. for half  $\alpha$  helical and half max  $\beta$  sheet structure

yielded best-fit  $f_1$  and  $f_2$  and  $\chi^2_{min}$  respectively within 0.01, 0.01, and 1 of the corresponding Fig. 6a values. For HFP-NC fitting, n.a.d. was underestimated by a short-range model that only considered n.a. sites separated by one or two bonds from each labeled site. This effect was also observed when fitting membrane-associated HFP data with the short-range n.a.d. model and led to best-fit  $f_1 = 0.22$  and  $f_2 = 0.57$  which were significantly higher than the Fig. 6a values. The  $\chi^2_{min} = 20$  using the short-range model was also higher than the  $\chi^2_{min}$  in Fig. 6a.

### Quantitative analysis of registry populations – unconstrained model

In addition to the fully constrained model for strand registries, an alternate “unconstrained” fitting model was also considered for which there was local mixing of: (1) 1→17/1→17 parallel registries; (2) 16→1/1→16 and 17→1/1→17 antiparallel registries; and (3)  $X$  registries not directly detected by any of our labeling schemes, Fig. 1c. Each pairwise registry type was labeled by  $t=1, 2,$  or  $3$  with fractional population  $f_t$ . For this unconstrained model, Fig. 5e–h displays schematics of three-strand registries with  $^{13}\text{C}$ O labeled HFP in the middle strand. Each e–h row has three-strand registries that were each a combination of two registries labeled by specific  $t_1$  and  $t_2$  which denote the respective  $t$  of the top/middle and middle/bottom strands. As with the fully constrained models, the registries in each e–h column corresponded to a particular  $^{13}\text{C}$ O HFP/ $^{15}\text{N}$  HFP arrangement which respective label  $v = 1, 2, 3,$  or  $4$ . The  $^{13}\text{C}$ O HFP: $^{15}\text{N}$  HFP population ratio of 1:2 correlated with a sum weighting of 1/9 for the  $v = 1$  registries with individual registry weighting  $w_1 = 1/(9 \times 9) = 1/81$ . The sum weightings for  $v = 2, 3,$  or  $4$  were respectively 2/9, 2/9, or 4/9 with respective individual weightings  $w_2 = 2/81, w_3 = 2/81,$  and  $w_4 = 4/81$ . Eq. 11 was modified for the unconstrained model:

$$S_{1u}^{lab}(\tau) = \sum_{t_1=1}^3 \sum_{t_2=1}^3 \left\{ f_{t_1} \times f_{t_2} \times \sum_{v=1}^4 \left[ w_v \times \gamma_{1t_1t_2uv}^{lab}(\tau) \right] \right\} \times S_0^{lab}(\tau) \quad (13)$$

Similar to the fully constrained model, many combinations of  $t_1, t_2, u,$  and  $v$  have  $r_{CN} > 7 \text{ \AA}$  with consequent  $d_{CN} \approx 0$  and  $\gamma_{1t_1t_2uv}^{lab}(\tau) = 1$ . In Fig. 5e–g, such registries are not enclosed by a box. Similar to results for the fully constrained model, the  $\gamma_{1t_1t_2uv}^{lab}(\tau)$  were similar for the two antiparallel registries and an average value was used.

The values of  $f_1, f_2,$  and  $f_3$  in the unconstrained model were the same for the HFP-P, HFP-A, and HFP-AP samples with  $f_3 = 1 - f_1 - f_2$ . Best-fit values of  $f_1$  and  $f_2$  were obtained with Eq. 12 and Fig. 6b displays a plot of  $\chi^2$  vs  $f_1$  and  $f_2$  with best-fit  $f_1 = 0.11$  and  $f_2 = 0.46$  and corresponding  $\chi^2_{min} = 8$ . The unconstrained model was reasonable as evidenced by a best-fit  $\chi^2$  which was smaller than the number of degrees of fitting, 19. Similar to the results of the fully constrained model, the  $f_1$  fractional parallel population was small and the  $f_2$  antiparallel population and  $f_3 \approx 0.4$  other population were significant.

This unconstrained model fitting was done with n.a.d. calculated with a long-range model and half min  $\beta$  and half max  $\beta$  structure. Similar to the fully constrained model, best-fit  $f_1, f_2,$  and  $\chi^2$  for the unconstrained model were: (1) negligibly affected by the structural distribution of the long-range n.a.d. model; and (2) significantly increased by use of a short-range n.a.d. model.

## DISCUSSION

This paper sets an upper limit of  $\sim 0.15$  on the fraction of membrane-associated HFP in in-register parallel  $\beta$  sheet structure and this result is supported by both qualitative analysis of the data, Fig. 3b, as well as quantitative analyses with fully constrained and unconstrained

models, Fig. 6. Both models fit the data well and yielded similar best-fit fractional population of parallel registries and similar populations of antiparallel registries. The small fractional parallel population agrees with some earlier SSNMR studies but differs from interpretations of other SSNMR and IR data which respectively reported ~0.5 and ~1.0 fractions of in-register parallel structure (27, 30–32). The present study used samples with sparse isotopic labeling while the earlier studies interpreted to support a large fraction of parallel structure used samples with extensive labeling. We think that there was ambiguity of interpretation in the studies of extensively labeled samples and that the data could also be reasonably interpreted in terms of small in-register parallel population. For example, the earlier SSNMR study also used the REDOR technique but with only a single  $\tau$  (24 ms) and with samples containing equimolar amounts of HFP  $^{13}\text{C}$  labeled on three sequential residues and HFP  $^{15}\text{N}$  labeled on three sequential residues. The typical  $(\Delta S/S_0)^{exp}$  was ~0.1 and was approximately independent of the positions of the labeled residues and also had some contribution from n.a.d. It was not possible to do unambiguous quantitative analysis of registry distributions because: (1) each sample was extensively labeled so that non-zero  $(\Delta S/S_0)$  was expected for many different registries; (2)  $(\Delta S/S_0)^{exp}$  were only measured for a single  $\tau$ ; and (3) a “HFP-NC”-type sample was not studied and n.a.d. was therefore not quantitatively modeled.

The samples for the IR study were also extensively labeled with backbone  $^{13}\text{C}$  labeling at either: (1) A1 to V3, G5 to L9; (2) F8 to G16; or (3) A1 to V3, G5 to G16. The authors' interpretation of their spectra to support predominant in-register parallel structure was based in part on expected effects of ( $^{13}\text{C}=\text{O}$  electric dipole)···( $^{13}\text{C}=\text{O}$  electric dipole) coupling on  $^{13}\text{C}=\text{O}$  vibrational wavenumber and intensity. However, their interpretation appeared to neglect the substantial intramolecular coupling between  $^{13}\text{C}=\text{O}$ s on adjacent residues and we note that this coupling is independent of registry. In addition to these “undiluted” samples, three “diluted” samples were studied that had an equimolar mixture of a labeled and unlabeled peptide. The wavenumber ( $\nu$ ) of a  $^{13}\text{C}=\text{O}$  vibration is sensitive to nearby (~5 Å away) C=O vibrations and is higher with  $^{12}\text{C}=\text{O}$  neighbors than with  $^{13}\text{C}=\text{O}$  neighbors. If there is hydrogen bonding between  $^{13}\text{C}$  labeled residues of adjacent strands in an undiluted sample, the corresponding diluted sample should have an increased fraction of  $^{13}\text{C}=\text{O}/^{12}\text{C}=\text{O}$  proximities, decreased fraction of  $^{13}\text{C}=\text{O}/^{13}\text{C}=\text{O}$  proximities, and  $\Delta \nu = \nu_{diluted} - \nu_{undiluted} > 0$ . If there were a major fraction of parallel 1→17/1→17 structure (as claimed by the authors), dilution of (1) A1-V3, G5-L9; (2) F8-G16; or (3) A1-V3, G5-G16 labeled HFps would have had comparable effect on proximities and resulted in similar  $\Delta \nu$ . However, the experimental  $\Delta \nu_{A1-V3, G5-L9} \approx \Delta \nu_{F8-G16} \approx (\Delta \nu_{A1-V3, G5-G16})/2$  which is inconsistent with a large fraction of in-register parallel structure. Like the earlier SSNMR study on extensively labeled samples, extensive labeling of the IR samples also meant that the IR data were consistent with many registry distributions and precluded more quantitative analysis of the distribution. Overall, the sparse labeling of the present SSNMR study allowed for much more unambiguous and quantitative determination of the populations of specific registries. This general approach can be applied in the future to determine the registry distributions of HFP constructs with very high or low fusogenicity such as HFPtr or V2E mutant, respectively.

Fig. 7 displays a structure/function model for HFP based on results from this and earlier studies. Prior to membrane binding, HFP is monomeric in aqueous solution and has random coil structure (11, 12, 18). HFP sequentially: (1) binds to membranes; (2) forms  $\beta$  sheet oligomers with a significant fraction of 16→1/1→16 and 17→1/1→17 antiparallel registries; and (3) induces membrane fusion as monitored by intervesicle lipid mixing (12, 14, 17, 18). It is also known that the A6 and L9 residues of  $\beta$  sheet HFP insert shallowly into the membrane with correlation between membrane insertion depth and both membrane perturbation and fusion rate (19, 36, 49). A global structure-function model is non-

transmembrane HFP insertion perturbs bilayer structure and moves the membrane on the fusion reaction coordinate towards the highly perturbed transition state with consequent reduction in fusion activation energy and increase in fusion rate.

There is functional and electron microscopic evidence that multiple gp41 trimers are required for fusion and Fig. 7(right) shows a  $\beta$  sheet HFP hexamer as would be reasonable for interleaved antiparallel fusion peptides from two gp41 trimers. While there are no data specifically supporting a HFP hexamer, HFP  $\beta$  sheet oligomers likely contain a small number of molecules because: (1) for ~90% of HFP molecules, there is an A1  $^{13}\text{C}$ -lipid  $^{31}\text{P}$  distance of ~5 Å, i.e. close contact of most HFPs with the membrane; and (2) significant temperature dependence of intensities of SSNMR spectra (19, 26). Interleaved antiparallel fusion peptides from multiple gp41 trimers may also be the fusogenic structure of HIV/host cell fusion. As noted above, this structure can insert into and perturb membranes which are likely requirements for HIV/host cell fusion.

The experimentally observed membrane insertion of the central  $\beta$  sheet region (e.g. A6 to L12) of HFP is consistent with  $\Delta G^{\text{insertion}} \approx -6$  kJ/mol for the fully constrained 17→1/1→17 registry as calculated by summing individual residue insertion energies for L12→A6/A6→L12, Fig. 1c (50). A similar calculation yielded  $\Delta G^{\text{insertion}} \approx -3$  kJ/mol for L12→G5/G5→L12 of the 16→1/1→16 registry. The  $\Delta G^{\text{insertion}} \approx -6$  kJ/mol for A6→L12/A6→L12 of the 1→17/1→17 parallel registry suggests that  $\Delta G^{\text{insertion}}$  does not underlie the preference for antiparallel over parallel structure. This preference may instead be due to  $\Delta G^{\text{electrostatic}}$  as the HFP N-terminus is located in the high water content lipid headgroup region and is therefore likely protonated. Closest intermolecular  $\text{NH}_3^+-\text{NH}_3^+$  distance is ~5 Å for the low population 1→17/1→17 registry and ~10 Å for the significantly populated 17→1/1→17 registry. For  $\epsilon^{\text{dielectric}} = 78$  and hexameric HFP,  $\Delta G^{\text{electrostatic}} \approx +5.1$  kJ/mol for the 1→17/1→17 registry and +1.5 kJ/mol for the 17→1/1→17 registry.

We expect that inclusion of the non-native C-terminal W(K)<sub>6</sub>A tag does not contribute to the preference for antiparallel over parallel registry because either registry would have similar minimized electrostatic repulsion energy. Such repulsion would be minimized by: (1) extensive solvation of the tag; and (2) large inter-tag distances that are possible because of random coil tag structure. Tag solvation is supported by the previously observed lack of membrane insertion of HFP beyond residue L12 and random coil tag structure is supported by broad NMR linewidths in the C-terminal region of HFP (14, 19). We also note that inclusion of the tag has minor effect on fusion activity and that similar REDOR  $\Delta S/S_0$  were observed for mixtures of triply  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled HFPs with or without the tag (12, 27).

In contrast to the reasonably large distribution of membrane-associated HFP registries, i.e. significant 16→1/1→16, 17→1/1→17, and *X* registries, SSNMR studies of  $\beta$  sheet registries of protein in amyloid fibrils have typically shown a single registry which is usually in-register parallel, e.g. 1→17/1→17 (28, 29). The width of a fibril is at most a few protein molecules and the length is >200 molecules and along the intermolecular  $\beta$  sheet hydrogen bonding direction. The amyloid fibrils are grown in aqueous solution (without lipid) and their greater registry homogeneity may reflect ordered fibril growth from seeds (51).

One distinctive feature of the present study is the development of a quantitative n.a.d. model that was experimentally validated. Accurate fitting of the HFP-NC data and fitting of the membrane-associated HFP data relied on a long-range n.a.d. model which included effects of n.a. nuclei <7 Å from each labeled nucleus. For this model, the n.a.d. was approximately independent of secondary and tertiary structure. The n.a.d. was systematically



underestimated by a short-range model which only considered n.a. nuclei separated by one or two bonds from each labeled nucleus.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

NMR and mass spectra were obtained at Michigan State University facilities. We acknowledge useful discussions with Dr. Wei Qiang, Dr. Kelly Sackett, and Dr. Charles Gabrys.

## ABBREVIATIONS

<b>CO</b>	carbonyl
<b><i>d</i></b>	magnetic dipole-dipole coupling
<b>DTPC</b>	1,2-di- <i>O</i> -tetradecyl- <i>sn</i> -glycero-3-phosphocholine
<b>DTPG</b>	1,2-di- <i>O</i> -tetradecyl- <i>sn</i> -glycero-3-phosp- <i>rac</i> -(1-glycerol) sodium salt
<b>FMOG</b>	9-fluorenylmethoxycarbonyl
<b>HEPES</b>	<i>N</i> -(2-hydroxyethyl)piperazine- <i>N</i> -2-ethanesulfonic acid
<b>HFP</b>	HIV fusion peptide
<b>HFPtr</b>	HIV fusion peptide trimer
<b>HIV</b>	human immunodeficiency virus
<b>IR</b>	infrared
<b>MAS</b>	magic angle spinning
<b>max n.a.d.</b>	maximum natural abundance dephasing
<b>min n.a.d.</b>	minimum natural abundance dephasing
<b>n.a.</b>	natural abundance
<b>n.a.d.</b>	natural abundance dephasing
<b>NMR</b>	nuclear magnetic resonance
<b>PDB</b>	Protein Data Bank
<b><i>r</i></b>	internuclear distance
<b>REDOR</b>	rotational-echo double resonance
<b>SIMPSON</b>	simulation program for solid-state NMR spectroscopy
<b>SSNMR</b>	solid-state NMR
<b>TFA</b>	trifluoroacetic acid
<b>TPPM</b>	two-pulse phase-modulation

## 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. [PubMed: 18568847]

2. Tan K, Liu J, Wang J, Shen S, Lu M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. U.S.A.* 1997; 94:12303–12308. [PubMed: 9356444]
3. Caffrey M, Cai M, Kaufman J, Stahl SJ, Wingfield PT, Covell DG, Gronenborn AM, Clore GM. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J.* 1998; 17:4572–4584. [PubMed: 9707417]
4. 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; 126:131–144. [PubMed: 10388624]
5. Buzon V, Natrajan G, Schibli D, Campelo F, Kozlov MM, Weissenhorn W. Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *Plos Pathogens.* 2010; 6:e1000880. [PubMed: 20463810]
6. Freed EO, Myers DJ, Risser R. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc. Natl. Acad. Sci. U.S.A.* 1990; 87:4650–4654. [PubMed: 2191297]
7. Freed EO, Delwart EL, Buchschacher GL Jr, Panganiban AT. 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.* 1992; 89:70–74. [PubMed: 1729720]
8. 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. [PubMed: 9394290]
9. Nieva JL, Agirre A. Are fusion peptides a good model to study viral cell fusion? *Biochim. Biophys. Acta.* 2003; 1614:104–115. [PubMed: 12873771]
10. Epand RM. Fusion peptides and the mechanism of viral fusion. *Biochim. Biophys. Acta.* 2003; 1614:116–121. [PubMed: 12873772]
11. Chang DK, Chien WJ, Cheng SF. The FLG motif in the N-terminal region of glucoprotein 41 of human immunodeficiency virus type 1 adopts a type-I beta turn in aqueous solution and serves as the initiation site for helix formation. *Eur. J. Biochem.* 1997; 247:896–905. [PubMed: 9288913]
12. 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; 87:1951–1963. [PubMed: 15345571]
13. 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; 73:1977–1986. [PubMed: 9336193]
14. 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; 40:8126–8137. [PubMed: 11434782]
15. 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; 130:5459–5471. [PubMed: 18370385]
16. 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. U.S.A.* 2006; 103:2641–2646. [PubMed: 16481622]
17. Buzon V, Padros E, Cladera J. Interaction of fusion peptides from HIV gp41 with membranes: A time-resolved membrane binding, lipid mixing, and structural study. *Biochemistry.* 2005; 44:13354–13364. [PubMed: 16201760]
18. 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; 126:14722–14723. [PubMed: 15535688]
19. Qiang W, Sun Y, Weliky DP. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:15314–15319. [PubMed: 19706388]

20. Yang XZ, Kurteva S, Ren XP, Lee S, Sodroski J. Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1. *J. Virol.* 2005; 79:12132–12147. [PubMed: 16160141]
21. Magnus C, Rusert P, Bonhoeffer S, Trkola A, Regoes RR. Estimating the stoichiometry of Human Immunodeficiency Virus entry. *J. Virol.* 2009; 83:1523–1531. [PubMed: 19019953]
22. Sougrat R, Bartesaghi A, Lifson JD, Bennett AE, Bess JW, Zabransky DJ, Subramaniam S. Electron tomography of the contact between T cells and SIV/HIV-1: Implications for viral entry. *PLOS Pathogens.* 2007; 3:571–581.
23. Lau WL, Ege DS, Lear JD, Hammer DA, DeGrado WF. Oligomerization of fusogenic peptides promotes membrane fusion by enhancing membrane destabilization. *Biophys. J.* 2004; 86:272–284. [PubMed: 14695269]
24. Pan JH, Lai CB, Scott WRP, Straus SK. Synthetic fusion peptides of tick-borne Encephalitis virus as models for membrane fusion. *Biochemistry.* 2010; 49:287–296. [PubMed: 20000438]
25. Han X, Tamm LK. A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc. Natl. Acad. Sci. U.S.A.* 2000; 97:13097–13102. [PubMed: 11069282]
26. 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; 42:187–194. [PubMed: 14745799]
27. Yang J, Weliky DP. Solid state nuclear magnetic resonance evidence for parallel and antiparallel strand arrangements in the membrane-associated HIV-1 fusion peptide. *Biochemistry.* 2003; 42:11879–11890. [PubMed: 14529300]
28. Benzinger TL, Gregory DM, Burkoth TS, Miller-Auer H, Lynn DG, Botto RE, Meredith SC. Propagating structure of Alzheimer's  $\beta$ -amyloid(10–35) is parallel  $\beta$ -sheet with residues in exact register. *Proc. Natl. Acad. Sci. U.S.A.* 1998; 95:13407–13412. [PubMed: 9811813]
29. Tycko R. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Quart. Rev. Biophys.* 2006; 39:1–55.
30. 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; 45:12960–12975. [PubMed: 17059213]
31. 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; 45:S247–S260.
32. 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. [PubMed: 15964015]
33. Lapatsanis L, Miliadis G, Froussios K, Kolovos M. Synthesis of N-2,2,2-(trichloroethoxycarbonyl)-L-amino acids and N-(9-fluorenylmethoxycarbonyl)-L-amino acids involving succinimidoxycarbonyl anion as a leaving group in amino-acid protection. *Synthesis.* 1983; 8:671–673.
34. Siminovitch DJ, Ruocco MJ, Makriyannis A, Griffin RG. The effect of cholesterol on lipid dynamics and packing in diether phosphatidylcholine bilayers. X-ray diffraction and  $^2\text{H}$ -NMR study. *Biochim. Biophys. Acta.* 1987; 901:191–200. [PubMed: 3607045]
35. Komatsu H, Rowe ES. Effect of cholesterol on the ethanol-induced interdigitated gel phase in phosphatidylcholine: Use of fluorophore pyrene-labeled phosphatidylcholine. *Biochemistry.* 1991; 30:2463–2470. [PubMed: 2001373]
36. Gabrys CM, Yang R, Wasniewski CM, Yang J, Canlas CG, Qiang W, Sun Y, Weliky DP. Nuclear magnetic resonance evidence for retention of a lamellar membrane phase with curvature in the presence of large quantities of the HIV fusion peptide. *Biochim. Biophys. Acta.* 2010; 1798:194–201. [PubMed: 19616505]
37. Huster D, Xiao LS, Hong M. Solid-state NMR investigation of the dynamics of the soluble and membrane-bound colicin Ia channel-forming domain. *Biochemistry.* 2001; 40:7662–7674. [PubMed: 11412120]
38. Gullion T, Schaefer J. Rotational-echo double-resonance NMR. *J. Magn. Reson.* 1989; 81:196–200.

39. Gullion T, Baker DB, Conradi MS. New, compensated Carr-Purcell sequences. *J. Magn. Reson.* 1990; 89:479–484.
40. Bennett AE, Rienstra CM, Auger M, Lakshmi KV, Griffin RG. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 1995; 103:6951–6958.
41. Long HW, Tycko R. Biopolymer conformational distributions from solid-state NMR: alpha-helix and 3(10)-helix contents of a helical peptide. *J. Am. Chem. Soc.* 1998; 120:7039–7048.
42. Fowler DJ, Weis RM, Thompson LK. Kinase-active signaling complexes of bacterial chemoreceptors do not contain proposed receptor-receptor contacts observed in crystal structures. *Biochemistry.* 2010; 49:1425–1434. [PubMed: 20088541]
43. Bevington, PR.; Robinson, DK. *Data Reduction and Error Analysis for the Physical Sciences.* 2nd ed.. Boston: McGraw-Hill; 1992.
44. Oas TG, Hartzell CJ, McMahon TJ, Drobny GP, Dahlquist FW. The carbonyl  $^{13}\text{C}$  chemical-shift tensors of 5 peptides determined from  $^{15}\text{N}$  dipole-coupled chemical shift powder patterns. *J. Am. Chem. Soc.* 1987; 109:5956–5962.
45. Bak M, Rasmussen JT, Nielsen NC. SIMPSON: A general simulation program for solid-state NMR spectroscopy. *J. Magn. Reson.* 2000; 147:296–330. [PubMed: 11097821]
46. Bak M, Schultz R, Vosegaard T, Nielsen NC. Specification and visualization of anisotropic interaction tensors in polypeptides and numerical simulations in biological solid-state NMR. *J. Magn. Reson.* 2002; 154:28–45. [PubMed: 11820824]
47. Jaroniec CP, Kaufman JD, Stahl SJ, Viard M, Blumenthal R, Wingfield PT, Bax A. Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry.* 2005; 44:16167–16180. [PubMed: 16331977]
48. Zhang HY, Neal S, Wishart DS. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR.* 2003; 25:173–195. [PubMed: 12652131]
49. Tristram-Nagle S, Chan R, Kooijman E, Uppamoochikkal P, Qiang W, Weliky DP, Nagle JF. HIV fusion peptide penetrates, disorders, and softens T-cell membrane mimics. *J. Mol. Biol.* 2010; 402:139–153. [PubMed: 20655315]
50. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, Heijne G. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature.* 2005; 433:377–381. [PubMed: 15674282]
51. Petkova AT, Leapman RD, Guo ZH, Yau WM, Mattson MP, Tycko R. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science.* 2005; 307:262–265. [PubMed: 15653506]

## (a) Peptides

	1	10	20	30
HFP-F8	AVGIGAL	FLGFLGAAGS	TMGARS	WKKKKKKA
HFP-L12	AVGIGAL	FLGFLGAAGS	TMGARS	WKKKKKKA
HFP-G5A6	AVGI	GALFLGFLGAAGS	TMGARS	WKKKKKKA
HFP-A6L7	AVGIGAL	FLGFLGAAGS	TMGARS	WKKKKKKA
HFP-L12G13	AVGIGAL	FLGFLGAAGS	TMGARS	WKKKKKKA
HFP-G13A14	AVGIGAL	FLGFLGAAGS	TMGARS	WKKKKKKA

## (b) Samples

HFP-NC	HFP-F8 + HFP-A6L7	(1:2)
HFP-P	HFP-L12/HFP-G13A14	(1:2)
HFP-A	HFP-L12/HFP-G5A6	(1:2)
HFP-AP	HFP-F8/HFP-L9G10	(1:2)

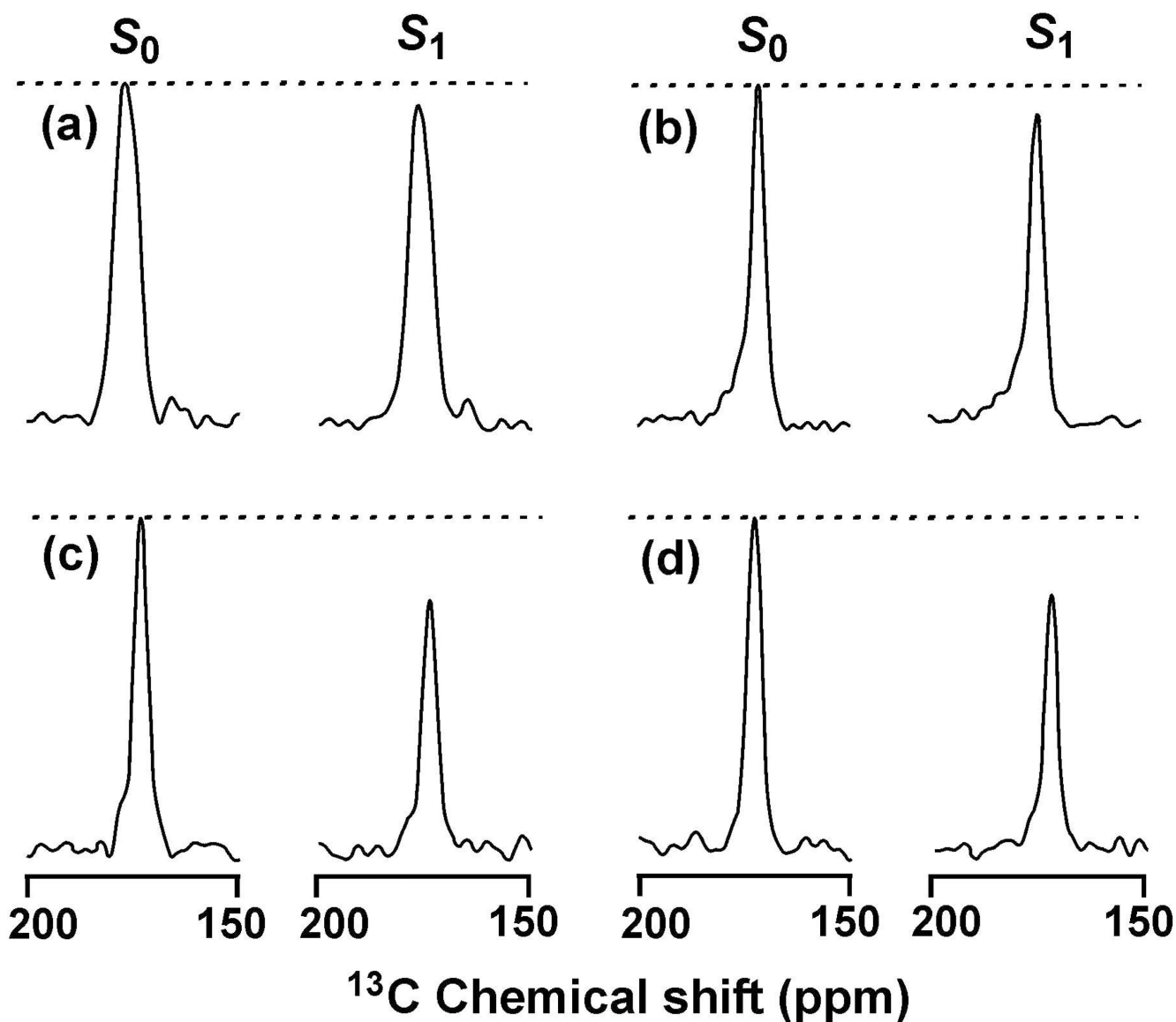
## (c) Registries

	Parallel 1→17/1→17	Parallel 2→17/1→16	Antiparallel 16→1/1→16	Antiparallel 17→1/1→17
HFP-P	AVGIGALFLGFLGAAGS AVGIGALFLGFLGAAGS	AVGIGALFLGFLGAAGS AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS
HFP-A	AVGI GALFLGFLGAAGS AVGIGALFLGFLGAAGS	AVGI GALFLGFLGAAGS AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS
HFP-AP	AVGIGALFLGFLGAAGS AVGIGALFLGFLGAAGS	AVGIGALFLGFLGAAGS AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS	SGAAGLFLGFLFLAGIGVA AVGIGALFLGFLGAAGS

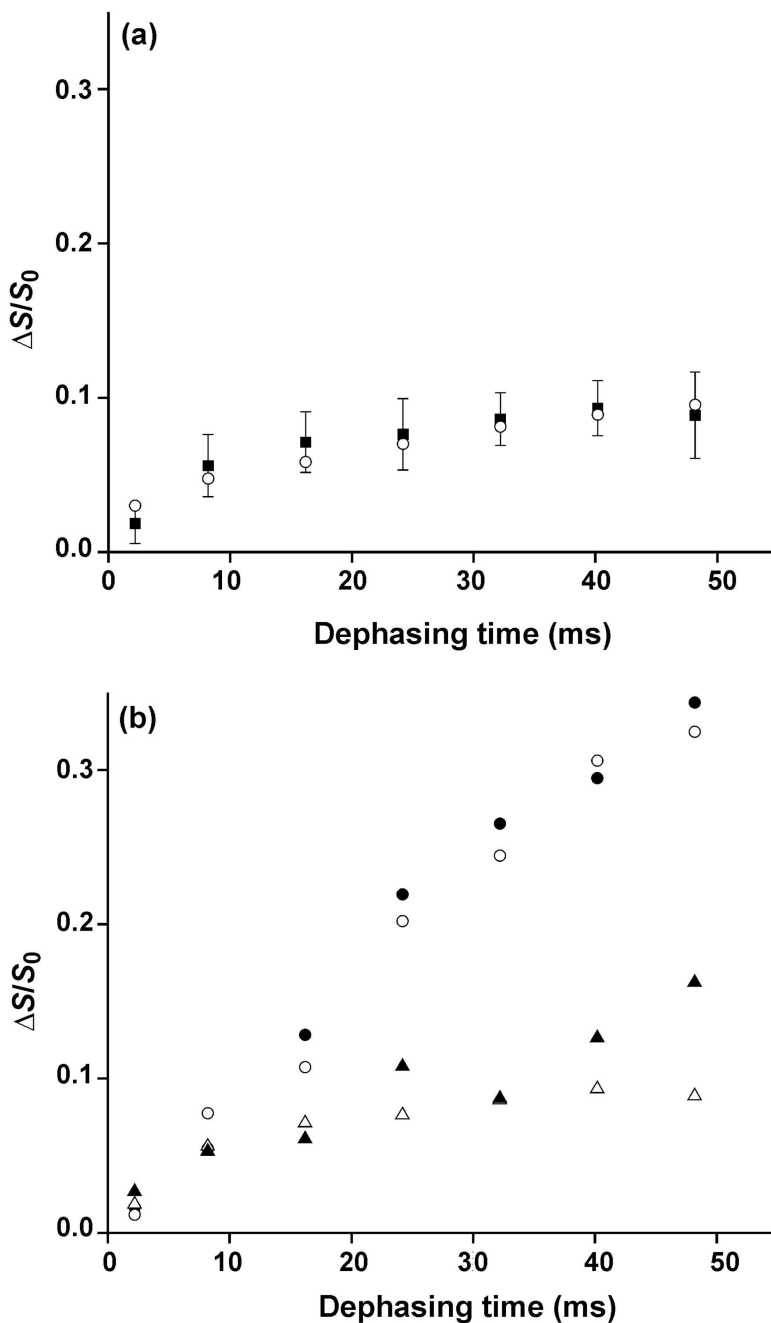
**Figure 1.**

(a) HFPs where red and blue correspond to  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled residues, respectively. (b) HFP-NC, HFP-P, HFP-A, and HFP-AP were SSNMR samples which each contained a mixture of  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled peptides in a 1:2 mol ratio. The HFP-NC sample was a mixture of HFP-F8 and HFP-A6L7 that had been lyophilized separately. The other samples were membrane-associated HFPs that formed  $\beta$  sheet structure with a molecular mixture of  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled peptides in the sample. (c) Registries probed by the SSNMR REDOR experiments and labeled  $^{13}\text{C}$ /labeled  $^{15}\text{N}$  proximities for the membrane-associated HFPs in these registries. Consideration of residue 1→16 or 1→17 registries is based on the fully extended conformation of this HFP region. For parallel sheets, there is CO(residue  $i$ ) – HN(residue  $i+1$ ) hydrogen bonding of adjacent strands.

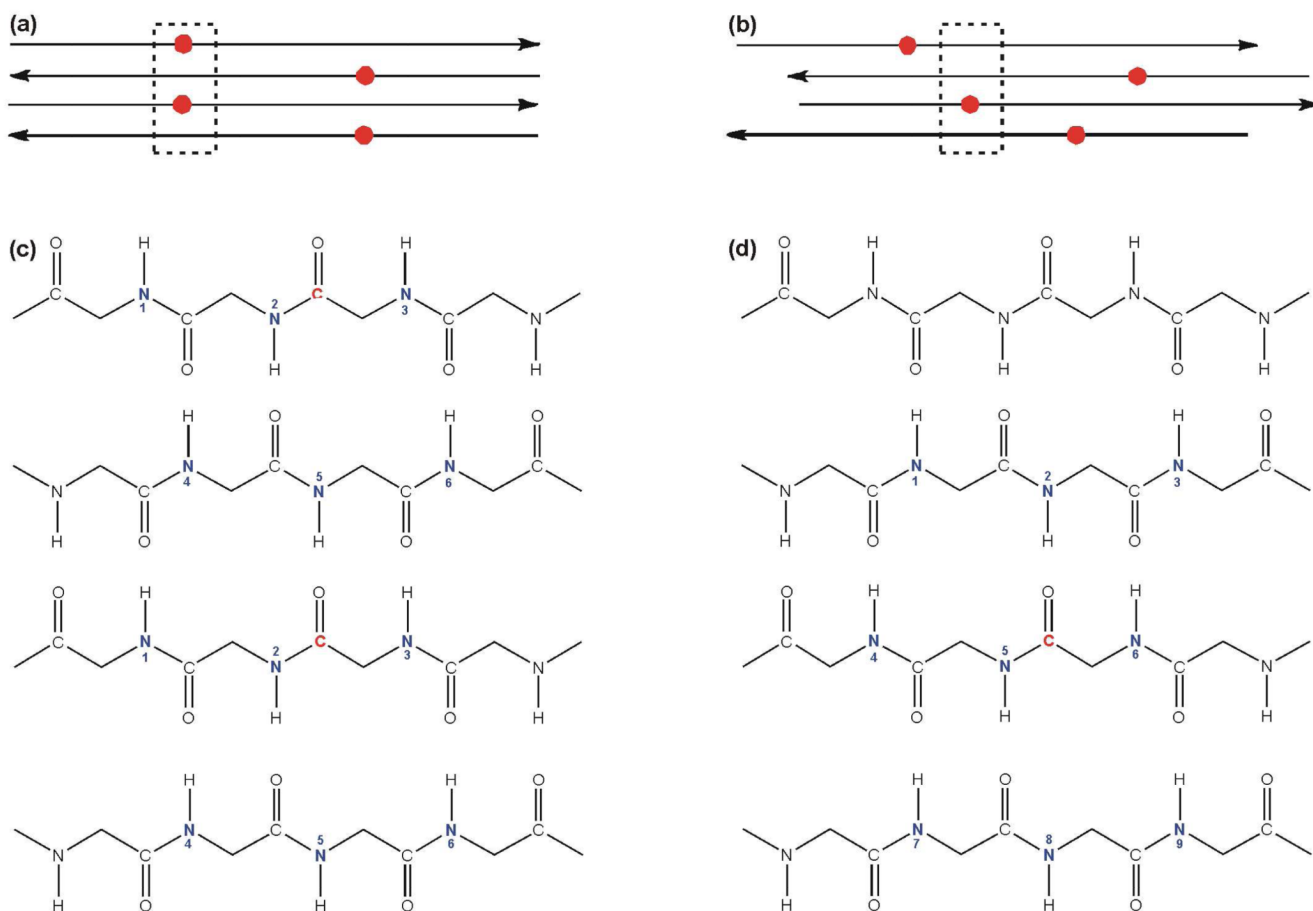




**Figure 2.** REDOR  $S_0$  and  $S_1$   $^{13}\text{C}$  SSNMR spectra at 32.2 ms dephasing time for (a) HFP-NC, (b) HFP-P, (c) HFP-A, or (d) HFP-AP. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 38624; (b) 23488; (c) 24914; or (d) 14240 scans. Relatively narrow  $^{13}\text{CO}$  signals were observed in the HFP-P, HFP-A, and HFP-AP samples because the HFPs were membrane-associated with predominant  $\beta$  sheet conformation at the labeled  $^{13}\text{CO}$  site. A broader  $^{13}\text{CO}$  signal was observed in the HFP-NC sample because there was no membrane and there were populations of lyophilized HFP with either  $\alpha$  helical or  $\beta$  sheet conformation at the labeled  $^{13}\text{CO}$  site.



**Figure 3.** (a) Plot of REDOR  $(\Delta S/S_0)^{exp}$  (filled squares with error bars) and  $(\Delta S/S_0)^{sim}$  (open circles) vs dephasing time for the lyophilized HFP-NC sample. The  $(\Delta S/S_0)^{sim}$  were calculated using a mixture of n.a.d. models with fractional populations:  $\alpha$  helical, 0.5; min  $\beta$  sheet, 0.25; max  $\beta$  sheet, 0.25. (b) Plots of  $(\Delta S/S_0)^{exp}$  vs dephasing time for: HFP-NC, open triangles; HFP-P, filled triangles; HFP-A, open circles; HFP-AP, filled circles. The typical  $\sigma^{exp}$  is  $\pm 0.02$ . Variation of  $\pm 0.02$  in  $(\Delta S/S_0)^{exp}$  was also observed between two different preparations of the same sample type, e.g. HFP-A.



**Figure 4.**

(a, b) Schematic diagrams of the HFP-F8 region of the HFP-NC sample in antiparallel  $\beta$  sheet structure with labeled  $^{13}\text{CO}$ s represented as red circles. Panel a shows a model that is fully constrained to a single registry while panel b shows multiple registries. (c, d)  $\beta$  sheet backbone representations of the respective boxed regions of panels a and b with labeled  $^{13}\text{CO}$ s in red and possible n.a.  $^{15}\text{N}$  sites in blue, i.e. sites for which a n.a.  $^{15}\text{N}$  is within  $7 \text{ \AA}$  of a labeled  $^{13}\text{CO}$ . A particular spin geometry will have only one  $^{15}\text{N}$ . The min n.a.d. model is shown in panel c and each spin geometry will have either one labeled  $^{13}\text{CO}$  and one n.a.  $^{15}\text{N}$  (#1, 2 or 3) or two labeled  $^{13}\text{CO}$ s and one n.a.  $^{15}\text{N}$  (#4, 5, or 6). The max n.a.d. model is shown in panel d and each spin geometry will have one labeled  $^{13}\text{CO}$  and one n.a.  $^{15}\text{N}$ .

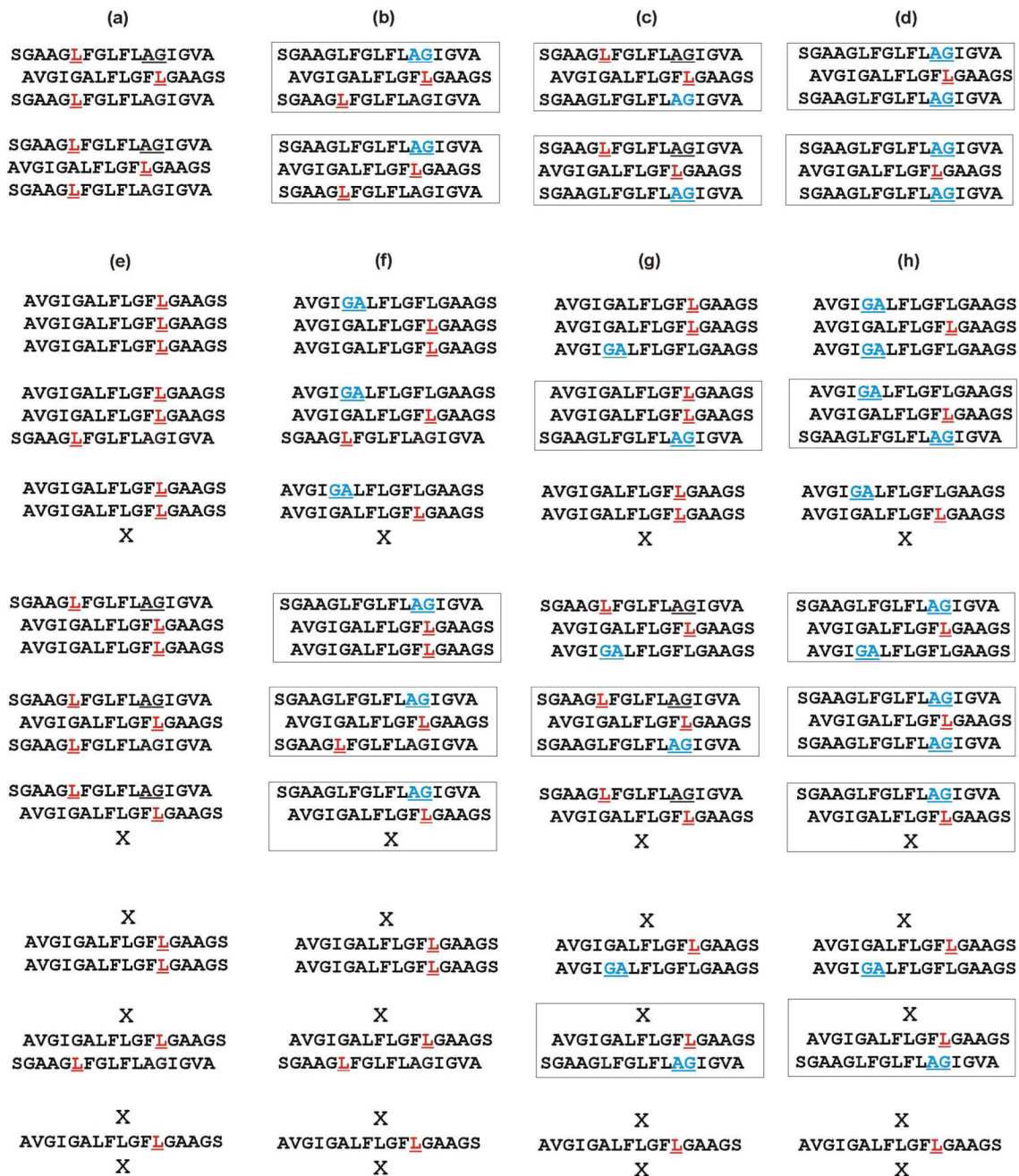
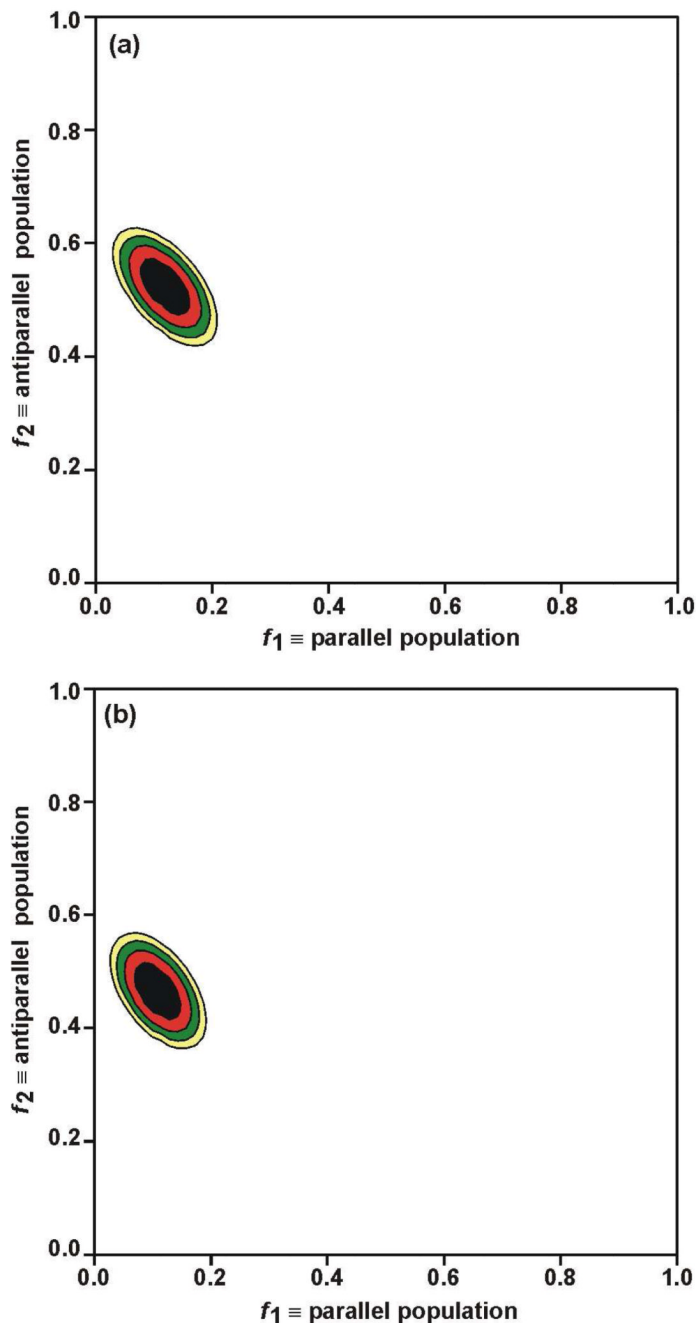


Figure 5.

Schematics of three adjacent HFPs for HFP-A, i.e.  $u = 2$ , in (a–d) fully constrained or (e–h) unconstrained models. Red and blue correspond to  $^{13}\text{C}$ O and  $^{15}\text{N}$  labeled residues, respectively, and the labeled  $^{13}\text{C}$ O in the middle strand was hydrogen bonded to the HN group of the residue in the top strand. Panels a–d display antiparallel  $16 \rightarrow 1/1 \rightarrow 16$  (top) or  $17 \rightarrow 1/1 \rightarrow 17$  (bottom) registries while panels e–h display parallel  $1 \rightarrow 17/1 \rightarrow 17$ , antiparallel  $16 \rightarrow 1/1 \rightarrow 16$ , and  $X$  registries where  $X$  refers to a registry for which the labeled  $r_{\text{CN}} > 7 \text{ \AA}$ , i.e. beyond the approximate detection limit of the SSNMR experiment, and which is not  $1 \rightarrow 17/1 \rightarrow 17$ ,  $2 \rightarrow 17/1 \rightarrow 16$ ,  $16 \rightarrow 1/1 \rightarrow 16$ , or  $17 \rightarrow 1/1 \rightarrow 17$ . Correspondence between columns and the index  $v$  are: a and e,  $v = 1$ ; b and f,  $v = 2$ ; c and g,  $v = 3$ ; d and h;  $v = 4$ .

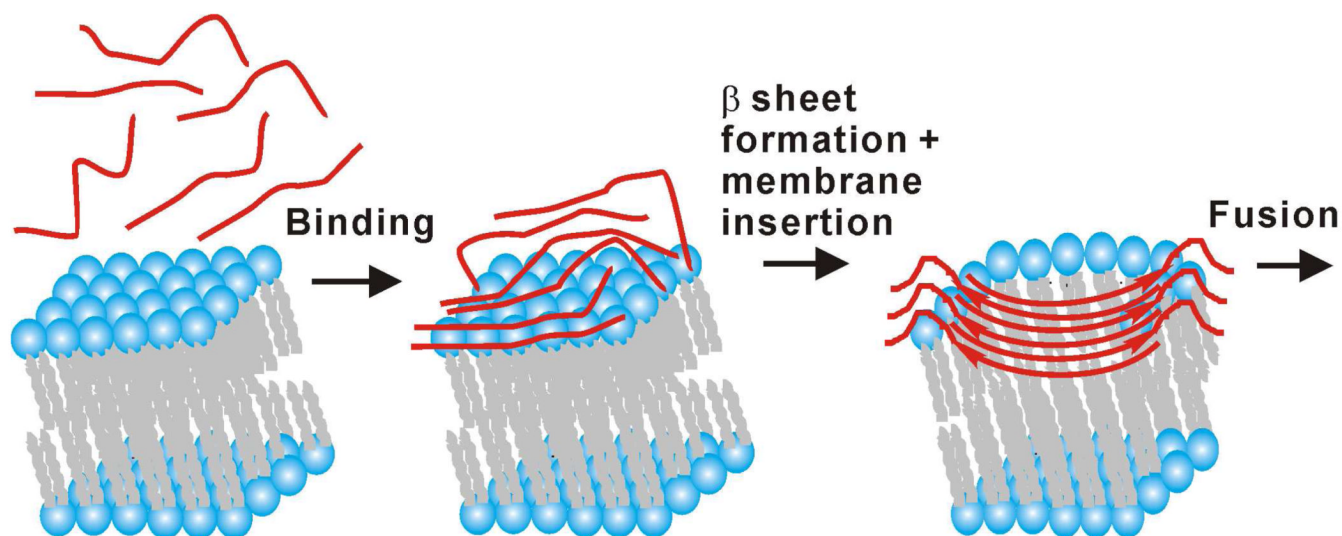
= 4. Both rows of three-strand arrangements in panels a–d correspond to  $t = 2$  and the row,  $t_1$ ,  $t_2$  correspondence in panels e–h is: row 1,  $t_1 = 1$ ,  $t_2 = 1$ ; row 2,  $t_1 = 1$ ,  $t_2 = 2$ ; row 3,  $t_1 = 1$ ,  $t_2 = 3$ ; row 4,  $t_1 = 2$ ,  $t_2 = 1$ ; row 5,  $t_1 = 2$ ,  $t_2 = 2$ ; row 6,  $t_1 = 2$ ,  $t_2 = 3$ ; row 7,  $t_1 = 3$ ,  $t_2 = 1$ ; row 8,  $t_1 = 3$ ,  $t_2 = 2$ ; row 9,  $t_1 = 3$ ,  $t_2 = 3$ . For each three-strand arrangement enclosed by a box, the  $\gamma_{1tuv}^{lab}(\tau)$  or  $\gamma_{1t_1t_2uv}^{lab}(\tau)$  were calculated by SIMPSON simulation. For arrangements with  $t$ ,  $t_1$ , or  $t_2 = 2$ , fitting to experiment was based on  $\gamma$ s that were the average of those calculated with  $16 \rightarrow 1/1 \rightarrow 16$  and  $17 \rightarrow 1/1 \rightarrow 17$  registries although the latter registry is not displayed in panels e–h. For any arrangement not enclosed by a box,  $\gamma_{1tuv}^{lab}(\tau) = 1$  or  $\gamma_{1t_1t_2uv}^{lab}(\tau) = 1$ .





**Figure 6.**

Contour plots of  $\chi^2$  vs  $f_1$  parallel and  $f_2$  antiparallel fractional populations for (a) fully constrained and (b) unconstrained models. In each plot,  $f_1$  is the sum of populations of  $1 \rightarrow 17/1 \rightarrow 17$  and  $2 \rightarrow 17/1 \rightarrow 16$  parallel registries and  $f_2$  is the sum of populations of  $16 \rightarrow 1/1 \rightarrow 16$  and  $17 \rightarrow 1/1 \rightarrow 17$  antiparallel registries. For plot a, the best-fit values were  $f_1 = 0.12 \pm 0.03$  and  $f_2 = 0.52 \pm 0.04$  with  $\chi^2_{min} = 11$ , and for plot b,  $f_1 = 0.11 \pm 0.03$ ,  $f_2 = 0.46 \pm 0.04$ , and  $\chi^2_{min} = 8$ . Parameter uncertainties were determined by the region of  $\chi^2$  within about three units of  $\chi^2_{min}$ . In plot a, the black, red, green, yellow, and white regions correspond to  $\chi^2 < 14$ ,  $14 < \chi^2 < 17$ ,  $17 < \chi^2 < 20$ ,  $20 < \chi^2 < 23$ , and  $\chi^2 > 23$  and in plot b, the regions correspond to  $\chi^2 < 11$ ,  $11 < \chi^2 < 14$ ,  $14 < \chi^2 < 17$ ,  $17 < \chi^2 < 20$ , and  $\chi^2 > 20$ .



**Figure 7.** Pictorial model of HFP (red lines) binding to membranes followed by antiparallel  $\beta$  sheet formation and membrane insertion and then fusion. Time increases from left-to-right. For reasons of clarity, some lipids are not shown in the right-most picture. Although there are no data yet on fusion peptide structure during HIV/host cell fusion, the antiparallel  $\beta$  sheet structure of the right-most picture is plausible because: (1) the structure is consistent with multiple trimers at the fusion site; and (2) the structure is membrane-inserted with deeper insertion positively correlated with increased membrane perturbation and vesicle fusion rate.

Table 1

## Indices and parameters

Index/parameter	Description	Values
$f_t$	fractional population of structure $t$	for membr. assoc. samples, determined by fitting
$j, k$	index for n.a. site $<7 \text{ \AA}$ from a labeled site: $j$ , n.a. $^{15}\text{N}$ near labeled $^{13}\text{CO}$ ; $k$ , n.a. $^{13}\text{CO}$ near labeled $^{15}\text{N}$	
$J, K$	number of n.a. sites $<7 \text{ \AA}$ from a labeled site: $J$ , n.a. $^{15}\text{N}$ s near labeled $^{13}\text{CO}$ ; $K$ , n.a. $^{13}\text{CO}$ s near labeled $^{15}\text{N}$	$\alpha$ helical structure, $J=9, K=10$ ; min $\beta$ sheet structure, $J=6, K=8$ ; max $\beta$ sheet structure, $J=9, K=12$
$l$	REDOR data type index	0 $\equiv$ no dipolar evolution 1 $\equiv$ dipolar evolution
$m$	datum index	1, 2, 3, 4, 5, 6, or 7
$S_j, S_{jp}, S_{jk}, S_{ju}$	REDOR signal intensity	determined by experiment or calculation
$t, t_1, t_2$	structural population index; for unconstrained model of membr. assoc. samples, $t_1$ indexes the top/middle registry and $t_2$ indexes the middle/bottom registry	for membr. assoc. samples: 1 $\equiv$ parallel registry 2 $\equiv$ antiparallel registry 3 $\equiv$ other "X" registry
$u$	membr. assoc. sample index	1 $\equiv$ HFP-P 2 $\equiv$ HFP-A 3 $\equiv$ HFP-AP
$v$	index for arrangement of three adjacent labeled HFPs in membr. assoc. samples – middle HFP has $^{13}\text{CO}$ labeling and the $^{13}\text{CO}$ is hydrogen bonded to HN of top HFP	1 $\equiv$ $^{13}\text{CO}$ HFP (top), $^{13}\text{CO}$ HFP (bottom) 2 $\equiv$ $^{15}\text{N}$ HFP (top), $^{13}\text{CO}$ HFP (bottom) 3 $\equiv$ $^{13}\text{CO}$ HFP (top), $^{15}\text{N}$ HFP (bottom) 4 $\equiv$ $^{15}\text{N}$ HFP(top), $^{15}\text{N}$ HFP(bottom)
$w_v$	fractional population of arrangement of three adjacent labeled HFPs in membr. assoc. samples	fully constrained model: $w_1 = 1/9, w_2 = 2/9, w_3 = 2/9, w_4 = 4/9$ ; unconstrained model: $w_1 = 1/81, w_2 = 2/81, w_3 = 2/81, w_4 = 4/81$
$\gamma_{iN}^{na}(\tau), \gamma_{iC}^{na}(\tau)$	$S(\tau)/S_0$ : $\gamma_{iN}(\tau)$ , labeled $^{13}\text{CO}$ -n.a. $^{15}\text{N}$ ; $\gamma_{iC}(\tau)$ , n.a. $^{13}\text{CO}$ -labeled $^{15}\text{N}$	$\gamma_{0N}(\tau) = 1; \gamma_{0C}(\tau) = 1; \gamma_{1N}(\tau)$ and $\gamma_{1C}(\tau)$ determined by calculation
$\gamma_{luv}^{lab}(\tau), \gamma_{l_1l_2uv}^{lab}(\tau)$	$S_{lu}^{lab}(\tau)/S_0^{lab}$ for arrangement of labeled $^{13}\text{CO}$ and $^{15}\text{N}$ nuclei: $\gamma_{luv}^{lab}(\tau)$ , fully constrained model; $\gamma_{l_1l_2uv}^{lab}(\tau)$ , unconstrained model	$\gamma_{0uv}^{lab}(\tau) = 1; \gamma_{0l_1l_2uv}^{lab}(\tau) = 1; \gamma_{1uv}^{lab}(\tau)$ and $\gamma_{1l_1l_2uv}^{lab}(\tau)$ determined by calculation
$\tau$	REDOR dephasing time	2.2, 8.2, 16.2, 24.2, 32.2, 40.2, or 48.2 ms