J|A|C|S
ARTICLES

Published on Web 03/28/2008

Solid-State NMR Spectroscopy of Human Immunodeficiency Virus Fusion Peptides Associated with Host-Cell-Like Membranes: 2D Correlation Spectra and Distance Measurements Support a Fully Extended Conformation and Models for Specific Antiparallel Strand Registries

Wei Qiang, Michele L. Bodner, and David P. Weliky*

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

Received September 20, 2007; E-mail: weliky@chemistry.msu.edu

Abstract: The human immunodeficiency virus (HIV) is "enveloped" by a membrane, and infection of a host cell begins with fusion between viral and target cell membranes. Fusion is catalyzed by the HIV gp41 protein which contains a functionally critical ~20-residue apolar "fusion peptide" (HFP) that associates with target cell membranes. In this study, chemically synthesized HFPs were associated with host-cell-like membranes and had "scatter-uniform" labeling (SUL), that is, only one residue of each amino acid type was U−^{13}C, ^{15}N labeled. For the first sixteen HFP residues, an unambiguous ^{13}C chemical shift assignment was derived from 2D ^{13}C/^{13}C correlation spectra with short mixing times, and the shifts were consistent with continuous β-strand conformation. ^{13}C−^{13}C contacts between residues on adjacent strands were derived from correlation spectra with long mixing times and suggested close proximity of the following residues: Ala-6/Gly-10, Ala-6/Phe-11, and Ile-4/Gly-13. Specific antiparallel β-strand registries were further tested using a set of HFPs that were ^{13}CO-labeled at Ala-14 and ^{15}N-labeled at either Val-2, Gly-3, Ile-4, or Gly-5. The solid-state NMR data were fit with 50–60% population of antiparallel HFP with either Ala-14/Gly-3 or Ala-14/Ile-4 registries and 40–50% population of structures not specified by the NMR experiments. The first two registries correlated with intermolecular hydrogen bonding of 15–16 apolar N-terminal residues and this hydrogen-bonding pattern would be consistent with a predominant location of these residues in the hydrophobic membrane interior. To our knowledge, these results provide the first residue-specific structural models for membrane-associated HFP in its β-strand conformation.

1. Introduction

Many viruses important in disease including human immunodeficiency virus (HIV) are "enveloped" by a membrane which is obtained after budding from a host cell. Infection of a new cell is accomplished by fusion between viral and target cell membranes with the end result of release of the viral nucleocapsid into the host cell cytoplasm. Fusion is catalyzed by the gp41 integral membrane protein of HIV that contains a ~170-residue "ectodomain" which lies outside the virus and has a ~20-residue apolar "fusion peptide" (HFP) at its N-terminus.[1,2] The HFP is believed to bind to the host cell membrane and to play an important role in fusion catalysis.[3–6] Peptides with the HFP sequence catalyze fusion between unilamellar lipid vesicles and the experimental correlation between the mutation/fusion activity relationships of HIV, and HFP-induced fusion provides evidence that the HFP is a useful model fusion system.[3–5,7–9]

X-ray and liquid-state nuclear magnetic resonance (NMR) structures have been determined for the "soluble ectodomain" which lacked the HFP and was soluble in nondetergent containing aqueous solution. Residues 30–147 of this domain had a trimeric bundle structure.[1,10,11] The structure of monomeric HFP in detergent micelles has been determined using liquid-state NMR and showed α helical structure which may be continuous between Ile-4 and Met-19.[12–15] Biophysical tech-

(1) Eckert, D. M.; Kim, P. S. *Annu. Rev. Biochem.* **2001**, *70*, 777–810.
(2) Jahn, R.; Lang, T.; Sudhof, T. C. *Cell* **2003**, *112*, 519–533.
(3) Freed, E. O.; Myers, D. J.; Risser, R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 4650–4654.
(4) Freed, E. O.; Delwart, E. L.; Buchschacher, G. L., Jr.; Panganiban, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 70–74.
(5) Schaal, H.; Klein, M.; Gehrmann, P.; Adams, O.; Scheid, A. *J. Virol.* **1995**, *69*, 3308–3314.
(6) Delahunty, M. D.; Rhee, I.; Freed, E. O.; Bonifacino, J. S. *Virology* **1996**, *218*, 94–102.
(7) Durell, S. R.; Martin, I.; Ruysschaert, J. M.; Shai, Y.; Blumenthal, R. *Mol. Membr. Biol.* **1997**, *14*, 97–112.
(8) Pereira, F. B.; Goni, F. M.; Muga, A.; Nieva, J. L. *Biophys. J.* **1997**, *73*, 1977–1986.
(9) Kliger, Y.; Aharoni, A.; Rapaport, D.; Jones, P.; Blumenthal, R.; Shai, Y. *J. Biol. Chem.* **1997**, *272*, 13496–13505.
(10) Tan, K.; Liu, J.; Wang, J.; Shen, S.; Lu, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12303–12308.
(11) Caffrey, M.; Cai, M.; Kaufman, J.; Stahl, S. J.; Wingfield, P. T.; Covell, D. G.; Gronenborn, A. M.; Clore, G. M. *EMBO J.* **1998**, *17*, 4572–4584.
(12) Chang, D. K.; Cheng, S. F.; Chien, W. J. *J. Virol.* **1997**, *71*, 6593–6602.
(13) Morris, K. F.; Gao, X. F.; Wong, T. C. *Biochim. Biophys. Acta* **2004**, *1667*, 67–81.
(14) Jaroniec, C. P.; Kaufman, J. D.; Stahl, S. J.; Viard, M.; Blumenthal, R.; Wingfield, P. T.; Bax, A. *Biochemistry* **2005**, *44*, 16167–16180.
(15) Li, Y. L.; Tamm, L. K. *Biophys. J.* **2007**, *93*, 876–885.

niques including infrared, circular dichroism, and solid-state NMR spectroscopies have been used to investigate the conformation of membrane-associated HFP.[8,16,17] Distinct populations of HFP were observed with either predominant α helical or β-strand conformations, and the relative ratio of these two populations was dependent on the peptide/lipid ratio, the membrane composition, and the concentrations of ions such as Ca^{2+}.[16,17] As one specific example, α helical conformation was favored in membranes lacking cholesterol and β-strand conformation was favored in membranes containing the ∼30 mol% cholesterol typical for host cells of HIV.[18–21]

For peptides and proteins, solid-state NMR has had great utility in developing and testing local structural models primarily through incorporation of a few specific isotopic labels and measurement of a few distances or torsion angles.[22–26] In a few cases, analysis of many differently labeled samples has led to more global structures.[27,28] During the past ten years, there has been a significant progress in the study of U−^{15}N, and U−^{13}C,^{15}N-labeled solid peptides and proteins with the goal of sequential ^{13}C, ^{15}N, and possibly ^{1}H assignments and the goal of three-dimensional structures based on chemical shifts and internuclear distance and torsion angle constraints. For example, there have been backbone structures of 20–60 residue peptides and proteins which were membrane-associated and strongly aligned in the NMR field.[29] In addition, there have been unambiguous sequential assignments of U−^{13}C,^{15}N solid peptides and small proteins determined from 2D and 3D magic angle spinning (MAS) correlation spectra as well as a few 3D structures based on these types of spectra.[30–40] In many cases, the peptides and proteins were in microcrystalline form with high structural homogeneity and narrow linewidths (e.g., ∼0.5 ppm) and high concentrations (e.g., ∼100 mM).[41] In a few cases, the peptides or proteins were in well-ordered fibrillar forms or were tightly bound to other proteins. For one case, the 52-residue membrane protein phospholamban was incorporated into membranes at protein/lipid of ∼0.05.[42] The solid-state NMR structure contained a helical C-terminal transmembrane domain and a disordered N-terminal domain and differed from the liquid-state NMR structure in micelles for which the N-terminal domain was helical and at right angles to the C-terminal helix.[43]

Analysis of 2D ^{13}C/^{13}C and ^{15}N/^{13}C spectra for membrane-associated HFP which was uniformly labeled yielded amino acid-type rather than sequential assignment because of large spectral overlap of different crosspeaks.[44] This overlap was due to the 1–3 ppm spectral linewidths and to the redundancy of amino acid types in the sequence, for example, six glycines and five alanines. Similar linewidths were observed in the spectra of the 40-residue β amyloid peptide in its fibrillar form and an unambiguous assignment was achieved with peptides synthesized with "scatter-uniform" labeling (SUL) in which only one residue of each amino acid type was U−^{13}C,^{15}N labeled.[45] The assignment was based in large part on ^{13}C/^{13}C correlation spectra with short (≤10 ms) mixing times that yielded only intraresidue crosspeaks whose chemical shifts could be assigned from the characteristic shifts of each amino acid. The chemical shifts were then correlated with residue-specific conformation, and β-strand and non-β-strand regions of the β-amyloid sequence were identified. The structural arrangements of adjacent SUL β-amyloid peptides were also determined in part from ^{13}C/^{13}C spectra with long (≥500 ms) mixing times for which crosspeaks could be observed between ^{13}C separated by up to 7 Å.[46] A similar approach was applied to membrane-associated SUL peptides representing the transmembrane domain of the HIV Vpu ion channel.[47]

In the present study, ^{13}C/^{13}C correlation spectra were obtained for SUL-HFPs associated with host-cell-like membranes and an unambiguous ^{13}C assignment was achieved for all of the labeled residues. In addition, interpeptide contacts were determined from SUL spectra and led to specific tertiary structure models which were subsequently tested and validated with internuclear distance measurements on specifically labeled HFPs. To our knowledge, these results provide the first residue-specific

(16) Rafalski, M.; Lear, J. D.; DeGrado, W. F. *Biochemistry* **1990**, *29*, 7917–7922.
(17) Nieva, J. L.; Nir, S.; Muga, A.; Goni, F. M.; Wilschut, J. *Biochemistry* **1994**, *33*, 3201–3209.
(18) Aloia, R. C.; Tian, H.; Jensen, F. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5181–5185.
(19) Wasniewski, C. M.; Parkanzky, P. D.; Bodner, M. L.; Weliky, D. P. *Chem. Phys. Lipids* **2004**, *132*, 89–100.
(20) Zheng, Z.; Yang, R.; Bodner, M. L.; Weliky, D. P. *Biochemistry* **2006**, *45*, 12960–12975.
(21) Brugger, B.; Glass, B.; Haberkant, P.; Leibrecht, I.; Wieland, F. T.; Krasslich, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2641–2646.
(22) Long, J. R.; Shaw, W. J.; Stayton, P. S.; Drobny, G. P. *Biochemistry* **2001**, *40*, 15451–15455.
(23) Murphy, O. J., III; Kovacs, F. A.; Sicard, E. L.; Thompson, L. K. *Biochemistry* **2001**, *40*, 1358–1366.
(24) Smith, S. O.; Eilers, M.; Song, D.; Crocker, E.; Ying, W. W.; Groesbeek, M.; Metz, G.; Ziliox, M.; Aimoto, S. *Biophys. J.* **2002**, *82*, 2476–2486.
(25) Toke, O.; Maloy, W. L.; Kim, S. J.; Blazyk, J.; Schaefer, J. *Biophys. J.* **2004**, *87*, 662–674.
(26) Abu-Baker, S.; Lorigan, G. A. *Biochemistry* **2006**, *45*, 13312–13322.
(27) Ketchem, R. R.; Hu, W.; Cross, T. A. *Science* **1993**, *261*, 1457–60.
(28) Lansbury, P. T., Jr.; Costa, P. R.; Griffiths, J. M.; Simon, E. J.; Auger, M.; Halverson, K. J.; Kocisko, D. A.; Hendsch, Z. S.; Ashburn, T. T.; Spencer, R. G.; Tidor, B.; Griffin, R. G. *Nat. Struct. Biol.* **1995**, *2*, 990–998.
(29) Marassi, F. M.; Opella, S. J. *Protein Sci.* **2003**, *12*, 403–411.
(30) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. *Chembiochem* **2001**, *2*, 272–281.
(31) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98–102.
(32) Bockmann, A.; Lange, A.; Galinier, A.; Luca, S.; Giraud, N.; Juy, M.; Heise, H.; Montserret, R.; Penin, F.; Baldus, M. *J. Biomol. NMR* **2003**, *27*, 323–339.
(33) Zech, S. G.; Wand, A. D.; McDermott, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 8618–8626.
(34) Marulanda, D.; Tasayco, M. L.; Cataldi, M.; Arriaran, V.; Polenova, T. *J. Phys. Chem. B* **2005**, *109*, 18135–18145.
(35) Franks, W. T.; Zhou, D. H.; Wylie, B. J.; Money, B. G.; Graesser, D. T.; Frericks, H. L.; Sahota, G.; Rienstra, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 12291–12305.
(36) Lange, A.; Giller, K.; Hornig, S.; Martin-Eauclaire, M. F.; Pongs, O.; Becker, S.; Baldus, M. *Nature* **2006**, *440*, 959–962.
(37) Reuther, G.; Tan, K. T.; Kohler, J.; Nowak, C.; Pampel, A.; Arnold, K.; Kuhlmann, J.; Waldmann, H.; Huster, D. *Angew. Chem. Intl. Ed.* **2006**, *45*, 5387–5390.
(38) Jaroniec, C. P.; MacPhee, C. E.; Astrof, N. S.; Dobson, C. M.; Griffin, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16748–16753.
(39) van Gammeren, A. J.; Hulsbergen, F. B.; Hollander, J. G.; de Groot, H. J. M. *J. Biomol. NMR* **2005**, *31*, 279–293.
(40) Todokoro, Y.; Yumen, I.; Fukushima, K.; Kang, S. W.; Park, J. S.; Kohno, T.; Wakamatsu, K.; Akutsu, H.; Fujiwara, T. *Biophys. J.* **2006**, *91*, 1368–1379.
(41) Andersson, K. M.; Hovmöller, S. *Acta Crystallogr. D* **2000**, *56*, 789–790.
(42) Andronesi, O. C.; Becker, S.; Seidel, K.; Heise, H.; Young, H. S.; Baldus, M. *J. Am. Chem. Soc.* **2005**, *127*, 12965–12974.
(43) Zamoon, J.; Mascioni, A.; Thomas, D. D.; Veglia, G. *Biophys. J.* **2003**, *85*, 2589–2598.
(44) Bodner, M. L.; Gabrys, C. M.; Struppe, J. O.; Weliky, D. P. *J. Chem. Phys.* **2008**, *128*, 052319.
(45) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742–16747.
(46) Paravastu, A. K.; Petkova, A. T.; Tycko, R. *Biophys. J.* **2006**, *90*, 4618–4629.
(47) Sharpe, S.; Yau, W. M.; Tycko, R. *Biochemistry* **2006**, *45*, 918–933.

structural model for HFP in its β-strand form conformation which is dominant when HFP associates with membranes whose lipid and headgroup composition is comparable to that of host cells of the virus.[18–21,48,49] There have been previous measurements of some of the backbone ^{13}CO (carbonyl) chemical shifts of specifically labeled HFPs but these data did not distinguish between the fully extended and hairpin structural models which have been proposed in the literature.[48,50,51] There are several conserved glycines in the sequence and either hairpin or fully extended conformations are plausible. In addition, there has been detection of distance proximity between ^{13}CO nuclei on one strand and ^{15}N nuclei on an adjacent strand in samples containing a HFP with ^{13}CO labeling at three consecutive residues and a HFP with ^{15}N labeling at three consecutive residues.[52] This study suggested a mixture of parallel and antiparallel arrangements but the registries were not clearly defined in part because of the multiple labels.

The number or number distribution of HFPs in the β-strand oligomer is not known although there is evidence that the number is small (<100). Evidence supporting the small size includes narrower linewidths in unfrozen samples relative to frozen samples and the 5–6 Å distances between the ^{31}Ps in the lipid headgroups and the ^{13}COs of the Ala-14 to Gly-16 residues.[53,54] In larger aggregates, thermally induced motional narrowing effects will be minimized and most of the HFPs would be segregated from the membrane lipids. The biological relevance of small HFP aggregates is supported by experimental and modeling evidence that there are clusters of gp41 trimers at the fusion site.[4,55]

2. Experimental Methods

Peptides. Resins and 9-fluorenylmethoxycarbonyl (FMOC) protected amino acids were purchased from Peptides International Inc. (Louisville, KY). Labeled amino acids were obtained from Cambridge Isotope Laboratories (Andover, MA) and were FMOC-protected using literature methods.[56] Peptide sequences and labeling are listed in Figure 1. All peptides began with the 23-residue N-terminal residues of gp41 (AVGIGALFLGFLGAAGSTMGARS) followed by a non-native C-terminal sequence that contained lysines to improve HFP aqueous solubility and tryptophan as a A_{280} chromophore. Although some HFPs contained a cysteine, they were predominantly non-cross-linked as judged by monomeric molecular weight in analysis of ultracentrifugation data.[57,58] HFP-A,B,C,D,E,F had SUL and HFP-G,H,I,J,K had selective ^{13}CO and ^{15}N labeling.

HFP-A,B,C,D,E were made with a peptide synthesizer (Applied Biosystems 431A, Foster City, CA), and HFP-F,G,H,I,J,K were synthesized using a 15 mL manual reaction vessel (Peptides International, Louisville, KY) and FMOC chemistry. Peptides were cleaved from the resin for 2–3 h using either a mixture of trifluoroacetic acid (TFA)/water/phenol/thioanisole/ethanedithiol/water in a 33:2:2:2:1 volume ratio or a mixture of TFA/thioanisole/ethanedithiol/anisole in a 90:5:3:2 volume ratio. TFA was removed from the cleavage filtrate with nitrogen gas and peptides were precipitated with cold *t*ert-butyl methyl ether. Peptides were purified by reversed-phased high performance liquid chromatography (HPLC) using a semipreparative C_{18} column and a water-acetonitrile gradient containing 0.1% TFA. Mass spectroscopy was used for peptide identification.

	1	6	11	16	21
HFP-A	AVGIGALFLGFLGAAGSTMGARSCKKKKKKW				
HFP-B	AVGIGALFLGFLGAAGSTMGARSCKKKKKKW				
HFP-C	AVGIGALFLGFLGAAGSTMGARSCKKKKKKW				
HFP-D	AVGIGALFLGFLGAAGSTMGARSCKKKKKKW				
HFP-E	AVGIGALFLGFLGAAGSTMGARSCKKKKKKW				
HFP-F	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				
HFP-G	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				
HFP-H	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				
HFP-I	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				
HFP-J	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				
HFP-K	AVGIGALFLGFLGAAGSTMGARSWKKKKKK				

Figure 1. Peptide sequences and labeling with blue, green, and red respectively, corresponding to ^{13}CO, ^{15}N, and U−^{13}C, ^{15}N labeling.

Preparation of Solid-State NMR Samples. HFP was incorporated into membranes in a manner comparable to that of functional fusion assays.[57] The samples were made with lipid and cholesterol mixtures reflecting the approximate lipid headgroup and cholesterol content of host cells infected by the HIV virus.[18] For HFP-A,B,C,D,E, the "LM3" mixture contained 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-L-serine] (POPS), phosphatidylinositol (PI), sphingomyelin, and cholesterol in a 10:5:2:1:2:10 molar ratio. LM3 contained the approximate lipid headgroup and cholesterol composition of membranes of host cells of HIV.[18,21] For HFP-F, G,H,I,J,K the "PC/PG/CHOL" mixture consisted of 1,2-di-*O*-tetradecyl-*sn*-glycero-3-phosphocholine (DTPC), 1,2-di-*O*-tetradecyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (DTPG), and cholesterol in a 8:2:5 molar ratio. Use of the ether-linked lipids DTPC and DTPG eliminated natural abundance lipid ^{13}CO signals and provided for more straightforward NMR analysis. Peptide conformation was not affected by the substitution.[20] Each sample preparation began with dissolution in chloroform of 30 total μmol of lipid and cholesterol. The chloroform was removed under a stream of nitrogen followed by overnight vacuum pumping. The lipid film was suspended in 2 mL of 5 mM *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid (HEPES) buffer with pH = 7.0 and 0.01% NaN_3 and was homogenized with ten freeze–thaw cycles. Large unilamellar vesicles were formed by extrusion through a 100 nm diameter polycarbonate filter (Avestin, Ottawa, ON). For most samples, a 0.8 μmol aliquot of HFP (as determined using ϵ_{280} = 5700 cm^{-1} M^{-1}) was dissolved in 2 mL of HEPES buffer, and the HFP and vesicle solutions were then gently vortexed together. The mixture was refrigerated overnight and ultracentrifuged at ~150000*g* for five hours. The membrane pellet with associated bound HFP was transferred to a 4 mm diameter MAS NMR rotor. The unbound HFP does not pellet under these conditions.[48]

Previous studies have shown that HFP forms β-strand oligomers or aggregates when associated with cholesterol-containing membranes.[52,59] Although HFP can aggregate in aqueous buffered solution under certain conditions, it is more biologically relevant that HFP be monomeric in solution prior to membrane binding so

(48) Yang, J.; Gabrys, C. M.; Weliky, D. P. *Biochemistry* **2001**, *40*, 8126–8137.
(49) Yang, J.; Parkanzky, P. D.; Bodner, M. L.; Duskin, C. G.; Weliky, D. P. *J. Magn. Reson.* **2002**, *159*, 101–110.
(50) Taylor, S. E.; Desbat, B.; Blaudez, D.; Jacobi, S.; Chi, L. F.; Fuchs, H.; Schwarz, G. *Biophys. Chem.* **2000**, *87*, 63–72.
(51) Haque, M. E.; Koppaka, V.; Axelsen, P. H.; Lentz, B. R. *Biophys. J.* **2005**, *89*, 3183–3194.
(52) Yang, J.; Weliky, D. P. *Biochemistry* **2003**, *42*, 11879–11890.
(53) Bodner, M. L.; Gabrys, C. M.; Parkanzky, P. D.; Yang, J.; Duskin, C. A.; Weliky, D. P. *Magn. Reson. Chem.* **2004**, *42*, 187–194.
(54) Qiang, W.; Yang, J.; Weliky, D. P. *Biochemistry* **2007**, *46*, 4997–5008.
(55) Munoz-Barroso, I.; Durell, S.; Sakaguchi, K.; Appella, E.; Blumenthal, R. *J. Cell Biol.* **1998**, *140*, 315–323.
(56) Lapatsanis, L.; Milias, G.; Froussios, K.; Kolovos, M. *Synthesis (Stuttgart)* **1983**, *8*, 671–673.
(57) Yang, R.; Prorok, M.; Castellino, F. J.; Weliky, D. P. *J. Am. Chem. Soc.* **2004**, *126*, 14722–14723.
(58) Bodner, M. L. Ph.D. Dissertation. Michigan State University, East Lansing, MI, 2006.
(59) Yang, J.; Prorok, M.; Castellino, F. J.; Weliky, D. P. *Biophys. J.* **2004**, *87*, 1951–1963.

that the oligomeric structure is a result of membrane binding.[48,60] A HFP construct containing three C-terminal lysines is predominantly monomeric in HEPES buffer with [HFP] $\approx$ 100 μM.[57] The HFP constructs of the present paper contained six lysines and should be monomeric at even higher concentrations. Most samples were made with $[HFP]_{initial} \approx 400$ μM but the effect of concentration was probed in a few selectively labeled samples by first dissolving the peptide in $\sim$30 mL of buffer so that $[HFP]_{initial} \approx 25$ μM.

NMR Experiments. Experiments were done on a 9.4 T solid-state NMR spectrometer (Varian Infinity Plus, Palo Alto, CA) with a temperature of –50 °C to enhance ^{13}C signal and to reduce motional averaging of dipolar couplings. It has previously been shown that HFP chemical shifts vary little as a function of temperature.[53] ^{13}C shifts were externally referenced to the methylene resonance of adamantane at 40.5 ppm.[61]

Proton-Driven Spin Diffusion (PDSD) Experiments. The probe had a double resonance configuration with ^{13}C and ^{1}H frequencies of 100.8 and 400.8 MHz, respectively. The PDSD pulse sequence contained an initial ^{1}H $\pi/2$ pulse followed by a ^{1}H$-^{13}$C cross polarization (CP), an evolution period t_1, a ^{13}C $\pi/2$ pulse that rotated the ^{13}C transverse magnetization to the longitudinal axis, a spin diffusion period τ during which ^{13}C magnetization was mixed among nearby ^{13}C nuclei, a second ^{13}C $\pi/2$ pulse that rotated the ^{13}C magnetization back to the transverse plane, and a detection period t_2. A 100 kHz ^{1}H decoupling field with two-pulse phase modulation (TPPM) was applied during t_1 and t_2, but not during τ.[62] The following parameters were typical for PDSD experiments: 6.8 kHz MAS frequency, 44–64 kHz ramp on the ^{13}C CP rf field; 62.5 kHz ^{1}H CP rf field; 2 ms CP contact time; 50 kHz ^{13}C $\pi/2$ pulse rf field; 25 μs t_1 dwell time; 200 t_1 values; 20 μs t_2 dwell time; and 1 s recycle delay. Hypercomplex data were obtained by acquiring two individual FIDs for each t_1 point with either a ^{13}C $(\pi/2)_x$ or $(\pi/2)_y$ pulse at the end of the t_1 evolution period.

Rotational-Echo Double-Resonance (REDOR) Experiments and Simulations. The triple resonance MAS probe was tuned to ^{13}C, ^{1}H, and ^{15}N frequencies of 100.8, 400.8, and 40.6 MHz, respectively, and the ^{13}C transmitter was at 152.4 ppm. The REDOR sequence contained in sequence: (1) a 44 kHz ^{1}H $\pi/2$ pulse; (2) 2.2 ms cross-polarization with 63 kHz ^{1}H field and 76–84 kHz ramped ^{13}C field; (3) a dephasing period of duration τ for which the "S_0" and "S_1" acquisitions contained 62 kHz ^{13}C π pulses at the end of each rotor cycle except the last cycle and for which the S_1 acquisition contained 27 kHz ^{15}N π pulses in the middle of rotor cycles; and (4) ^{13}C detection.[20,52,63–65] *XY*-8 phase cycling was applied to the ^{13}C and ^{15}N pulses during the dephasing period, TPPM ^{1}H decoupling of $\sim$95 kHz was applied during the dephasing and detection periods, the recycle delay was 1 s, and the MAS frequency was 8000 $\pm$ 2 Hz. REDOR experiments were calibrated using a lyophilized "I4" peptide with sequence AcAE-AAAKEAAAKEAAAKA-NH_2 and a ^{13}CO label at Ala-9 and a ^{15}N label at Ala-13. For the predominant α helical conformation of I4, the labeled ^{13}CO$-^{15}$N distance is $\sim$4.1 Å.[20,66]

The S_0 REDOR spectrum contained all ^{13}C signals while the S_1 spectrum had reduced signals from ^{13}C with proximal ^{15}N and therefore appreciable ^{13}C$-^{15}$N dipolar coupling (d). The equation $d = 3100/r^3$ expresses the relation between d in Hz and ^{13}C$-^{15}$N distance (r) in Å. The data analysis focused on integrated S_0 and S_1 intensities in the labeled ^{13}CO region that were denoted as "S_0" and "S_1", respectively. An experimental fractional dephasing $(\Delta S/S_0)^{exp} = (S_0^{exp} - S_1^{exp})/S_0^{exp}$ was calculated for each τ. The $(\Delta S/S_0)^{exp}$ provided the experimental basis for determination of d and r. The σ^{exp} uncertainty in $(\Delta S/S_0)^{exp}$ was calculated by

$$\sigma^{exp} = \frac{\sqrt{\left(S_0^2 \times \sigma_{S_1}^2\right) + \left(S_1^2 \times \sigma_{S_0}^2\right)}}{S_0^2} \quad (1)$$

where σ_{S_0} and σ_{S_1} are the experimental root-mean-squared noise of the S_0 and S_1 spectra, respectively.[67] All of the $(\Delta S/S_0)$ in this paper were calculated using integration over 1 ppm which was the region of maximum peak intensity. The effect of integration width was assessed by also doing integration over 3 ppm which represented the approximate full-width at half-maximum line width. There was a typical difference of 0.01 between the $(\Delta S/S_0)^{exp}$ determined with 3 ppm integration and the $(\Delta S/S_0)^{exp}$ determined with 1 ppm integration and for all data, the difference was less than the σ^{exp} determined with 1 ppm integration.

Calculations of $(\Delta S/S_0)$ as a function of spin geometry were denoted $(\Delta S/S_0)^{sim}$ and were made using the SIMPSON program.[68] The calculations were based on two or three spins where one of the spins was the Ala-14 ^{13}CO in a central β strand and the other one or two spins were labeled ^{15}N on adjacent strands. To make meaningful comparison between the $(\Delta S/S_0)^{sim}$ which were based only on labeled nuclei and experimental data which included contributions from both labeled and natural abundance nuclei, $(\Delta S/S_0)^{cor}$ values were calculated from the $(\Delta S/S_0)^{exp}$ and reflected removal of the natural abundance contribution. This contribution was estimated using the fractional natural abundances and known local ^{13}CO$-^{15}$N distances and associated dipolar couplings of peptides.[20] For each $(\Delta S/S_0)^{cor}$, a σ^{cor} was calculated and σ^{cor} were $\approx 1.4 \times \sigma^{exp}$. A detailed description of the calculation of $(\Delta S/S_0)^{cor}$ and σ^{cor} is provided in the Supporting Information.

There is experimental evidence from a previous study that the natural abundance correction factors are accurate. REDOR data were analyzed for a membrane-associated HIV fusion peptide with a ^{13}CO label at Leu-7 and a ^{15}N label at Phe-11.[20] Unlike the membranes used in the present study, the model membranes in the earlier study did not contain cholesterol and the Leu-7 ^{13}CO chemical shift was consistent with helical rather than strand conformation. It was also shown that the $(\Delta S/S_0)^{cor}$ derived from the REDOR data could be fitted well to a 4.1 $\pm$ 0.1 Å ^{13}CO$\cdots^{15}$N distance which is the expected distance between the Leu-7 and Phe-11 nuclei in an α helix. The natural abundance correction factor used in the earlier study is almost identical to the correction factor used in the present study. In addition, the earlier study included analysis of REDOR $(\Delta S/S_0)^{cor}$ values of the I4 model helical peptide which had a ^{13}CO label at residue 9 and a ^{15}N label at residue 13. The best-fit ^{13}CO$\cdots^{15}$N distance was 4.11 $\pm$ 0.01 Å and was consistent with the expected distance between these nuclei in an α helix. The correction factor for this model peptide was very similar to the one used in the present study.

Input parameters to the SIMPSON program included the ^{13}CO$-^{15}$N dipolar couplings, the Ala-14 ^{13}CO chemical shift and chemical shift anisotropy (CSA) principal values, and sets of Euler angles which reflected the orientations of ^{13}CO$-^{15}$N dipolar coupling and ^{13}CO CSA principal axis systems (PASs) in the fixed crystal frame. The ^{13}CO chemical shift was 175 ppm and CSA principal values were set to 241, 179, and 93 ppm, respectively.[69] Determination of Euler angles was based on atomic coordinates of the labeled nuclei and these coordinates were taken from crystal structure coordinates of outer membrane protein G (OMPG) (PDB

(60) Slepushkin, V. A.; Andreev, S. M.; Sidorova, M. V.; Melikyan, G. B.; Grigoriev, V. B.; Chumakov, V. M.; Grinfeldt, A. E.; Manukyan, R. A.; Karamov, E. V. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 9–18.
(61) Morcombe, C. R.; Zilm, K. W. *J. Magn. Reson.* **2003**, *162*, 479–486.
(62) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951–6958.
(63) McDowell, L. M.; Holl, S. M.; Qian, S. J.; Li, E.; Schaefer, J. *Biochemistry* **1993**, *32*, 4560–4563.
(64) Anderson, R. C.; Gullion, T.; Joers, J. M.; Shapiro, M.; Villhauer, E. B.; Weber, H. P. *J. Am. Chem. Soc.* **1995**, *117*, 10546–10550.
(65) Gullion, T. *Concepts Magn. Reson.* **1998**, *10*, 277–289.
(66) Long, H. W.; Tycko, R. *J. Am. Chem. Soc.* **1998**, *120*, 7039–7048.
(67) Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed.; McGraw-Hill: Boston, MA, 1992.
(68) Bak, M.; Rasmussen, J. T.; Nielsen, N. C. *J. Magn. Reson.* **2000**, *147*, 296–330.
(69) Gabrys, C. M.; Yang, J.; Weliky, D. P. *J. Biomol. NMR* **2003**, *26*, 49–68.

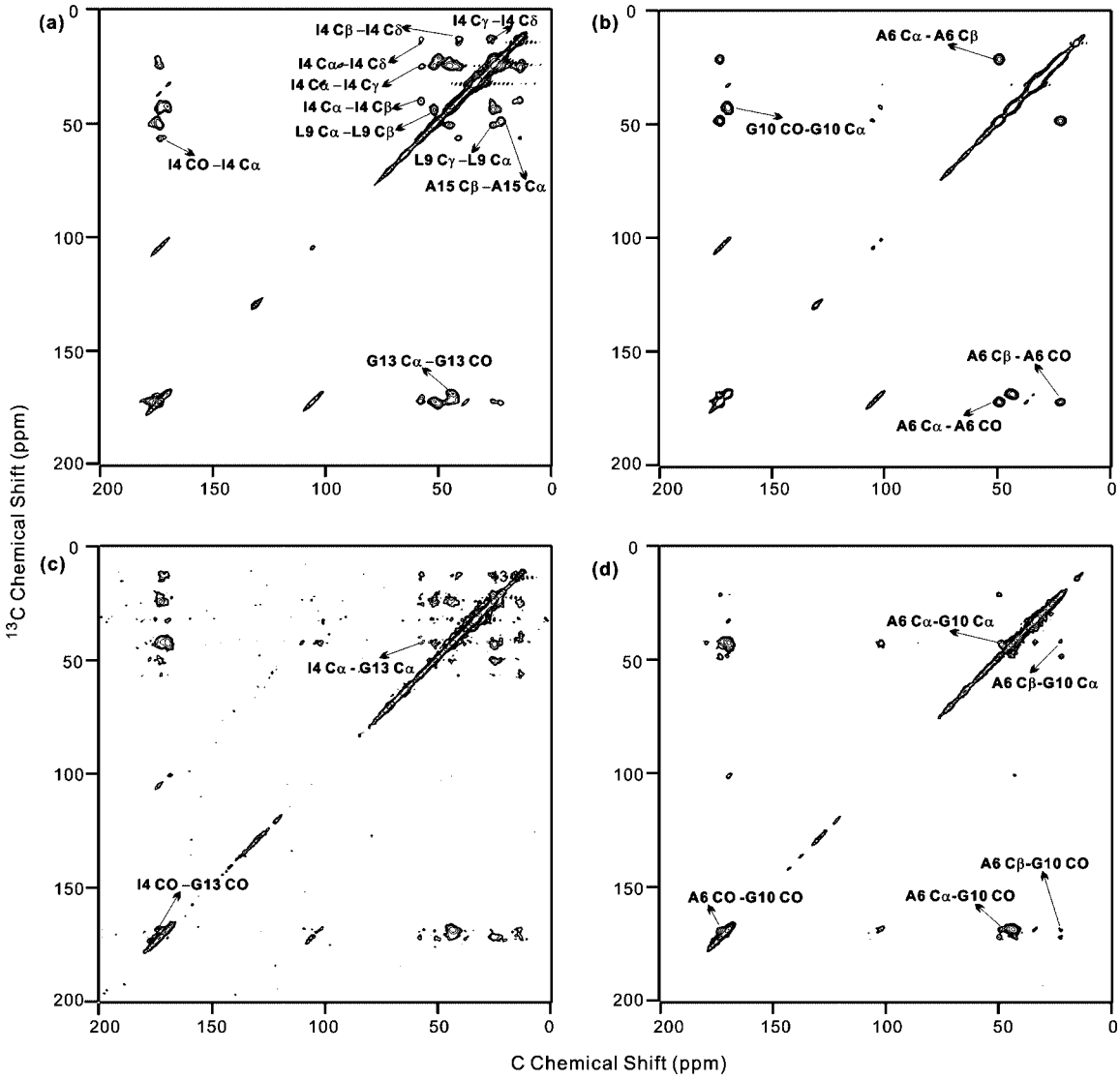

Figure 2. 2D $^{13}C-^{13}C$ PDSD spectra of membrane-associated (a, c) HFP-C and (b, d) HFP-D. The magnetization exchange time was either (a, b) 10 ms or (c, d) 1000 ms. All spectra were processed with 100 Hz Gaussian line broadening and baseline correction in the f_2 (horizontal) and f_1 (vertical) dimensions. The total numbers of scans were (a, b) 102400 and (c, d) 204800. Some of the (a, b) intraresidue and (c, d) inter-residue peak assignments are listed using the convention of assignment in f_2 − assignment in f_1.

file 2IWW).[70,71] OMPG was chosen because the REDOR experiments probed antiparallel β-strand structure in HFP and this was the predominant OMPG structural motif. After the ^{13}CO coordinates were obtained from a specific residue in OMPG, ^{15}N coordinates were obtained from nearby residues in the two adjacent strands. The Results section includes more detail about the specific choices of these nearby residues. For the two-spin simulations, the (α, β, γ) Euler angles of the dipolar coupling PAS were (0, 0, 0) and for the three-spin simulations, the angles for one dipolar PAS was (0, 0, 0) and for the other PAS were (0, θ, 0) where θ was the angle between two $^{13}CO-^{15}N$ vectors. The Euler angles for the ^{13}CO CSA PAS were calculated using the known orientation of the PAS relative to the ^{13}CO chemical bonds and the OMPG-derived orientation of these chemical bonds relative to the crystal frame.[72]

(70) Yildiz, O.; Vinothkumar, K. R.; Goswami, P.; Kuhlbrandt, W. *EMBO J.* **2006**, *25*, 3702–3713.
(71) Mehring, M. *Principles of high-resolution NMR in solids*, 2nd ed.; Springer: Berlin, 1983.
(72) Oas, T. G.; Hartzell, C. J.; McMahon, T. J.; Drobny, G. P.; Dahlquist, F. W. *J. Am. Chem. Soc.* **1987**, *109*, 5956–5962.

3. Results

PDSD Spectra and Chemical Shift Assignment. For the HFP-A,B,C,D,E,F samples, 2D PDSD spectra were obtained with exchange time $\tau = 10$, 100, 500, and 1000 ms. Figure 2 displays example PDSD spectra for (a, c) HFP-C and (b, d) HFP-D with $\tau =$ (a, b) 10, and (c, d) 1000 ms and Figure S2 in the Supporting Information contains representative slices from these four spectra. For spectra with $\tau = 10$ or 100 ms, (1) most possible intraresidue ^{13}C crosspeaks for labeled residues were observed and (2) no inter-residue ^{13}C crosspeaks were detected. Observation 1 was consistent with rapid exchange of ^{13}C magnetization within the intraresidue networks of directly bonded ^{13}C backbone and side chain nuclei which had $^{13}C-^{13}C$ dipole couplings >2 kHz. Observation 2 was due to the presence of at least one unlabeled residue between each pair of labeled residues in the SUL samples.[53] The resulting inter-residue $^{13}C-^{13}C$ distances were >4.5 Å and correlated with <100 Hz dipolar couplings and slow exchange of ^{13}C magnetization between labeled residues. Each group of crosspeaks which

Table 1. ^{13}C Chemical Shifts of LM3-Associated HFP[a]

residue	Cα	Cβ	Cγ	Cδ	C1	C2–6	CO
Ala-1	52.0	21.3					173.5
Val-2	60.5	36.2	22.3				174.7
Gly-3	45.5						171.5
Ile-4	59.6	42.9	28.7/18.8[b]	15.8			174.5
Gly-5	45.6						170.3
Ala-6	51.0	23.9					174.2
Leu-7	53.8	47.5	27.8	24.4			174.2
Phe-8	56.0	44.3			139.7	131.1	173.8
Leu-9	53.6	47.1	27.6	24.9			174.7
Gly-10	44.8						170.7
Phe-11	56.5	44.6			140.0	131.2	173.3
Leu-12	53.7	47.1	27.6				174.4
Gly-13	45.6						171.4
Ala-14	51.8	24.3					174.9
Ala-15	52.0	24.6					175.6
Gly-16	45.5						172.4
Ala-21	52.3	18.9					176.7

[a] Shifts are reported in ppm units with ± 0.5 ppm uncertainty. [b] The chemical shift of C(γ)H$_2$ is 28.7 ppm and the chemical shift of C(γ)H$_3$ is 18.8 ppm.

shared common chemical shifts was assigned to a specific labeled amino acid type (e.g., Ala CO/Cα, CO/Cβ, Cα/Cβ). The assignment was facilitated by knowledge of the typical chemical shift of each ^{13}C in a particular amino acid.[73] Because of SUL, the resulting ^{13}C chemical shift assignments were unambiguous (cf. Figure 2a,b and Table 1).

The assignment in Table 1 was based on crosspeaks in the spectra of LM3-associated HFP-A, HFP-B, HFP-C, HFP-D, and HFP-E. Analysis of the PC:PG:CHOL-associated HFP-F sample spectra yielded the following assignments (in ppm): Ile-4, Cα 59.6, Cβ 43.2, Cγ 28.5/18.1, Cδ 15.4, CO 174.4; Ala-6, Cα 51.4, Cβ 24.3, CO 175.0; Phe-11, Cα 56.7, Cβ 45.3, C1 138.2, *C*2–6 131.4, CO 174.0; Gly-13, Cα 45.4, CO 171.9. These residues were also labeled in the HFP-B, HFP-C, and HFP-D samples and the typical variation between an HFP-F shift and the corresponding shift in the other samples was 0.5 ppm and showed that there was little difference in shifts and presumably peptide structure between peptides with different C-terminal tags or between peptides associated with either LM3 or with PC:PG:CHOL membranes. After taking into account different chemical shift referencing, the ^{13}CO shifts in Table 1 were also consistent with earlier measurements using selectively labeled HFPs bound to LM3.[48]

Within the range of shifts for each amino acid-type Cα, Cβ, and CO, there are additional narrower subranges of shifts for helical and β-strand conformations. These ranges were first noted in spectra of solid homopolymers of amino acids of defined conformation and were defined with better precision by correlation of liquid-state ^{13}C chemical shifts and local conformation in proteins of known structure.[74,75] Further corroboration has been obtained with the close agreement between the ^{13}C chemical shifts of the same protein in soluble or microcrystalline form.[32,34,35,76] Figure 3 displays the differences between the experimental ^{13}C shifts and the consensus ^{13}C shifts expected for helical or β-strand conformations. For residues between

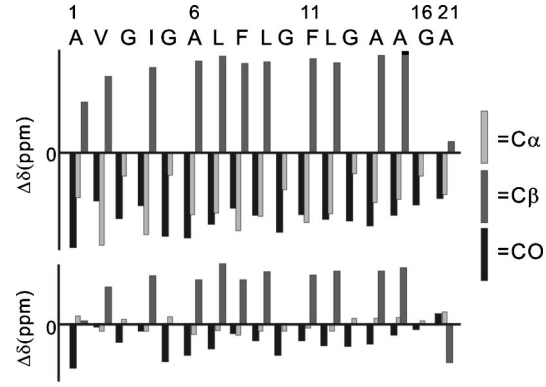

Figure 3. Differences in chemical shift (Δδ) between experimental ^{13}C chemical shifts for membrane-associated HFP and characteristic helical (top) or β-strand (bottom) ^{13}C shifts.[74] Each bar in the legend represents 3 ppm. There appears to be better agreement with the β-strand shifts.

Table 2. Comparison of TALOS-Derived Dihedral Angles for LM3-Associated HFP and Distributions of Dihedral Angles in β Strands in Protein Structures

	TALOS angles[a]		parallel β strand[b]		antiparallel β strand[b]	
residue	φ	ψ	φ	ψ	φ	ψ
Val-2	−130(9)	143(15)	−118(13)	128(12)	−121(14)	133(15)
Gly-3	−146(16)	152(31)				
Ile-4	−135(14)	154(16)	−115(13)	126(12)	−119(14)	131(14)
Gly-5	−149(13)	159(21)				
Ala-6	−136(14)	150(15)	−122(22)	137(17)	−130(21)	144(15)
Leu-7	−133(16)	143(11)	−112(16)	125(13)	−115(16)	132(14)
Phe-8	−127(15)	143(14)	−114(19)	129(17)	−124(19)	141(17)
Leu-9	−142(11)	147(18)	−112(16)	125(13)	−115(16)	132(14)
Gly-10	−150(12)	160(21)				
Phe-11	−137(10)	152(13)	−114(19)	129(17)	−124(19)	141(17)
Leu-12	−143(10)	148(18)	−112(16)	125(13)	−115(16)	132(14)
Gly-13	−149(13)	159(21)				
Ala-14	−141(9)	149(13)	−122(22)	137(17)	−130(21)	144(15)
Ala-15	−145(8)	149(12)	−122(22)	137(17)	−130(21)	144(15)
Gly-16	−130(26)	154(16)				
Ala-21	−76(9)	146(9)	−122(22)	137(17)	−130(21)	144(15)

[a] Best-fit angles and uncertainties in parentheses are reported in degrees and were determined using the ≥5 best-fit matches from the TALOS database. [b] Distributions with average angle and standard deviations in parentheses were determined from 1042 X-ray structures in the Protein Data Bank with resolutions ≤2.0 Å.[92]

Ala-1 and Gly-16, there were much smaller differences with β-strand shifts and particularly good agreement with β-strand ^{13}Cα shifts. Table 2 lists the dihedral angles derived from a TALOS program comparison between the experimental shifts and a large database of shifts of proteins of known structure.[77] For the residues between Val-2 and Gly-16, the angles were consistent with β-strand conformation. Table 2 also includes database distributions of angles for non-Gly residues in parallel and antiparallel β strands obtained from analysis of a large number of high-resolution protein structures. There is better overall agreement between the HFP angles and the antiparallel distributions. This analysis does not rule out the parallel structure but is interesting in the context of the subsequent experiments to determine the antiparallel registry.

The ^{13}C shifts of Ala-21 were different from the shifts of other HFP alanines such as Ala-6. For example, the Ala-21 Cβ and CO shifts were, respectively, 4.0 and 2.5 ppm higher than those of Ala-6. The TALOS-derived $\varphi = -76°$ for Ala-21 was

(73) Evans, J. N. S. *Biomolecular NMR Spectroscopy*; Oxford Press: New York, 1995.
(74) Zhang, H. Y.; Neal, S.; Wishart, D. S. *J. Biomol. NMR* **2003**, *25*, 173–195.
(75) Kricheldorf, H. R.; Muller, D. *Macromolecules* **1983**, *16*, 615–623.
(76) Igumenova, T. I.; McDermott, A. E.; Zilm, K. W.; Martin, R. W.; Paulson, E. K.; Wand, A. J. *J. Am. Chem. Soc.* **2004**, *126*, 6720–6727.
(77) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289–302.

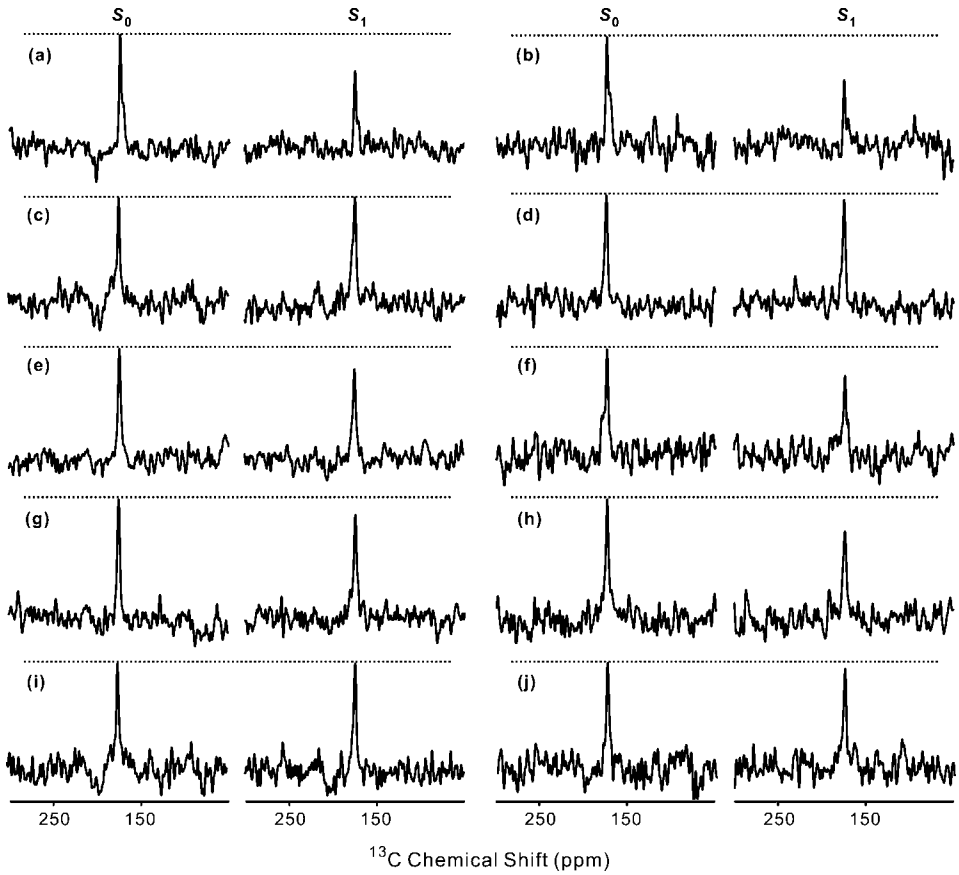

Figure 4. REDOR S_0 and S_1 spectra for membrane-associated (a, b) HFP-G, (c, d) HFP-H, (e, f) HFP-I, (g, h) HFP-J and (i, j) HFP-K. Spectra a, c, e, g, i were obtained with 24 ms dephasing time and spectra b, d, f, h, j were obtained with 32 ms dephasing time. Each spectrum was processed with 200 Hz Gaussian line broadening and baseline correction. Each S_0 or S_1 spectrum was the sum of (a) 41328, (b) 56448, (c) 45920, (d) 81460, (e) 55936, (f) 79744, (g) 30898, (h) 81856, (i) 45920 or (j) 71040 scans.

~70° lower than the values of φ derived for Val-2 through Gly-16 and the Ala-21 linewidths were also broader than those of the more N-terminal residues.[48] All these data suggested that the conformational distribution of Ala-21 was different and less well-defined than the conformations of the more N-terminal residues. These conclusions were consistent with a HFP model composed of (1) an apolar N-terminal HFP region (approximately Ala-1 to Gly-16) which has regular secondary structure because it is predominantly located in the membrane interior and must form intra or interpeptide hydrogen bonds in this low water concentration environment; and (2) a polar C-terminal HFP region (approximately Ser-17 to Ser-23) which is more disordered because it is located at the membrane/water interface or in water and can adopt irregular secondary structures with hydrogen-bonding to water.[48,54]

For τ = 500 or 1000 ms, inter-residue crosspeaks were observed between I4 and G13 of HFP-C (Figure 2c), A6 and G10 of HFP-D (Figure 2d), and A6 and F11 of HFP-F (Supporting Information Figure S1). Inter-residue crosspeaks were not detected in comparable spectra of HFP-A, HFP-B, and HFP-E. For a regular β-sheet conformation, the closest intrapeptide/inter-residue A6/G10 $^{13}C-^{13}C$ distance is ~11 Å and the A6/F11 and I4/G13 distances are even longer. The experimental crosspeaks were likely due to short interpeptide $^{13}C-^{13}C$ distances and would be consistent with a significant population of antiparallel peptides with adjacent strand crossing near F8 and L9. For these registries, there would be A6/G10, A6/F11, and I4/G13 $^{13}C-^{13}C$ distances of <5 Å, as judged by measurement of distances between adjacent antiparallel β strands of proteins with high-resolution structures.

REDOR Spectra and Data Analysis. The TALOS-derived dihedral angles and inter-residue PDSD crosspeaks were consistent with a population of antiparallel HFP with adjacent strand crossing near Phe-8 and Leu-9. Accurate quantitation of tertiary structure $^{13}C-^{13}C$ dipolar couplings and long-range distances from PDSD crosspeak intensities is difficult, and particularly so in the case of $U-^{13}C$ labeled residues for which there are also strong intraresidue $^{13}C-^{13}C$ dipolar couplings. The PDSD results did provide the basis for specific labeling of HFP-G, HFP-H, HFP-I, HFP-J, and HFP-K and application of more quantitative REDOR methods of $^{13}C-^{15}N$ distance determination. Experiments were first carried out on HFP-G that contained ^{13}CO labeling at Ala-14, Ala-15, and Gly-16 and ^{15}N labeling at Ala-1, Val-2, and Gly-3. This labeling scheme was chosen because (1) if there were adjacent strand crossing between Phe-8 and Leu-9, the ^{13}CO-labeled residues on one strand would be hydrogen bonded to the ^{15}N labeled residues on an adjacent strand with concomitant $^{13}C-^{15}N$ dipolar couplings of ~45 Hz; and (2) intramolecular $^{13}CO-^{15}N$ couplings are negligible. Representative spectra are displayed in Figure 4a,b and the respective $(\Delta S/S_0)^{exp}$ were ~0.41 and ~0.50 for τ = 24 and 32 ms. These values suggested that there was a large population of HFP with the putative antiparallel strand registry. Unambiguous analysis of these data was challenging because there were contributions of three distinct

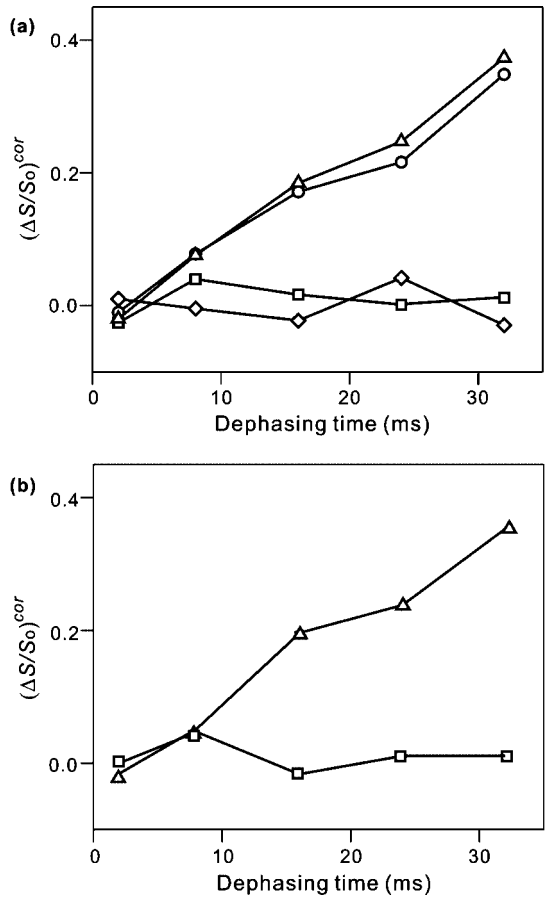

Figure 5. Plots of $(\Delta S/S_0)^{cor}$ vs dephasing time for membrane-associated HFP samples prepared with $[HFP]_{initial}$ of (a) 400 or (b) 25 μM. The symbol legend is diamonds, HFP-H; triangles, HFP-I; circles, HFP-J; and squares, HFP-K. The σ^{cor} were ~0.04.

^{13}COs and because different combinations of antiparallel strand registries could fit the data.

Ambiguity was reduced by studying samples for which HFP had only a single ^{13}CO and a single ^{15}N label. Four HFPs were prepared and all had a ^{13}CO label at Ala-14. This residue had been ^{13}CO labeled in HFP-G and had been previously observed to give a fairly sharp signal.[54] The ^{15}N label was at either Val-2 (HFP-H), Gly-3 (HFP-I), Ile-4 (HFP-J), or Gly-5 (HFP-K). The variation of the REDOR data among the different HFPs was striking (cf. Figure 4c−j and Figure 5a). For $\tau = 32$ ms, the $(\Delta S/S_0)^{exp}$ were ~0.3 for the HFP-I and HFP-J samples and ~0 for the HFP-H and HFP-K samples. These data suggested that there were two antiparallel registries which could be classified: (1) Ala-14 on one strand opposite Gly-3 on the adjacent strand; and (2) Ala-14 on one strand opposite Ile-4 on the adjacent strand. These two registries were denoted A and B and are displayed in Figure 6a. The σ^{cor} in Figure 5 were ~0.04 and explicit error bars are not displayed for visual clarity. Figure S5 in the Supporting Information displays the same plot of $(\Delta S/S_0)^{cor}$ versus dephasing time with explicit error bars as well as the corresponding plot of $(\Delta S/S_0)^{exp}$ versus dephasing time. For $(\Delta S/S_0)^{exp}$ appreciably greater than 0, the typical value of $(\Delta S/S_0)^{cor}/(\Delta S/S_0)^{exp}$ was ~1.2.

The samples used to obtain data for Figures 4 and 5a were made with $[HFP]_{initial} \approx 400$ μM. To check for possible effects of HFP self-association in aqueous solution prior to membrane binding, two additional HFP-I and HFP-K samples were made with $[HFP]_{initial} \approx 25$ μM which is a concentration for which HFP is known to be monomeric in the HEPES buffer.[59] Figure 5a,b illustrates that very similar $(\Delta S/S_0)^{cor}$ were obtained for both values of $[HFP]_{initial}$ and the apparent strand registries appear to be due to membrane-association. The similar values also support the reproducibility of the large differences in $(\Delta S/S_0)^{cor}$ as a function of the ^{15}N labeling site.

More quantitative analysis of the $(\Delta S/S_0)^{cor}$ of the samples was done using calculations of $(\Delta S/S_0)^{sim}$ based on different models for registries of three adjacent strands with the overall goal of quantitation of the populations of the different registries. The strands were denoted, "top", "middle", and "bottom". Figure 6b displays the models as well as spin geometries specific to the HFP-I sample. The models were focused on registries at the middle strand Ala-14 whose ^{13}CO group was hydrogen-bonded to an amide proton in the top strand. Each model was labeled by two letters which were either A, B, or X. The first letter described the registry relating the middle strand and the top strand and the second letter described the registry relating the middle strand and the bottom strand. For registry A, Ala-14 in the middle strand was across from Gly-3 in the adjacent strand and for registry B, Ala-14 in the middle strand was across from Ile-4 in the adjacent strand (cf. Figure 6a). Registry X was defined as any structure for which the interpeptide ^{13}CO−^{15}N distance was large in the HFP-H, HFP-I, HFP-J, and HFP-K samples so that $d \approx 0$. Registry X could include the in-register parallel strand arrangement. Such registry has been proposed for membrane-associated gp41 constructs which contain the HFP.[78] Consideration of the two strands adjacent to the central strand is based on the labeled ^{15}N which would be close to the Ala-14 ^{13}CO and is not meant to imply that HFP forms trimers.

Model XX had $(\Delta S/S_0)^{sim} = 0$ for all dephasing times while models AX, XA, BX, and XB resulted in two-spin systems for which $(\Delta S/S_0)^{sim}$ were primarily dependent on the ^{13}CO−^{15}N distance. Models AA, BA, AB, and BB were three-spin systems for which $(\Delta S/S_0)^{sim}$ depended both on the two ^{13}CO−^{15}N distances and on the angle between the two ^{13}CO−^{15}N vectors.[79] For all samples and all models, $(\Delta S/S_0)^{sim}$ were calculated for each of the five experimental dephasing times.

The fractional populations of each of the models were calculated with fitting of the $(\Delta S/S_0)^{sim}$ and the $(\Delta S/S_0)^{cor}$. The fitting was primarily based on the data from the HFP-I and HFP-J samples because many of the $(\Delta S/S_0)^{cor}$ for these samples were appreciably positive. Fitting was accomplished with the following equations:

$$\chi^2 = \sum_{j=1}^{2}\sum_{k=1}^{5} \frac{\left\{(\Delta S/S_0)_{j,k}^{calcd} - (\Delta S/S_0)_{j,k}^{cor}\right\}^2}{(\sigma_{j,k}^{cor})^2} \quad (2)$$

$$(\Delta S/S_0)_{j,k}^{calcd} = \sum_{l=1}^{9} f_l \times (\Delta S/S_0)_{j,k}^{sim} \quad (3)$$

for which j was the index of the sample, k was the index of the dephasing time, l was the index of the model, and f_l was the fractional population of model l.

Three types of fitting were done and differed in the choice of which f_l were fitted and which were set to zero. For all fittings, $\Sigma f_l = 1$. For "unconstrained" fitting, there was no correlation between the registry of the middle and top strands and the

(78) Sackett, K.; Shai, Y. *J. Mol. Biol.* **2005**, *350*, 790–805.
(79) Vogt, F. G.; Gibson, J. M.; Mattingly, S. M.; Mueller, K. T. *J. Phys. Chem. B* **2003**, *107*, 1272–1283.

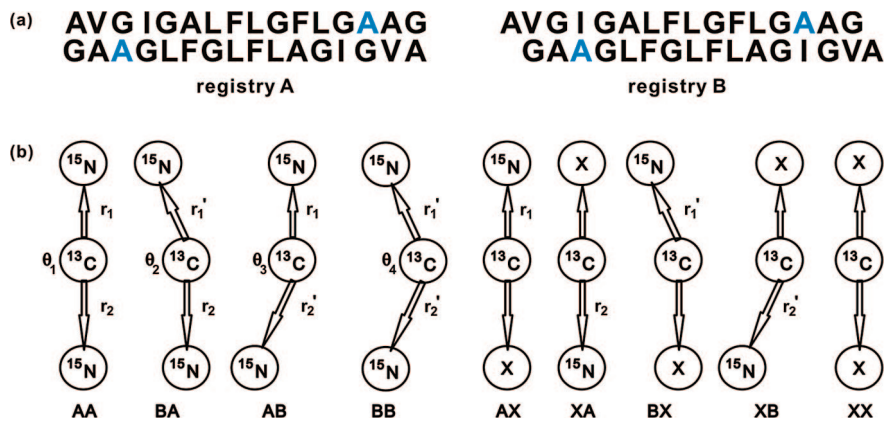

Figure 6. (a) Two antiparallel registries of residues 1–16 of HFP that were consistent with the REDOR data shown in Figure 5. The registries are denoted A and B and the ^{13}CO labeled Ala-14 residue is highlighted in blue. (b) Models used to calculate $(\Delta S/S_0)^{sim}$ and spin geometries specific for the HFP-I sample. Each model includes nuclei from three adjacent strands with the Ala-14 ^{13}CO always in the middle strand and ^{15}N in the top and/or bottom strands. Consideration of the two strands adjacent to the central strand is based on the labeled ^{15}N which would be close to the Ala-14 ^{13}CO and is not meant to imply that HFP forms trimers. The first letter in the labeling of each model refers to the middle strand/top strand registry and the second letter refers to the middle strand/bottom strand registry. Registry X is any registry for which the interpeptide ^{13}CO−^{15}N distance was large in the HFP-H, HFP-I, HFP-J, or HFP-K samples so that $d \approx 0$. The Ala-14 ^{13}CO is hydrogen bonded to an amide proton in the top strand. Relevant labeled ^{13}C−^{15}N distances and ^{15}N−^{13}C−^{15}N angles are $r_1 = 4.063$ Å; $r_1' = 5.890$ Å; $r_2 = 5.455$ Å; $r_2' = 6.431$ Å; $\theta_1 = 161.1°$; $\theta_2 = 131.9°$; $\theta_3 = 130.2°$; and $\theta_4 = 117.0°$. Each parameter value was the average of 10 specific values taken from the crystal structure of outer membrane protein G.

registry of the middle and bottom strands. All f were therefore fitted and each f was a function of "a" and "b" which were defined as the fractional probabilities of two adjacent strands having A or B registries, respectively. The fractional probability of the X registries was then $1 - a - b$. Each f was the product of the fractional probabilities of the middle strand/top strand registries and middle strand/bottom strand registries with resulting $f_{AA} = a^2$, $f_{BA} = ab$, $f_{AB} = ab$, $f_{BB} = b^2$, $f_{AX} = a(1 - a - b)$, $f_{XA} = a(1 - a - b)$, $f_{BX} = b(1 - a - b)$, $f_{XB} = b(1 - a - b)$, and $f_{XX} = (1 - a - b)^2$. "Partially constrained" fitting was done based on the idea that there were domains of antiparallel strand registry and domains of X registry so that $f_{AA} = a^2$, $f_{BA} = ab$, $f_{AB} = ab$, $f_{BB} = b^2$, $f_{AX} = 0$, $f_{XA} = 0$, $f_{BX} = 0$, $f_{XB} = 0$, and $f_{XX} = 1 - (a + b)^2$. For partially constrained fitting, physically meaningful expressions of a and b included (1) a/b which was the ratio of probability that two adjacent strands had A registry to the probability that they had B registry; and (2) $(a + b)^2/[(1 - (a + b)]^2$ which was the ratio of the total population of the A and B antiparallel structures to the population of the X structures. For "fully constrained" fitting, it was assumed that β strand domains would form with only A or only B or only X registries so that $f_{AA} = a^2$, $f_{BA} = 0$, $f_{AB} = 0$, $f_{BB} = b^2$, $f_{AX} = 0$, $f_{XA} = 0$, $f_{BX} = 0$, $f_{XB} = 0$, and $f_{XX} = 1 - a^2 - b^2$. In this fitting, the fractional populations of the A, B, and X strand arrangements were a^2, b^2, and $1 - a^2 - b^2$, respectively.

The results of unconstrained fitting are displayed in Figure 7a as a 2D contour plot of χ^2 versus a and b. The best-fit $a = 0.22$ and $b = 0.31$ with $\chi^2_{min} = 16.5$ and good-fit a and b represented in the black region.[67,69] The good-fit regions of the plot showed negative correlation between a and b as might be expected from the positive correlation between $(\Delta S/S_0)^{calcd}$ and either a or b for both the HFP-I and HFP-J samples. The $(\Delta S/S_0)^{calcd}$ were also computed for the HFP-H and HFP-K samples using the best-fit a and b. At $\tau = 32$ ms, maximum $(\Delta S/S_0)^{calcd}$ of 0.08 and 0.09 were obtained for the HFP-H and HFP-K samples, respectively, and can be compared to the maximum $(\Delta S/S_0)^{cor} = 0.05 \pm 0.04$ for these samples. Figure 7b displays the 2D contour plot of partially constrained fitting with best-fit $a = 0.31$, $b = 0.42$, and $\chi^2_{min} = 15.1$. At $\tau = 32$ ms, these a and b values led to $(\Delta S/S_0)^{calcd} = 0.11$ and 0.13 for the HFP-H and HFP-K samples, respectively. Figure 7c displays the 2D contour plot of fully constrained fitting with best-fit $a^2 = 0.26$, $b^2 = 0.33$, and $\chi^2_{min} = 12.7$ and $(\Delta S/S_0)^{calcd} = 0.09$ and 0.12 at $\tau = 32$ ms for the HFP-H and HFP-K samples, respectively. For all three fittings, the χ^2_{min} are reasonable, as evidenced by being within a factor of 2 of 8, the number of degrees of freedom of the fitting. This suggests that each model is plausible. The limits of the good-fit black regions have been generously set and include all parameter space with χ^2 2–3 units higher than χ^2_{min}.

The best-fit f of the three fittings were used to calculate P_A, P_B, and P_X which were fractional populations of the A, B, and X registries, respectively: $P_A = f_{AA} + (f_{BA} + f_{AB} + f_{AX} + f_{XA})/2$; $P_B = f_{BB} + (f_{BA} + f_{AB} + f_{BX} + f_{XB})/2$; and $P_X = f_{XX} + (f_{AX} + f_{XA} + f_{BX} + f_{XB})/2$ with $P_A + P_B + P_X = 1$. The resulting fractional populations were (1) unconstrained fitting, $P_A = 0.22$, $P_B = 0.31$, and $P_X = 0.47$; (2) partially constrained fitting, $P_A = 0.23$, $P_B = 0.31$, and $P_X = 0.46$; and (3) fully constrained fitting, $P_A = 0.26$, $P_B = 0.33$, and $P_X = 0.41$. An overall result of the three fittings was therefore $P_A \approx 0.25$, $P_B \approx 0.30$, and $P_X \approx 0.45$. In addition, examination of the values in the black regions of the three plots showed that the approximate range of reasonable values for the sum $P_A + P_B$ was 0.5−0.6 and that the corresponding range for P_X was 0.4−0.5.

The determination of P_X relied on quantitative determination of $(\Delta S/S_0)^{cor}$. Although some REDOR studies in the literature show smaller $(\Delta S/S_0)^{cor}$ than would be predicted by simulation, we think that our $(\Delta S/S_0)^{cor}$ are quantitative based on the results of an earlier study by our group.[20,52,80] In this study, REDOR data were analyzed for a membrane-associated HIV fusion peptide with a ^{13}CO label at Leu-7 and a ^{15}N label at Phe-11. Unlike the membranes used in the present study, the model membranes in the earlier study did not contain cholesterol and the Leu-7 ^{13}CO chemical shift was consistent with helical rather than strand conformation. It was also shown that the ^{13}CO/^{15}N REDOR data could be fitted well to a 4.1 ± 0.1 Å ^{13}CO···^{15}N

(80) Sharpe, S.; Kessler, N.; Anglister, J. A.; Yau, W. M.; Tycko, R. *J. Am. Chem. Soc.* **2004**, *126*, 4979–4990.

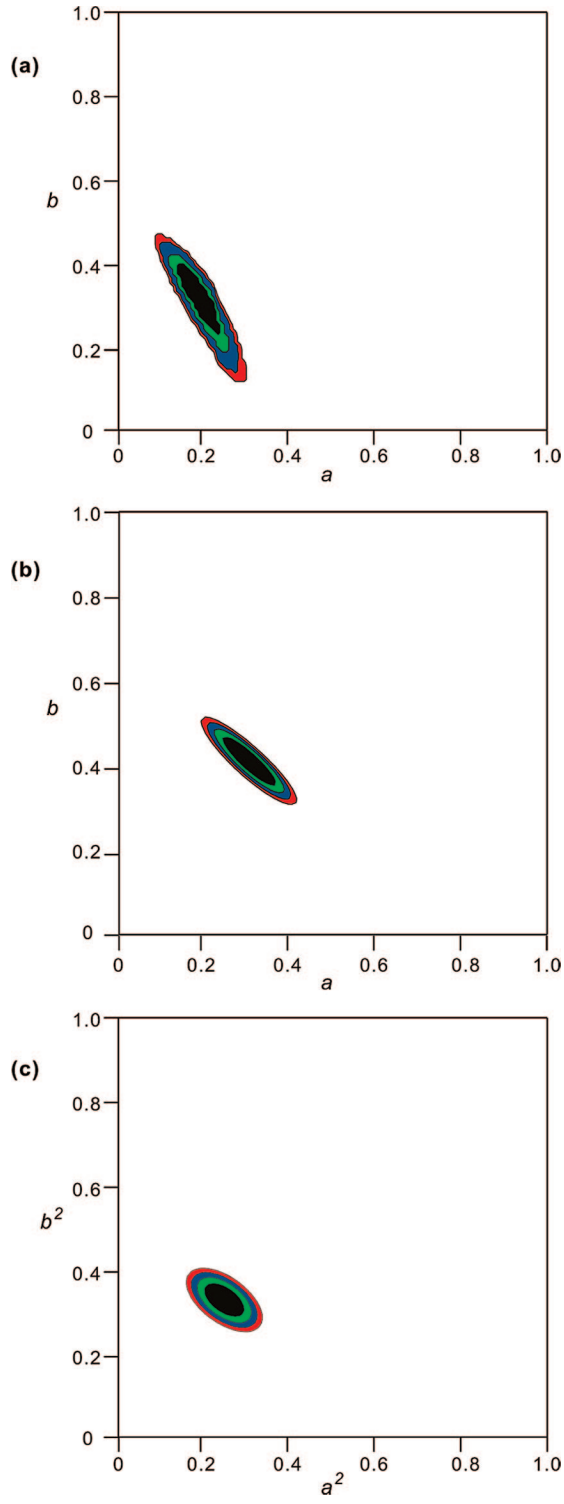

Figure 7. Contour plots of χ^2 vs strand fitting parameters for (a) unconstrained; (b) partially constrained; and (c) fully constrained fittings. The a, b, a^2, and b^2 parameters refer to probabilities for different adjacent strand arrangements. In plot a, the black, green, blue, red, and white regions respectively correspond to $\chi^2 < 19$, $19 < \chi^2 < 21$, $21 < \chi^2 < 23$, $23 < \chi^2 < 25$, and $\chi 2 < 25$. In plot b, the regions respectively correspond to $\chi^2 < 18$, $18 < \chi^2 < 20$, $20 < \chi^2 < 22$, $22 < \chi^2 < 24$, and $\chi 2 < 24$, and in plot c, the regions respectively correspond to $\chi^2 < 15$, $15 < \chi^2 < 17$, $17 < \chi^2 < 19$, $19 < \chi^2 < 21$, and $\chi^2 < 21$. Best-fit parameters were: (plot a) $a = 0.22$, $b = 0.31$, $\chi^2 = 16.5$; (plot b) $a = 0.31$, $b = 0.42$, $\chi^2 = 15.1$; and (plot c) $a^2 = 0.26$, $b^2 = 0.33$, $\chi^2 = 12.7$. In plot a, the a and b parameters are the fractional probabilities of adjacent strands having A or B registries, respectively. In plot c, the a^2 and b^2 parameters are the fractional probabilities of domains of A or B registries, respectively.

distance which is the expected distance between the Leu-7 and Phe-11 nuclei in a regular α helix. A natural abundance correction factor very similar to the one in the present paper was applied prior to fitting and the corrected data had $(\Delta S/S_0)^{cor} = 1.0$ at $\tau = 32$ ms which is the value predicted by simulation. We expect that there were similar ^{13}C T_2s in the earlier and the present studies because both samples were membrane-associated HIV fusion peptides at the same temperature. The correction factors and the $^{13}CO-^{15}N$ dipolar couplings were also very similar and we therefore think that the $(\Delta S/S_0)^{cor}$ in the present study are quantitative.

4. Discussion

Membrane-associated HFP is known to adopt either helical or β-strand conformation and membrane composition is one factor which impacts conformation. The goal of the current study was to develop a more detailed structural model for the β-strand form of the HFP and relied on previous studies which showed that this was the dominant conformation in membranes which contained a significant amount of cholesterol.[20,48,51,52,78,81] The β strand conformation may be a physiologically relevant HFP structure because membranes of host cells of HIV contain ~30 mol % cholesterol and because HFP fuses vesicles whose membranes contain cholesterol.[18,21,48]

The first aim of our study was to determine which HFP residues adopted β strand conformation and which residues adopted non-β-strand conformation. This aim was accomplished with analysis of $^{13}C-^{13}C$ correlation spectra of SUL samples and resulted in an unambiguous ^{13}C assignment for the Ala-1 to Gly-16 and the Ala-21 residues. The $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO shifts of Ala-1 to Gly-16 were more consistent with β-strand conformation than with helical conformation, and the good-fit φ, ψ dihedral angles derived from TALOS analysis of these shifts were closer to the centers of the distributions of angles of antiparallel β strands than to the centers of the distributions of parallel β strands. The Ala-21 ^{13}C shifts were less clearly β strand and the linewidths were broader than those of other residues.[48] The overall results of the chemical shift analysis were (1) continuous β strand over the Ala-1 to Gly-16 residues and (2) greater disorder at Ala-21. Infrared structural investigations of membrane-associated HFP have generally been consistent with predominant antiparallel β-sheet conformation and were based on analysis of the wavenumbers of the amide I transition.[17,51,81,82] One infrared study proposed that there was β-hairpin structure in the Ala-1 to Gly-16 region but the present work did not support this model because there were no residues in this region with non-β-strand ^{13}C shifts.[51] To our knowledge, the complete ^{13}C shift assignment of the present study is the first definitive evidence for a fully extended conformation.

One rationale for the conformational results of our study is based on HFP membrane location. The first sixteen residues of HFP are all apolar and could be predominantly located in the membrane interior. Because the membrane interior has a small dielectric constant and low water content, regular peptide conformation which maximizes intra or interpeptide hydrogen bonding would be favored. The more C-terminal HFP residues are more polar and might be located near the lipid headgroups or in aqueous solution. More disordered

(81) Castano, S.; Desbat, B. *Biochim. Biophys. Acta.* **2005**, *1715*, 81–95.
(82) Gordon, L. M.; Mobley, P. W.; Pilpa, R.; Sherman, M. A.; Waring, A. J. *Biochim. Biophys. Acta* **2002**, *1559*, 96–120.

conformation would be possible because the peptide CO and NH could hydrogen bond to water. This HFP membrane location model has been generally supported by a solid-state NMR study which showed that there were 5–6 Å distances between lipid ^{31}Ps and ^{13}COs of the Ala-14 to Gly-16 residues while the corresponding distances for the Gly-5 to Gly-13 residues were >8 Å.[54]

The second aim of this study was to determine the β-strand registry. The first set of experiments was detection of interpeptide/inter-residue crosspeaks in $^{13}C-^{13}C$ correlation spectra of SUL samples with long mixing times. Crosspeaks between Ala-6 and Gly-10 and between Ile-4 and Gly-13 were consistent with antiparallel strands with the A or B registries. These results demonstrated how experiments using SUL samples could aid development of tertiary structure models. More specific labeling for REDOR experiments probed adjacent strand $^{13}CO-^{15}N$ distances in registries consistent with the SUL data. REDOR data for HFPs with a ^{13}CO label at Ala-14 and a ^{15}N label at either Val-2, Gly-3, Ile-4, or Gly-5 were qualitatively clear with large $(\Delta S/S_0)^{cor}$ observed with a Gly-3 or Ile-4 ^{15}N label and $(\Delta S/S_0)^{cor} \approx 0$ with a Val-2 or Gly-5 ^{15}N label. One appealing aspect of these registries was that they would result in complete or nearly complete interpeptide hydrogen bonding of the apolar Ala-1 to Gly-16 apolar region of HFP as might be favored in the membrane interior.

The REDOR data were more quantitatively analyzed to yield ~25% population of antiparallel adjacent strands with Ala-14/Gly-3 registry A, ~30% antiparallel population with Ala-14/Ile-4 registry B, and ~45% population with a structure which was not Ala-14/Val-2, Ala-14/Gly-3, Ala-14/Ile-4, or Ala-14/Gly-5 registry (registry X). This result is significant because to our knowledge, it provides the first residue-specific structural model for β-strand HFP. As highlighted earlier, there would be complete (registry A) or nearly complete (registry B) interpeptide hydrogen bonding for residues Ala-1 to Gly-16 which form the apolar region of the HFP. These hydrogen bonding patterns would be favored if this region were predominantly located in the membrane interior. The existence of multiple β-strand structures is also consistent with a recent ^{13}C and ^{15}N assignment of a membrane-associated HFP with SUL at Phe-8, Leu-9, and Gly-10.[44] There were two crosspeaks of comparable intensity for the Leu-9 ^{13}CO/Gly-10 ^{15}N correlation and two crosspeaks of comparable intensity for the Gly-10 $^{13}C\alpha$/Gly-10 ^{15}N correlation. For a given pair, the two ^{13}C shifts differed by ~0.5 ppm and were both consistent with β-strand conformation whereas the Gly-10 ^{15}N shifts were 107 and 111 ppm. The two crosspeaks may correlate with the multiple β-strand structures inferred from analysis of the REDOR data in the present paper.

It is interesting to compare the antiparallel registries detected in the present study with the antiparallel registry suggested by a previous REDOR study.[52] In this study, the samples contained an equimolar mixture of a HFP with three sequential ^{13}CO labels and an HFP with three sequential ^{15}N labels. Data were only acquired for a single dephasing time ($\tau = 24$ ms) and the best-guess antiparallel registry had Ala-14 hydrogen bonded with Leu-7 which is different than the registries consistent with the data of the present study. The Ala-14/Leu-7 registry could be one of the X structures but it is noted that there was significant uncertainty in the determination of this registry because of the multiple ^{13}CO and ^{15}N labels. Because of the single site ^{13}CO and ^{15}N labeling in the present paper, there was definitive determination of the A and B registries and as discussed earlier, these registries are biophysically reasonable.

The REDOR results confirmed that the earlier $^{13}C-^{13}C$ correlation experiments with long mixing times on the SUL samples had provided accurate information about possible registries of antiparallel β strands. For these earlier experiments, it was significant that (1) there were no sequential labeled residues so that inter-residue/intrapeptide crosspeak intensities were attenuated; and (2) crosspeaks were observed between a few but not most residues and the observed crosspeaks were consistent with a small number of strand registries. A time-saving advantage of the approach was that the same SUL samples could be used for ^{13}C assignment as well as for semiquantitative $^{13}C-^{13}C$ distance determination. This type of SUL is an alternative to $U-^{13}C$ labeling for assignment and alternate site ^{13}C labeling for distance determination.[31,42,45–47] The SUL and assignment and distance determination method should be useful for concentration-limited systems such as membrane-associated peptides for which (1) reasonable signal-to-noise can be obtained with 2D but not 3D spectra and (2) intrinsic ^{13}C linewidths are 2–3 ppm so that assignment is ambiguous with $U-^{13}C$ labeling. The method is restricted to peptides and proteins which can be chemically synthesized.

It is interesting to consider models A and B in the context of the full gp41 protein. The gp41 soluble ectodomain structures to-date show a symmetric trimer with an in-register parallel coiled-coil extending over residues 30–80.[11] The residues N-terminal of Ala-30 are disordered and the soluble ectodomain constructs also lacked the N-terminal HFP. Although there is no evidence that the oligomeric state of the membrane-associated HFP of the present study is a trimer, it is interesting to consider the antiparallel β-sheet structure of HFP in the context of the putative trimeric state of intact gp41. It is difficult to understand this structure in the context of a single gp41 trimer, but this structure could be understood considering two trimers denoted "C" and "D" with respective HFP strands C_1, C_2, and C_3 and D_1, D_2, and D_3. A $C_1D_3C_2D_2C_3D_1$ antiparallel β-sheet structure could be formed with the C_1, C_2, and C_3 strands parallel to one another, the D_1, D_2, and D_3 strands parallel to one another, and the C and D strands antiparallel to one another with D_3 hydrogen bonded to C_1 and C_2, C_2 hydrogen bonded to D_3 and D_2, etc. There is some support for this model from internuclear distance measurements on a HFP trimer construct composed of three HFP strands chemically cross-linked at their C-termini. The $^{13}C-^{13}C$ and $^{13}C-^{15}N$ distances determined for this membrane-associated trimer were consistent with the A antiparallel registry deduced from the present study.[20]

Detection of multiple registries for a membrane-associated peptide is to our knowledge rare. Peptides which form amyloid fibrils can adopt different antiparallel registries at different pHs but we are not aware of a case for which two registries were formed in a single amyloid sample.[83] For a membrane-inserted HFP aggregate, one factor favoring the formation of the A registry is interpeptide hydrogen bonding for all of the residues between Ala-1 and Gly-16. This hydrogen bonding would reduce the unfavorable Born energy of CO and NH dipoles in the low dielectric environment of the membrane interior. For the B registry, Ala-1 is not part of the hydrogen bonded β-sheet registry and if the HFP N-terminus is charged, better charge

(83) Petkova, A. T.; Buntkowsky, G.; Dyda, F.; Leapman, R. D.; Yau, W. M.; Tycko, R. *J. Mol. Biol.* **2004**, *335*, 247–260.

solvation might be achieved relative to the A registry. Ala-1 could adopt a broader range of conformations in the B registry which might facilitate the location of the charged N-terminus in a solvated environment. A greater distribution of conformations for Ala-1 is supported by linewidths which were broader than those of residues in the central region of HFP.[48] Although the ionization state(s) of membrane-associated β-strand HFP have not yet been experimentally determined, there is evidence for a charged amino terminus in the related influenza fusion peptide in its helical conformation.[84]

Either HIV or HFP with the Val-2 → Glu-2 point mutation is nonfusogenic.[4,8,9] In the context of our results, this lack of fusion activity may be related to a change in strand registries arising from the charged glutamic acid side chain. The Val-2 → Glu-2 mutation is trans-dominant, that is, mixtures of wild-type and mutant proteins correlated with fusion activities which were reduced much more than would be expected from the fraction of mutant protein. This effect could be explained by registry changes for several strands near the mutant HFP which might affect HFP oligomerization and/or membrane location.[9]

It might be expected that the combination of the A, B, and X β-strand structures would result in broad NMR linewidths. However, the ^{13}C linewidths observed to-date for helical HFP are comparable to those of β-strand HFP.[19] It is also noted that a liquid-state NMR study of detergent-associated HFP indicated multiple helical structures.[85] Overall, the structural picture of membrane-associated HFP is complex with either predominant helical or β-strand conformation and two or more β-strand registries. This structural plasticity may be related to peptide flexibility needed as the lipid molecules move during the fusion process.[86,87]

The data in our study restrict the X structures to structures other than the Ala-14/Val-2, Gly-3, Ile-4, or Gly-5 antiparallel registries. There are therefore many possibilities for the X structures. One reasonable possibility is parallel β-strand structure either in-register or close to in-register. This structural model is appealing because most of the residues in the Ala-1 to Gly-16 region would have interpeptide hydrogen bonds and this region could therefore be located in the membrane interior. Previous solid-state NMR ^{13}C−^{15}N distance measurements were consistent with some population of in-register parallel strand structure over residues Gly-5 to Gly-13 in addition to antiparallel population over residues Gly-5 to Gly-16.[52,88] In addition, infrared studies on constructs containing the first 34 or first 70 residues of gp41 were consistent with predominant in-register parallel β-sheet structure from residues Ala-1 to Gly-16.[78] The interpretation of the infrared data was based on shifts in peak wavenumbers of ^{13}C labeled relative to natural abundance peptides.

The HFPs are monomeric in aqueous solution and β-sheet aggregates form upon association with the membrane.[59] The numbers of molecules in an aggregate have not been directly determined but this number is probably small and probably less than 100 based on the following experimental observations: (1) For ~30% of the molecules in the aggregates, the ^{13}COs of residues between Ala-14 and Gly-16 are <6 Å from the lipid ^{31}P.[54] The membranes also remain in the bilayer phase for HFP/lipid ≤0.1.[89] This close contact between HFP and lipid bilayer headgroups is more reasonable for a smaller aggregate than for a larger aggregate. (2) Relative to frozen samples, spectra of unfrozen samples yield significantly lower ^{13}C cross-polarization signal intensity and narrower ^{13}C linewidths.[53] Both phenomena are consistent with greater motion in the unfrozen samples. A larger dependence of motion on frozen versus unfrozen state is expected for smaller aggregates. Future solid-state NMR experiments might provide more detailed information about membrane-associated aggregate size.[90]

5. Conclusions

An unambiguous ^{13}C assignment was obtained for residues Gly-1 to Gly-16 and residue Ala-21 of the membrane-associated HFP. SUL and 2D ^{13}C−^{13}C correlation spectroscopy were used to obtain this assignment. The ^{13}C shifts and associated TALOS-derived φ, ψ dihedral angles were consistent with fully β-strand conformation for residues Ala-1 to Gly-16. Less definitive β-strand shifts and broader linewidths were observed for Ala-21 and indicated a broader distribution of conformations. Unambiguous assignments and detailed conformational analysis of SUL HFPs with C-terminal cross-linking should also be possible and should provide greater biological significance because the topology of HFP strands in these cross-linked constructs is thought to mimic the HFP topology in the gp41 protein.[57] In addition, the approach should be applicable to membrane-associated HFP which is helical and should provide information about whether HFP forms a continuous helix as was observed in detergent micelles or whether there is a helix-turn-helix motif as has been detected for membrane-associated influenza fusion peptide.[14,91]

^{13}C−^{13}C correlation experiments with long mixing times on the SUL samples provided information about possible registries of antiparallel β strands, and these registry models were tested with REDOR ^{13}CO−^{15}N distance measurements on a few selectively labeled samples. Two of the registries were shown to have significant population and both registries were consistent with complete or nearly complete interpeptide hydrogen bonding for the apolar N-terminal domain of the HFP. This hydrogen bonding scheme would be favored if a significant part of this domain were located in the membrane interior where there is low water content.

The development of a detailed structural model for β-strand HFP is significant because this is the observed conformation in cholesterol-containing membranes which reflect the composition of membranes of host cells of HIV. HFP fusion activity is also observed for vesicles with this membrane composition and the β-strand conformation may therefore be a physiologically relevant HFP structure.

(84) Zhou, Z.; Macosko, J. C.; Hughes, D. W.; Sayer, B. G.; Hawes, J.; Epand, R. M. *Biophys. J.* **2000**, *78*, 2418–2425.
(85) Gabrys, C. M.; Weliky, D. P. *Biochim. Biophys. Acta.* **2007**, *1768*, 3225–3234.
(86) Hofmann, M. W.; Weise, K.; Ollesch, J.; Agrawal, P.; Stalz, H.; Stelzer, W.; Hulsbergen, F.; de Groot, H.; Gerwert, K.; Reed, J.; Langosch, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14776–14781.
(87) Reichert, J.; Grasnick, D.; Afonin, S.; Buerck, J.; Wadhwani, P.; Ulrich, A. S. *Eur. Biophys. J.* **2007**, *36*, 405–413.
(88) Zheng, Z.; Qiang, W.; Weliky, D. P. *Magn. Reson. Chem.* **2007**, *245*, S247–S260.
(89) Yang, J.; Parkanzky, P. D.; Khunte, B. A.; Canlas, C. G.; Yang, R.; Gabrys, C. M.; Weliky, D. P. *J. Mol. Graph. Model.* **2001**, *19*, 129–135.
(90) Luo, W.; Hong, M. *J. Am. Chem. Soc.* **2006**, *128*, 7242–7251.
(91) Han, X.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K. *Nat. Struct. Biol.* **2001**, *8*, 715–720.
(92) Hovmoller, S.; Zhou, T.; Ohlson, T. *Acta Crystallogr. D* **2002**, *58*, 768–776.

Acknowledgment. This work was supported by NIH award AI47153 to D.P.W. The Mass Spectroscopy, Macromolecular Synthesis and Structure, and Max T. Roger NMR facilities at Michigan State University were used to carry out this research. Useful discussions with Dr. Charles Gabrys are also acknowledged.

Supporting Information Available: 2D PDSD spectra of HFP-F at 10 and 500 ms exchange times; 1D slices for the 2D PDSD spectra of HFP-C and HFP-D; REDOR spectra at 24 and 32 ms dephasing times for HFP-I and HFP-K samples prepared with 25 μM initial HFP concentration; derivation of $(\Delta S/S_0)^{cor}$ from $(\Delta S/S_0)^{exp}$ and σ^{cor} from σ^{exp}; comparison of $(\Delta S/S_0)^{exp}$ with $(\Delta S/S_0)^{cor}$ for the HFP-H, I, J, and K samples; model geometries for the HFP-H, HFP-J, and HFP-K samples; and $(\Delta S/S_0)^{sim}$ for all models for the HFP-H, HFP-I, HFP-J, and HFP-K samples. This material is available free of charge via the Internet at http://pubs.acs.org.

JA077302M

Supporting Information of "Solid-State NMR Spectroscopy of HIV 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" by Wei Qiang, Michele L. Bodner, and David P. Weliky

1. PDSD spectra of PC:PG:CHOL-associated HFP-F

Figure S1a

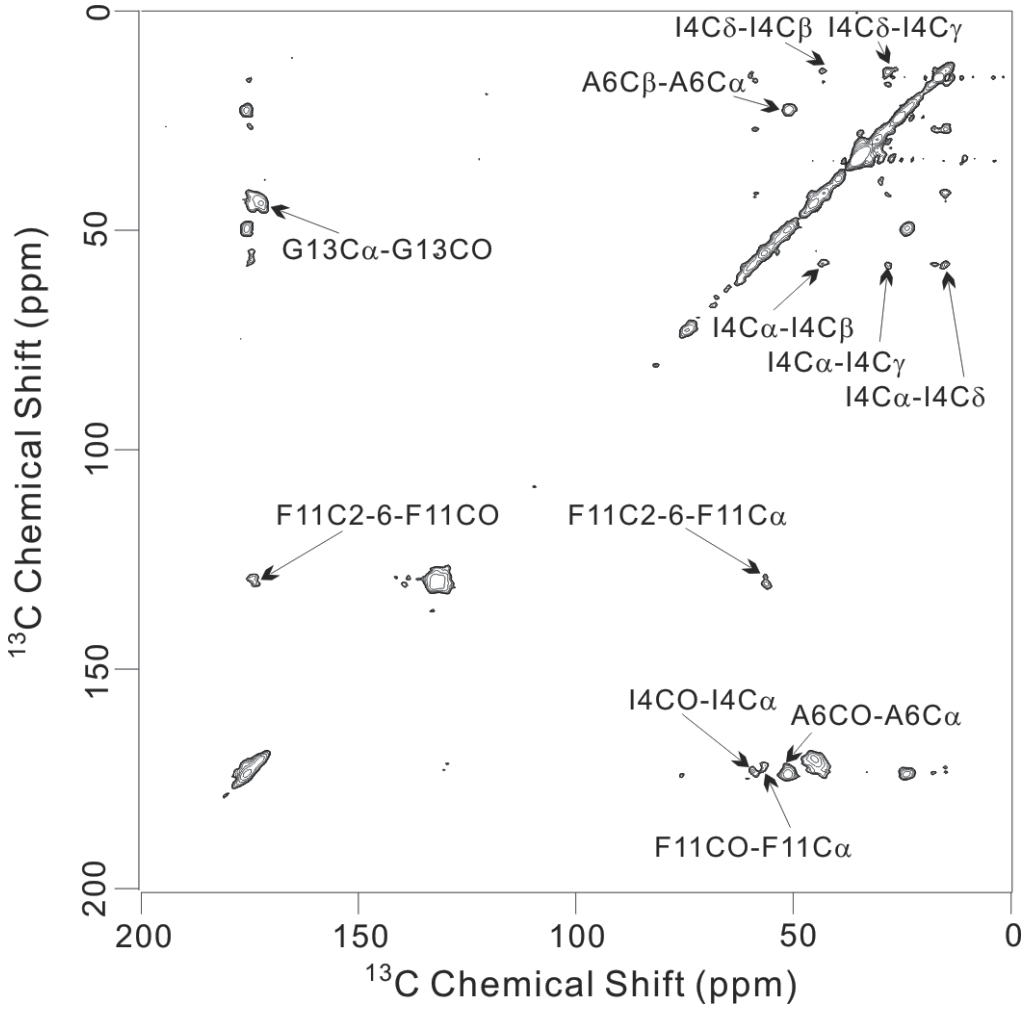

Figure S1b

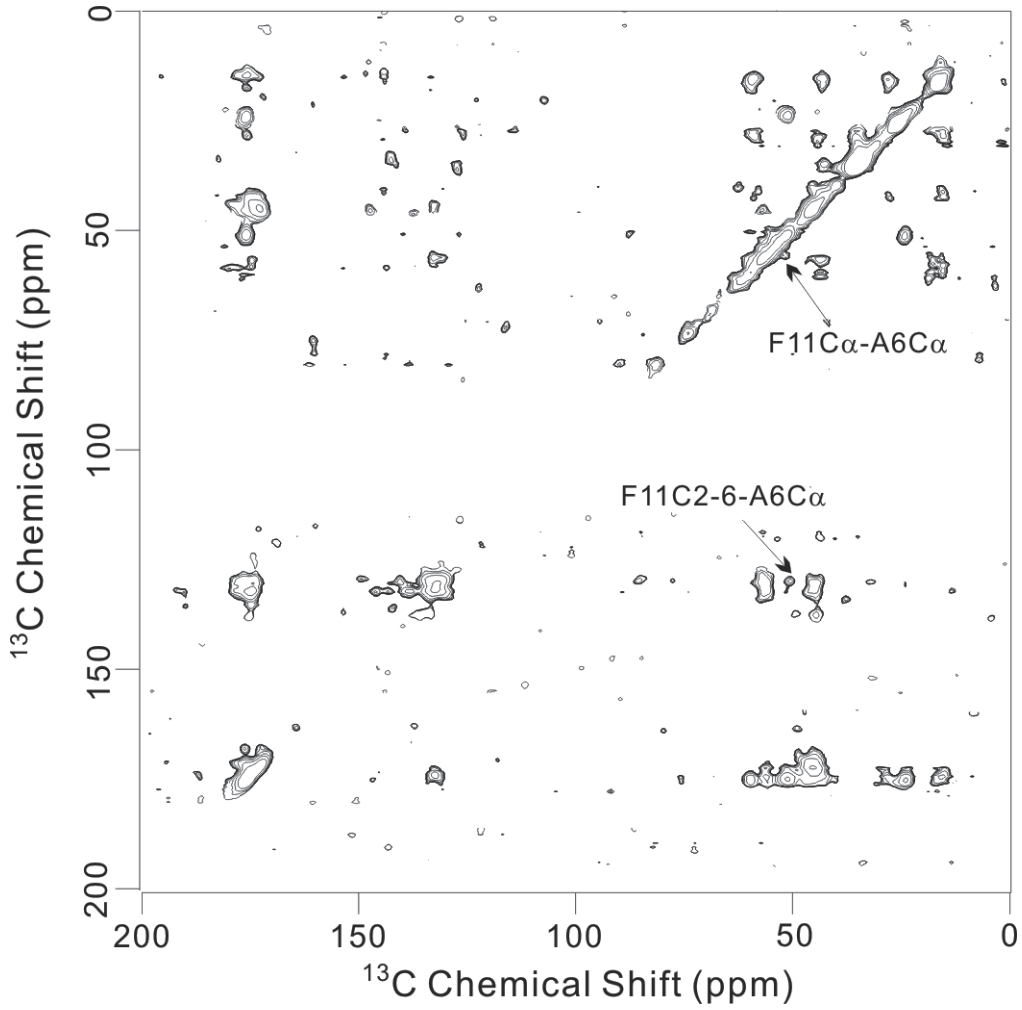

Figure S1. PDSD spectra for PC:PG:CHOL-associated HFP-F at (a) 10 ms and (b) 500 ms exchange times. The spectra were taken at –50 °C with MAS frequency of 10000±2 Hz. Spectra were processed with 100 Hz Gaussian line broadening and polynomial baseline correction in the f_2 (horizontal) and f_1 (vertical) dimensions. The total numbers of scans were (a) 102400 and (b) 409600. Some of the (a) intra-residue and (b) inter-residue peak assignments are listed using the convention of assignment in f_2 – assignment in f_1.

2. 1D-slices of PDSD spectra

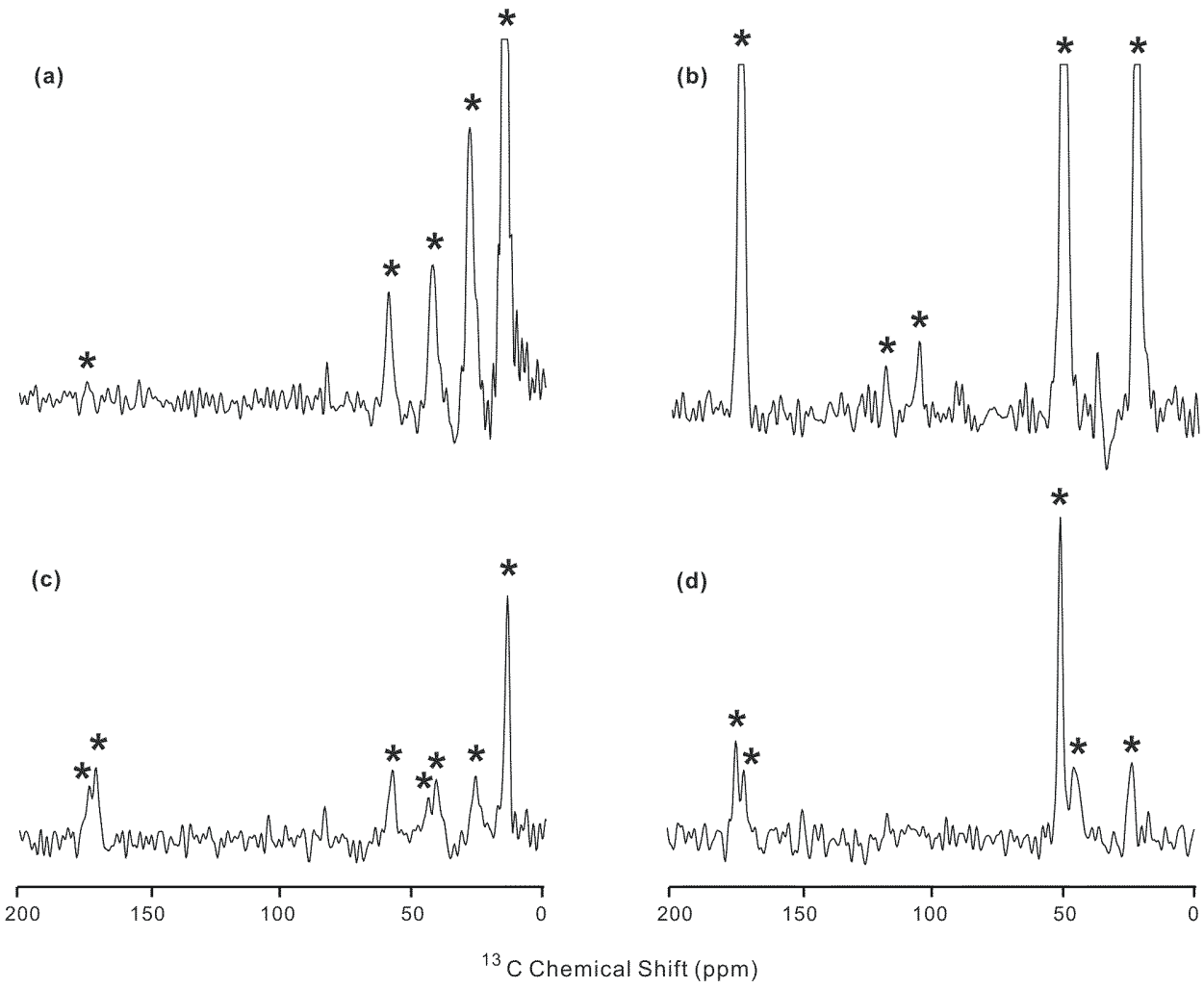

Figure S2. Slices of the 2D PDSD spectra displayed in Fig. 2 in the main text. Slices a-d are along the f_2 dimension and correspond to Fig. 2a-d, respectively, with: (**a**, **c**) f_1 = 15.8 ppm which is the shift of Ile-4 Cδ; and (**b**, **d**) f_1 = 51.0 ppm which is the shift of Ala-6 Cα. The vertical scales were adjusted so that the noise levels are comparable in the four spectra. Assignments of the left-to-right asterisked peaks are: **a**, Ile-4 CO, Cα, Cβ, Cγ, and Cδ; **b**, Ala-6 CO, spinning sidebands of Cα and CO, Cα, and Cβ; **c**, Ile-4 CO, Gly-13 CO, Ile-4 Cα, Gly-13 Cα, Ile-4 Cβ, Cγ, and Cδ; and **d**, Ala-6 CO, Gly-10 CO, Ala-6 Cα, Gly-10 Cα, and A6 Cβ.

3. REDOR spectra of HFP-I and HFP-K samples prepared with $[HFP]_{initial}$ = 25 μM

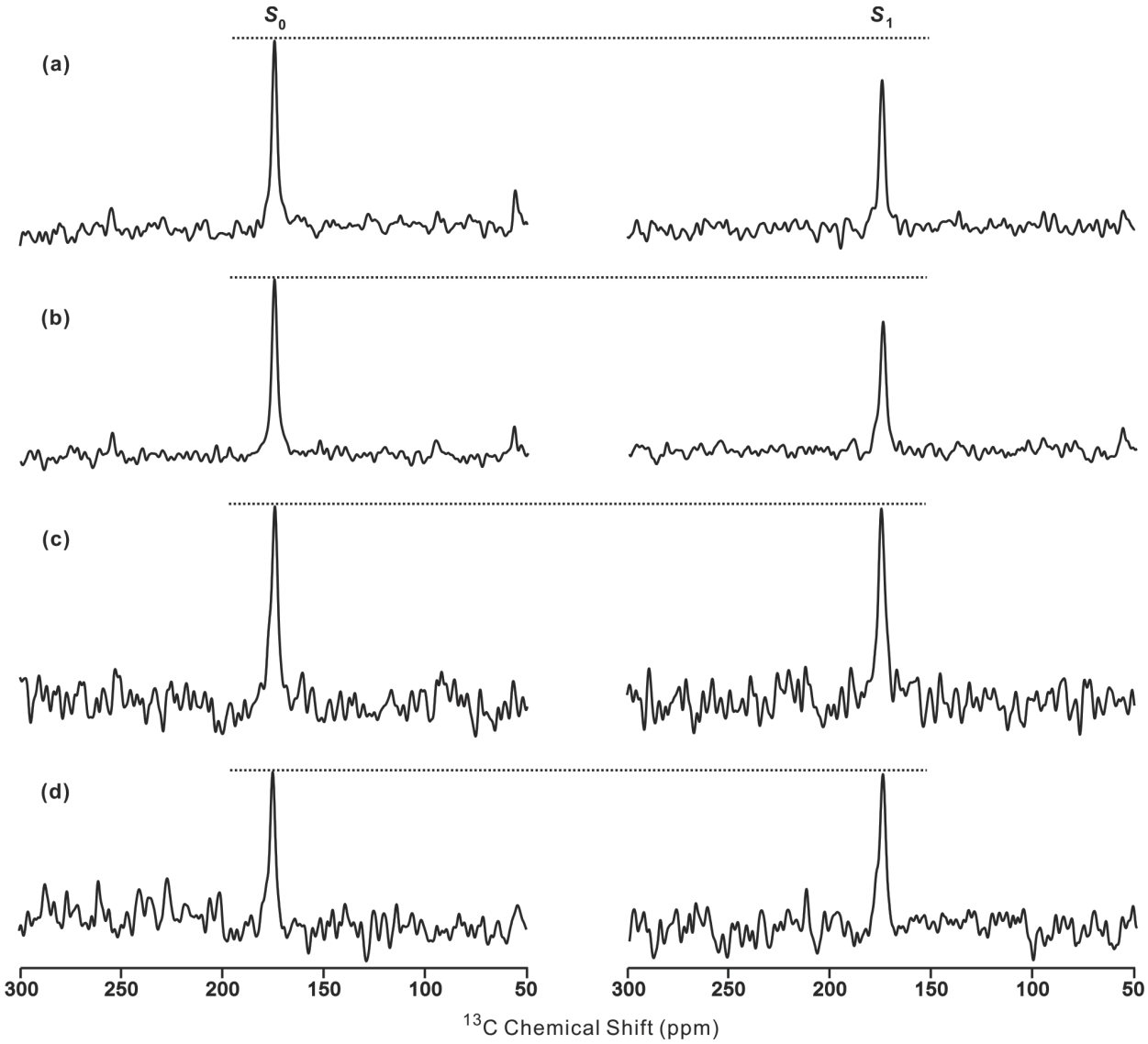

Figure S3. REDOR S_0 and S_1 spectra for membrane-associated (a, b) HFP-I and (c, d) HFP-K samples prepared with $[HFP]_{initial}$ = 25 μM. Spectra a and c were obtained with 24 ms dephasing time and spectra b and d were obtained with 32 ms dephasing time. All spectra were taken at –50 °C with MAS frequency of 8000±2 Hz and were processed with 200 Hz line broadening and polynomial baseline correction. Each S_0 or S_1 spectrum was the sum of (a) 40960, (b) 102400, (c) 30720, or (d) 40960 scans.

4. Derivation of $(\Delta S/S_0)^{cor}$ from $(\Delta S/S_0)^{exp}$ and σ^{cor} from σ^{exp}

The following parameters/approximations are used:

A1. There is 99% labeling of the Ala-14 ^{13}CO and Val-2, Gly-3, Ile-4 or Gly-5 ^{15}N sites. $S_1 = S_0$ for a labeled Ala-9 ^{13}CO in a molecule with a Val-2, Gly-3, Ile-4 or Gly-5 ^{14}N.

A2. Effects of natural abundance ^{15}N on ^{13}CO S_1 signals are evaluated using the following criteria: (1) $S_1 = 0$ for a labeled Ala-14 ^{13}CO separated by one or two bonds from a natural abundance ^{15}N at Ala-15 and Ala-14. Ala-14 S_1 is not affected by other natural abundance ^{15}N. (2) $S_1 = 0$ for natural abundance backbone ^{13}COs at Ala-1 and Val-2, Val-2 and Gly-3, Gly-3 and Ile-4, or Ile-4 and Gly-5 which are separated by one or two bonds from the labeled Val-2, Gly-3, Ile-4 or Gly-5 ^{15}N, respectively. $S_1 = S_0$ for other natural abundance backbone ^{13}CO sites. Criteria (1) and (2) are based on the close distance (≤ 2.5 Å) and consequent strong (≥ 200 Hz) dipolar coupling of ^{13}CO and ^{15}N nuclei separated by one or two bonds.

Fig. S4 displays a flow chart for the determination of $(\Delta S/S_0)^{cor}$ for HFP-H with ^{13}CO labeled Ala-14 and ^{15}N labeled Val-2. $(\Delta S/S_0)^{cor}$ for the other HFP samples were derived based on the same flow chart but only with different ^{15}N labelings.

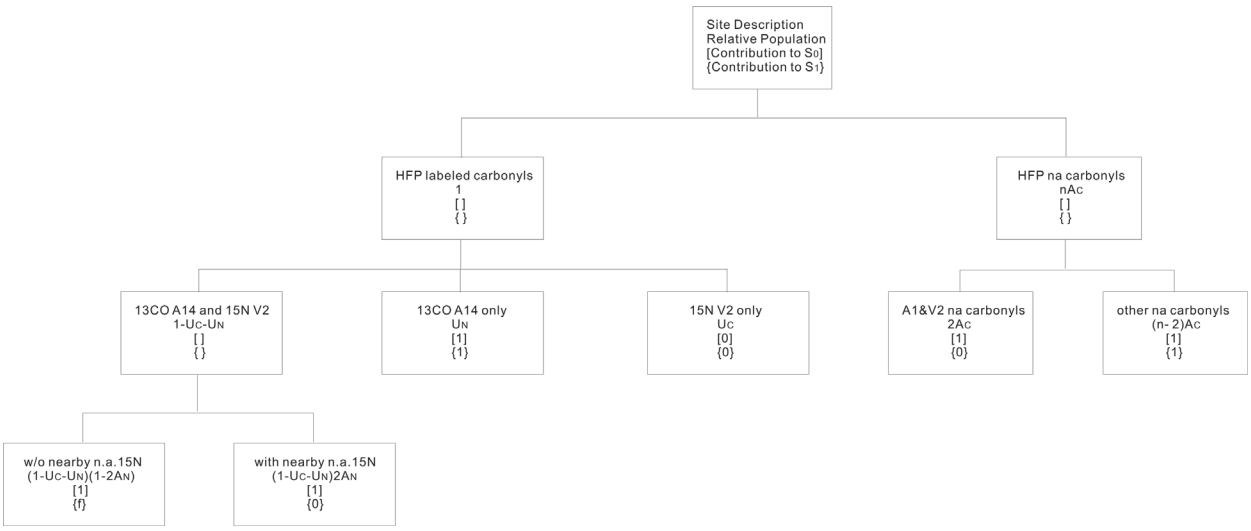

Figure S4. Flow chart of derivation of $(\Delta S/S_0)^{cor}$ for REDOR of HFP-H. The four rows in each box are in sequence: the site description, its relative population, and its contributions to S_0 and S_1.

A complete derivation of $(\Delta S/S_0)^{cor}$ follows:

$$\left(\frac{\Delta S}{S_0}\right)^{exp} = \frac{S_0^{exp} - S_1^{exp}}{S_0^{exp}} \tag{S1}$$

S_0^{exp} is expressed as the sum of contributions from labeled ^{13}CO nuclei (S_0^{lab}) and from natural abundance ^{13}CO nuclei ($S_0^{n.a.}$):

$$S_0^{exp} = S_0^{lab} + S_0^{n.a.} = 1 - U_C + n\,A_C \tag{S2}$$

where $1 - U_C$ is the fractional Ala-14 ^{13}CO labeling, A_C is the fractional ^{13}C natural abundance, and n is the total number of unlabeled peptide backbone CO sites in an HFP molecule. S_1^{exp} is also expressed as the sum of contributions from labeled ^{13}CO nuclei (S_1^{lab}) and from natural abundance ^{13}CO nuclei ($S_1^{n.a.}$):

$$S_1^{exp} = S_1^{lab} + S_1^{n.a.} \tag{S3}$$

with:

$$S_1^{lab} = (1 - U_C - U_N)(1 - 2A_N) f + U_N \tag{S4}$$

and:

$$S_1^{n.a.} = (n-2) A_C \tag{S5}$$

where $1 - U_N$ is the fractional ^{15}N labeling of the Val-2, Gly-3, Ile-4 or Gly-5 residue for HFP-H, HFP-I, HFP-J and HFP-K respectively, A_N is the fractional ^{15}N natural abundance and the parameter f:

$$f = \frac{S_1^{cor}}{S_0^{cor}} = 1 - \frac{S_0^{cor} - S_1^{cor}}{S_0^{cor}} = 1 - \left(\frac{\Delta S}{S_0}\right)^{cor} \tag{S6}$$

Incorporate Eq. S6 into Eq. S4:

$$\begin{aligned} S_1^{lab} &= (1 - U_C - U_N)(1 - 2A_N)\left[1 - \left(\frac{\Delta S}{S_0}\right)^{cor}\right] + U_N \\ &= (1 - U_C - U_N)(1 - 2A_N) - (1 - U_C - U_N)(1 - 2A_N)\left(\frac{\Delta S}{S_0}\right)^{cor} + U_N \end{aligned} \tag{S7}$$

U_C, U_N, and $2A_N$ are much less than 1 so that:

$$(1 - U_C - U_N)(1 - 2A_N) \cong 1 - U_C - U_N - 2A_N \tag{S8}$$

and:

$$S_1^{lab} \cong 1 - U_C - 2A_N - (1 - U_C - U_N - 2A_N)\left(\frac{\Delta S}{S_0}\right)^{cor} \tag{S9}$$

Incorporate Eqs. S5 and S9 in Eq. S3:

$$S_1^{\exp} = 1 - U_C - 2A_N - (1 - U_C - U_N - 2A_N)\left(\frac{\Delta S}{S_0}\right)^{cor} + (n-2) A_C \tag{S10}$$

Combine Eqs. S2, S3, S4, and S10:

$$S_0^{exp} - S_1^{exp} = \left[1 - U_C + n\,A_C\right] - \left[1 - U_C - 2A_N - \left(1 - U_C - U_N - 2A_N\right)\left(\frac{\Delta S}{S_0}\right)^{cor} + \left(n-2\right)A_C\right] \quad \text{(S11)}$$

and simplify:

$$S_0^{exp} - S_1^{exp} = 2A_C + 2A_N + \left(1 - U_C - U_N - 2A_N\right)\left(\frac{\Delta S}{S_0}\right)^{cor} \quad \text{(S12)}$$

Combine Eqs. S2 and S12:

$$\left(\frac{\Delta S}{S_0}\right)^{exp} = \frac{2A_C + 2A_N + \left(1 - U_C - U_N - 2A_N\right)\left(\frac{\Delta S}{S_0}\right)^{cor}}{1 - U_C + n\,A_C} \quad \text{(S13)}$$

and rewrite:

$$\left(\frac{\Delta S}{S_0}\right)^{cor} = \frac{1 - U_C + n\,A_C}{\left(1 - U_C - U_N - 2A_N\right)}\left(\frac{\Delta S}{S_0}\right)^{exp} - \frac{2A_C + 2A_N}{\left(1 - U_C - U_N - 2A_N\right)} \quad \text{(S14)}$$

Expressions in Eq. S14 were numerically evaluated using $A_C = 0.011$, $A_N = 0.0037$, $n = 29$, and $U_C = U_N = 0.01$ which were based on 0.99 fractional labeling of the Ala-14 ^{13}CO sites and 0.99 fractional labeling of the Val-2, Gly-3, Ile-4 or Gly-5 ^{15}N sites:

$$\left(\frac{\Delta S}{S_0}\right)^{cor} = 1.360\left(\frac{\Delta S}{S_0}\right)^{exp} - 0.030 \quad \text{(S15)}$$

Eq. S15 resulted in:

$$\sigma^{cor} = 1.360\,\sigma^{exp} \quad \text{(S16)}$$

5. Comparison of $(\Delta S/S_0)^{exp}$ with $(\Delta S/S_0)^{cor}$ for the HFP-H, I, J, and K samples

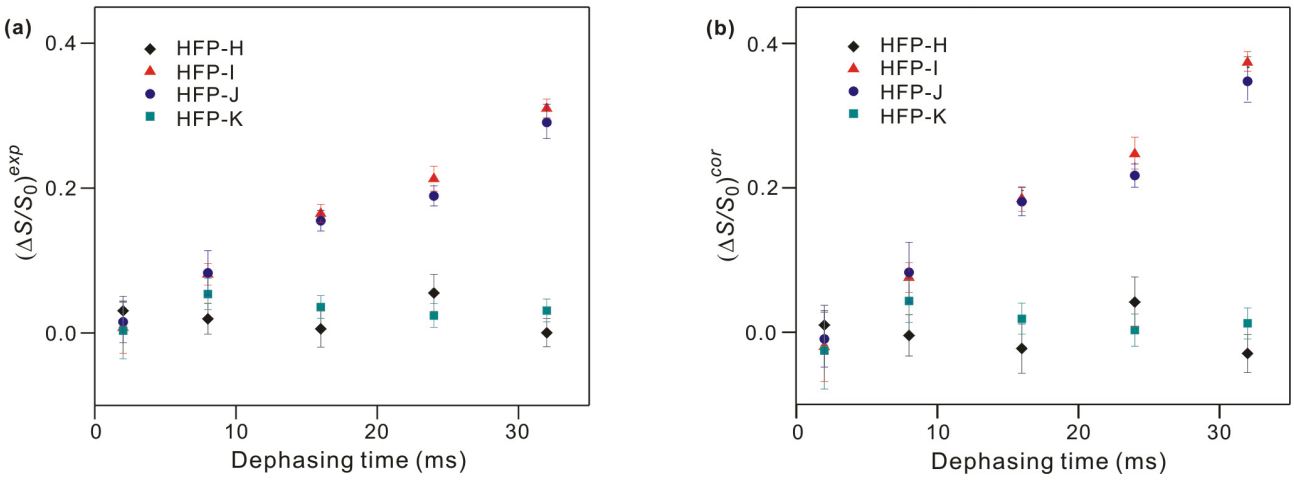

Figure S5. Plots of (a) $(\Delta S/S_0)^{exp}$ and (b) $(\Delta S/S_0)^{cor}$ vs dephasing time. The $(\Delta S/S_0)^{cor}$ plot is also shown in Fig. 5a in the main text and the symbol code is the same in the two plots. The error bars correspond to σ^{exp} (left panel) and σ^{cor} (right panel). The largest differences between $(\Delta S/S_0)^{exp}$ and $(\Delta S/S_0)^{cor}$ are at longer dephasing times.

6. Model geometries for the HFP-H, HFP-J and HFP-K samples

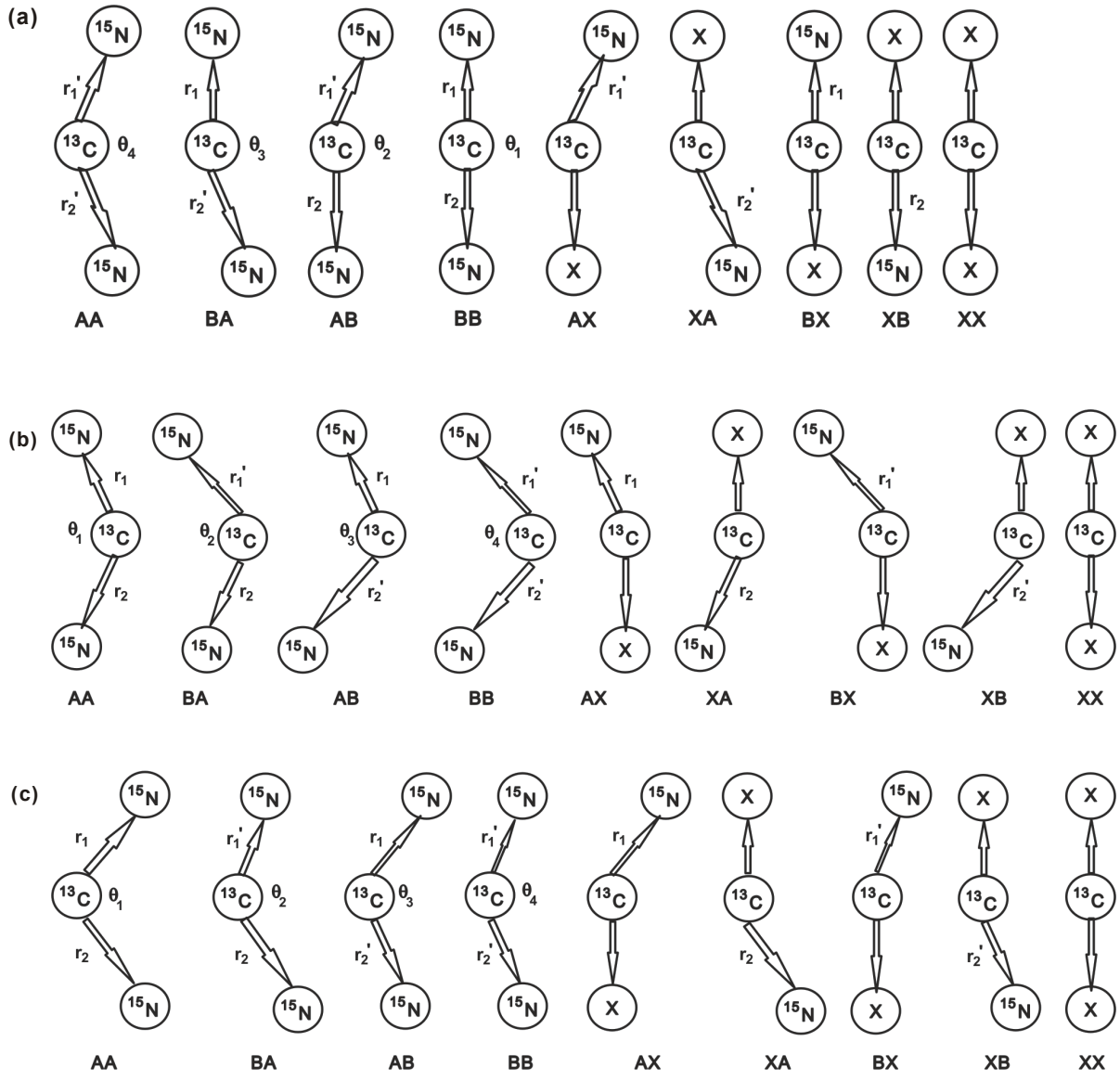

Figure S6. Models used to calculate $(\Delta S/S_0)^{sim}$ and spin geometries specific for the (a) HFP-J, (b) HFP-H, and (c) HFP-K samples. Each model includes nuclei from three adjacent strands with the Ala-14 ^{13}CO always in the middle strand and ^{15}N in the top and/or bottom strands. The first letter in the labeling of each model refers to the middle strand/top strand registry and the second letter refers to the middle strand/bottom strand registry. Registry *X* is any registry for which the

interpeptide ^{13}CO-^{15}N distance was large in the HFP-H, HFP-I, HFP-J, or HFP-K samples so that $d \approx 0$. The Ala-14 ^{13}CO is hydrogen bonded to an amide proton in the top strand. Relevant labeled ^{13}C-^{15}N distances and ^{15}N-^{13}C-^{15}N angles are: (a) r_1= 4.063 Å, r_1' = 5.502 Å, r_2= 5.455 Å, r_2' = 6.519 Å, θ_1 = 117.2˚, θ_2 = 131.1˚, θ_3 = 131.9˚, θ_4= 160.7˚; (b) r_1= 5.890 Å, r_1' = 8.925 Å, r_2= 6.431 Å, r_2' = 9.463 Å, θ_1 = 117.0˚, θ_2 = 104.2˚, θ_3 = 105.1˚, θ_4= 88.9˚; and (c) r_1= 9.319 Å, r_1' = 5.502 Å, r_2= 9.872 Å, r_2' = 6.519 Å, θ_1 = 89.0˚, θ_2 = 105.0˚, θ_3 = 104.7˚, and θ_4= 117.2˚. Each parameter value was the average of 10 specific values taken from the crystal structure of outer membrane protein G.

7. REDOR $(\Delta S/S_0)^{sim}$ for all models for the HFP-H, HFP-I, HFP-J, and HFP-K samples

Table S1. $(\Delta S/S_0)^{sim}$ for HFP-H

Dephasing time (ms)	**Model**								
	AA	BA	AB	BB	AX	XA	BX	XB	XX
2	0.00186	0.00079	0.00125	0.00016	0.00185	0.00069	0.00001	0.00001	0
8	0.02483	0.01064	0.01677	0.00223	0.02475	0.00927	0.00131	0.00092	0
16	0.09383	0.04082	0.06391	0.00862	0.09359	0.03562	0.00508	0.00356	0
24	0.20007	0.08915	0.13819	0.01910	0.19977	0.07809	0.01128	0.00792	0
32	0.33311	0.15348	0.23455	0.03357	0.33297	0.13510	0.01990	0.01398	0

Table S2. $(\Delta S/S_0)^{sim}$ for HFP-I

Dephasing time (ms)	**Model**								
	AA	BA	AB	BB	AX	XA	BX	XB	XX
2	0.01265	0.00302	0.01151	0.00186	0.01083	0.00117	0.00185	0.00069	0
8	0.15768	0.04012	0.14711	0.02483	0.13880	0.01567	0.02475	0.00927	0
16	0.48511	0.14915	0.48623	0.09383	0.46326	0.05982	0.09359	0.03562	0
24	0.75294	0.30944	0.83165	0.20007	0.80445	0.12971	0.19977	0.07809	0
32	0.86261	0.49606	1.02552	0.33311	1.00981	0.22098	0.33297	0.13510	0

Table S3. $(\Delta S/S_0)^{sim}$ for HFP-J

Dephasing time (ms)	**Model**								
	AA	BA	AB	BB	AX	XA	BX	XB	XX
2	0.00234	0.01103	0.00361	0.01265	0.00171	0.00064	0.01083	0.00117	0
8	0.03119	0.14342	0.04772	0.15768	0.02279	0.00855	0.13880	0.01567	0
16	0.11709	0.47845	0.17599	0.48511	0.08635	0.03290	0.46326	0.05982	0
24	0.24680	0.80566	0.36042	0.75294	0.18495	0.07221	0.80445	0.12971	0
32	0.40451	0.99752	0.56752	0.86261	0.30974	0.12514	1.00981	0.22098	0

Table S4. $(\Delta S/S_0)^{sim}$ for HFP-K

Dephasing time (ms)	Model								
	AA	BA	AB	BB	AX	XA	BX	XB	XX
2	0.00013	0.00184	0.00071	0.00234	0.00001	0.00001	0.00171	0.00064	0
8	0.00175	0.02451	0.00957	0.03119	0.00102	0.00072	0.02279	0.00855	0
16	0.00677	0.09268	0.03674	0.11709	0.00397	0.00281	0.08635	0.03290	0
24	0.01502	0.19777	0.08042	0.24680	0.00882	0.00626	0.18495	0.07221	0
32	0.02642	0.32954	0.13885	0.40451	0.01556	0.01105	0.30974	0.12514	0