

HIGH RESOLUTION TERTIARY STRUCTURE OF THE MEMBRANE-ASSOCIATED HIV
FUSION PEPTIDES BY SOLID STATE NUCLEAR MAGNETIC RESONANCE

By

Scott Schmick

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

CHEMISTRY

2012

ABSTRACT

HIGH-RESOLUTION TERTIARY STRUCTURE OF THE MEMBRANE-ASSOCIATED HIV FUSION PEPTIDES BY SOLID STATE NUCLEAR MAGNETIC RESONANCE

By

Scott Schmick

HIV gp41 protein catalyzes fusion between viral and host cell membranes, and its apolar N-terminal region or “fusion peptide” binds to host cell membranes and plays a key role in viral and host cell membrane fusion. Gp41 fusion can be dominantly inhibited by dilute amounts of V2E mutant gp41, but a structural basis for this inhibition has not been demonstrated. “HFP” is a construct containing the fusion peptide sequence that induces membrane vesicle fusion, and V2E mutant HFP (V2E-HFP) has reduced membrane vesicle fusion rates. Earlier solid-state NMR (SSNMR) studies showed that when HFP or V2E-HFP are associated with membranes with ~30 mol% cholesterol (mHFP or mV2E-HFP), the apolar N-terminal regions of these constructs have predominant β strand secondary structure. In mHFP, a fraction of the strands form antiparallel β sheet structure with residue $16 \rightarrow 1/1 \rightarrow 16$ or $17 \rightarrow 1/1 \rightarrow 17$ registries of adjacent strands (i.e $t = 16$ and $t = 17$ registries). 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 many isotopic labels and other structural models were also consistent with the data.

The tertiary structure of mHFP was studied using SSNMR with the rotational-echo double resonance (REDOR) pulse sequence to measure a sample’s average ^{13}CO - ^{15}N dipolar couplings. Experimental data were collected for samples with sparser ^{13}CO and ^{15}N labeling and were compared to simulated NMR data. The in-register parallel β sheet fraction was ≤ 0.15 , and a much greater fraction of antiparallel registries were identified. The accuracy of the quantitative measurements was enhanced by inclusion of “long range” natural abundance contributions in the

data analysis, and the validity of this approach was supported by a negative control sample. Furthermore, mHFP samples were prepared with a single ^{13}CO and a single ^{15}N label for which the closest ^{13}CO - ^{15}N interstrand proximity resulted from a distinct registry. These experimental data were compared to simulated data that incorporated fractional populations, f_t , of 17 different registries. These f_t were globally fit using a χ^2 metric which identified a broad distribution of antiparallel β sheet registries ($11 < t < 21$). Sequential hydrophobic residues in HFP result in intrastrand hydrophobic patches and interstrand overlap of these patches result in interstrand hydrophobic regions. These regions may insert into the vesicle membranes, and a hydrophobicity or insertion energy metric, ΔG_t^{\min} , was developed to quantify each registry's insertion energy. In general, registries present in our NMR samples had a negative ΔG_t^{\min} while registries that were not present generally had a positive ΔG_t^{\min} . A similar set of experiments were run with mV2E-HFP, and mV2E-HFP had a narrower distribution of registries where the $t = 20$ registry was significantly more populated in mV2E-HFP than in mHFP. The hydrophobic residues of HFP are located within the first 12 N-terminal amino acids, and the $t = 12$ registry was more populated in mHFP than mV2E-HFP. The $t = 12$ registry localizes hydrophobic residues which may result in deeper membrane insertion and increased vesicle fusion rates compared to the $t = 20$ registry. The $t = 20$ registry delocalizes the interstrand proximity of N-terminal hydrophobic residues which may result in shallower membrane insertion and reduced membrane fusion rates. These results provide a new, experimentally-based structural model for transdominant inhibition where co-mixing of V2E mutant gp41 and wild type gp41 may energetically favor a non-native registry distribution shifted toward longer registries for the FP region of wild type gp41.

TABLE OF CONTENTS

TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS.....	xviii
Chapter I. Introduction	1
Chapter II. Materials and Methods	14
2.1 HFP Synthesis and Sample Preparation	14
1. General HFP Preparation	14
2. HFP Solutions	17
3. Trimeric HFP (HFP _{tr}) Synthesis	18
2.2 General NMR Background	22
1. NMR Spectroscopy.....	22
2. Rotating Frame and Bloch Equations	24
2.3 Experimental Setup	26
1. General Concepts for Setting the Pulses.....	26
2. Setup Compounds.....	32
3. Setting the Magic Angle	33
4. Setting Proton $\pi/2$ Pulse.....	35
5. Proton Decoupling Field	38
6. Proton-Carbon Cross Polarization.....	41
7. Carbonyl π Pulse	42
8. Nitrogen π Pulse	43
9. Proton TPPM Decoupling to Nitrogen Rabi Frequency Ratio in HCN REDOR	44
10. Chemical Shift Referencing	46
2.4 MAS Solid-State NMR (REDOR)	46
2.5 SIMPSON Simulations.....	52
Chapter III. Natural Abundance Calculations for Solid State NMR REDOR Experiments and Quantitative Determination of In-Register Parallel β Sheet Registries in Membrane-Associated HFP	58
3.1 Background.	58
3.2 Materials and Methods.	63
1. SSNMR samples.....	63
2. Modeling.	65
3.3 Results.	67
1. Qualitative Analysis of the REDOR Data.	67
2. Natural Abundance Models.....	70
3. Quantitative Analysis of Registry Populations – Fully Constrained Model.	77
4. Quantitative Analysis of Registry Populations – Unconstrained Model.	85

3.4 Discussion	87
CHAPTER IV. Quantitative Identification of the Antiparallel Sheet Registry Distribution in Membrane Associated HFP and V2E-HFP Samples	94
4.1 Background	94
4.2 Materials and Methods	99
4.3 Results.	105
1. Fully Constrained Model mHFP Registry Distribution.....	105
2. Fully Constrained Model mV2E-HFP Registry Distribution	117
4.4 Discussion	125
1. Modeled Membrane Insertion Depth mHFP	125
2. Relevance of Broad Distribution.....	130
Chapter V. Dissertation Summary and Future Work	135
5.1 Summary	135
5.2 Membrane Location	139
5.3 Resin Bound Structure.....	145
APPENDICES	149
Appendix I. Files Checklist	150
Appendix II. Current HIV Inhibitor Drugs.....	155
Appendix III. Simple Number of Strands for Fusion Model.....	156
Appendix IV. RP-HPLC Purification, Optimization and Troubleshooting	160
1. Specific Problems.....	166
1.1 Well separated peak contains “impurities” by MALDI-TOF analysis	166
1.2 Forgot to deprotect Fmoc group.	166
1.3 Column pressure is increasing over time.	167
Appendix V. HFPdm Data and Lyophilized HFP	168
Appendix VI. SIMMOL, SIMPSON, and Fortran Files.	172
1. Sample SIMMOL Files	172
1.1 SIMMOL input file for Leu-132 ¹³ CO from the 2IWW.pdb file.	172
1.2. SIMMOL output file for the Leu-132 residue ¹³ CO from the 2IWW.pdb file.....	172
2. Sample SIMPSON Files	173
2.1. Input file 5 spin (NNCNN) SIMPSON File for Ala-12 ¹³ CO from 2WII.pdb.....	173
2.2 Output file 5 spin (NNCNN) SIMPSON File for Ala-12 ¹³ CO from 2WII.pdb	177
2.3 Input file 3 spin (CNN) SIMPSON File for Ala-12 ¹³ CO from 2WII.pdb	177
2.4 Output file 3 spin (CNN) SIMPSON File for Ala-12 ¹³ CO from 2WII.pdb.....	181
2.5 Input file 3 spin (CNC) SIMPSON File from 2WII.pdb	181
2.6 Output file 3 spin (CNC) SIMPSON File from 2WII.pdb.....	185
2.7 Input file 2 spin (CN) SIMPSON File from 2WII.pdb.....	185

2.8 Output file 2 spin (CN) SIMPSON File from 2WII.pdb	189
3. Sample Fortran Input Script Files For Global χ^2 Fittings	189
3.1 HFP 3 Registry Fitting qsub Script, "x2_fixed"	189
3.2 HFP 3 Registry Fitting Main Script, "HFP.f"	190
3.3 HFP 5 Registry Fitting qsub Script, "x2_HFP"	207
3.4 HFP 5 Registry Fitting Main Script, "HFP_5var.f"	208
3.5 V2E-HFP 3 Registry Fitting qsub Script, "x2_V2E"	226
3.6 V2E-HFP 3 Registry Fitting Main Script, "V2E.f"	226
3.7 V2E-HFP 5 Registry Fitting qsub Script, "x2_V2E"	244
3.8 V2E-HFP 5 Registry Fitting Main Script, "V2E_5var.f"	245
Appendix VII. Chapter III Table of $(\Delta S/S_0)^{exp}$ values and $\gamma_{1tuv}^{lab}(\tau)$ or $\gamma_{1t1t2uv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON.	264
Appendix VIII. Five Registry Fittings	272
Appendix IX. Chapter IV Unconstrained Fitting.....	273
Appendix X. Freed Mutations	276
Appendix XI. Raw Data for mHFP and mV2E-HFP.....	277
Appendix XII. Boltzmann Fraction Populations.	280
Appendix XIII. L9R Mutant Discussion	281
Appendix XIV. HFP, V2E, and L9R ΔG_t^{min} i_t , and n_t Values.	288
Appendix XV. Summary of Hessa Biological Hydrophobicity Scale	289
Appendix XVI. Chapter IV Tables for $\gamma_{1tuv}^{lab}(\tau)$ or $\gamma_{1t1t2uv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON for the three registry fittings.....	296
Appendix XVII. Chapter IV Tables for the unique $\gamma_{1tuv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON for the five registry fittings.	300
REFERENCES	301

LIST OF TABLES

Table 1. HFP construct labeling schemes.	17
Table 2. Oligomeric HFP synthesis summary.	20
Table 3. Error analysis for V2E-F8CG13N.....	52
Table 4. Error analysis for HFP-L12CA6N.	52
Table 5. Chapter III indices and parameters.....	73
Table 6. Chapter IV indices and parameter.....	100
Table 7. Three registry fittings for mHFP and mV2E-HFP.	106
Table 8. Five registry fittings for mHFP and mV2E-HFP.	107
Table 9. HFP-F8CG13N data sets used in global fittings.....	108
Table 10. mHFP and mV2E-HFP $^{13}\text{CO } D_{h,i_t,n_t}^{avg}$	129
Table 11. ΔG_t^{\min} values for HFP constructs.	139
Table 12. Freed fusion activity ²¹	158
Table 13. Strands Model 1.....	158
Table 14. Strands Model 2.....	159
Table 15. HFPdm and L12CA6NmN Lyophilized $(\Delta S/S_0)^{exp}$ and σ^{exp}	171
Table 16. Chapter III $(\Delta S/S_0)^{exp}$ and rms error.	264
Table 17. Unconstrained model mHFP.....	274
Table 18. mHFP $\Delta(S/S_0)^{exp}$	277
Table 19. mV2E-HFP $\Delta(S/S_0)^{exp}$	278
Table 20. mHFP fully constrained model and Boltzmann distribution based f_t	280
Table 21. Energy minimized membrane insertion energy parameters.....	288

LIST OF FIGURES

Figure 1. HIV infection model (left) and freeze fracture electron microscopy (right) of (a) binding (b) hemi-fused viral and host cell membranes (c, d) pore formation with HIV infection of the host cell (modified from literature⁸).....2

Figure 2. A summary of the gp41 sequence and regions defined from literature where FP = fusion peptide, FPPR = fusion peptide proximal region, NTH = N-terminal helix, CTH = C-terminal helix, MPER = membrane proximal region, and TMR = transmembrane region. This dissertation primarily focuses on the role of FP in membrane fusion. The ~16 residue FP sequence followed by C-terminal amino acids (HFP) was studied in this dissertation. This figure was adapted from literature¹⁵. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.....5

Figure 3. Proposed fusion mechanism (Figure provided by Dr. Kelly Sackett). This model for membrane fusion suggests that sequential events occur during viral and host cell membrane fusion: (A → B) The ectodomain of gp41 is fusion inactive prior to contact with the host cell membrane. Contact between the host cell and viral membrane is initiated when gp120 binds to CD4 and a co-receptor¹². This induces a conformational change in gp41 to the proposed pre-hairpin intermediate structure and FP binds to the host cell membrane. To my knowledge, there is currently no direct evidence that FP binds to the host cell, but synthetic fusion peptides bind and insert into model membranes^{25,26}. Additionally, no structural data have been collected to elucidate the proposed PHI structure; (B → C) The host cell and viral membrane undergo Hemi-Fusion; (C → D) Prior to completion of membrane fusion, a fusion pore is formed within the host cell membrane. HFP constructs have been shown to insert into model membranes²⁶ which appears to be driven by local hydrophobicity of HFP secondary and tertiary structures (See Chapter IV). Similarly, FP may insert into host cell membranes which may enable FP to be involved in pore formation. The membrane fusion pore is enlarged after or during the formation of hairpin structure²⁷. Additionally, peptides with the CTH sequence inhibit membrane fusion by a proposed mechanism where these peptides bind to the NTHs in the Pre-Hairpin Intermediate structure (B and C) which prevents the formation of the 6 helix bundle structure (D). This lead to development of the membrane fusion inhibitor drug Enfuvirtide²².....6

Figure 4. (a) Electron micrograph of HIV and a T cell where the width of the HIV membrane is ~100 nm. The increased electron density between the HIV and host cell membranes suggests that multiple proteins from the HIV membrane are in contact with the T cell membrane. (b) Model of HIV contact with a T cell based upon electron micrographs where multiple proteins are believed to be between the HIV and T cell membranes. This figure was modified from literature⁴⁶.....9

Figure 5. The $t = 16$, $t = 17$, and in-register parallel β sheet registries are illustrated. Isotopic labeling schemes to detect these registries are illustrated where red underlined atoms indicate

^{13}CO labeled residues and blue underlined atoms indicate ^{15}N labeled residues. These specific labeling schemes are effective because they have ^{13}CO - ^{15}N proximities of $\sim 4 \text{ \AA}$ 12

Figure 6. Purification of HFPtr synthesis. The product of each peak was identified by MALDI-TOF mass spectrometry, Figure 7, and each peak corresponded to the following: (1) HFPWK₅CA; (2) HFPdm(Lys) and HFPdm(Cys); (3) HFPdm(Lys) and HFPdm(Cys); (4) HFPtr; and (5) HFPte. Additionally, peaks (4) and (5) were not present in HFPdm(Cys) which further supports that these peaks result from HFPtr and/or HFPte. Peak (5) was not used for any experiments in this dissertation while Peak (4) was collected for the one HFPtr experiment in Chapter IV. The identity of HFPte would have to be confirmed before meaningful experimental data is collected for this peptide. 20

Figure 7. MALDI-TOF mass spectroscopy from the HFPtr purification displayed in Figure 6. ... 21

Figure 8. A dipole moment in the laboratory frame (x, y, and z axes) where the B_0 external magnetic field is aligned along the z-axis. A new axis, x' , rotates at the Larmor frequency only if $\omega_0 = \omega_1$ 26

Figure 9. FID's for cross polarization acquisitions of the ^{13}CO labeled i4 peptide where the red horizontal line signifies the zero value or baseline for the y-axis. (a) The FID decays to the baseline. (b) The FID decays to a positive y-value as evident by the zero point value of the FID ($\sim 2.5 \text{ ms}$) being above the red line. 31

Figure 10. A pulse of a single radiofrequency is illustrated in the time domain (a,c) and the frequency domain (b,d). The relationship between pulse length (τ_p), radiative field (B_1), and the energy of the B_1 field are illustrated. The excitation frequency is centered about the transmitter frequency, ω_1 . The figure was made using concepts from the literature where the frequency domain is represented as the Fourier transform of the time domain where the Fourier transform of a step function can be represented by a sinc function (b), (d) ⁶⁸ 32

Figure 11. The chemical structure of adamantane. 33

Figure 12. Magic angle spectrum FID of KBr with 64 acquisitions and a $\text{pd} = 0.10 \text{ s}$. The transmitter was moved to the ^{79}Br resonant frequency of KBr, and an exponential decay is observed in the free induction decay (FID). 35

Figure 13. Proton pulse length arrayed using the I4 peptide, a $\text{pd} = 1.0 \text{ sec}$ and 10 acquisitions per spectrum. The H rf ampl parameter was set to 0.2300 and the H 90 pulse parameter was arrayed from 1.0 to 15.0 μs by 1.0 μs increments. Maximum signal was observed at $\sim 5 \mu\text{s}$ which approximately corresponds to a $\pi/2$ pulse. Zero signal should be observed between 10.0 and 11.0 μs which would correspond to a π pulse. The change in signal intensity is the greatest at points surrounding the π pulse so it is more accurate to set the $\pi/2$ ^1H pulse length by identifying the π pulse length and dividing pulse length time by 2. To do this for the above data, the H 90 pulse parameter could be arrayed from 10.0 to 11.0 μs by increments of 0.1 μs and increase the number

of scans per spectrum to enhance the signal to noise ratio. Alternatively, a better approach would be to set the H 90 pulse parameter to 10.0 μ s and array the H rf ampl parameter. Upon determining which H rf ampl that yields zero signal, change the 10.0 μ s to 5.0 μ s to set the $\pi/2$ pulse. While the proton amplifier output voltage is approximately linear with respect to the H rf ampl parameter, it is not exactly linear so you need to divide the pulse length by 2 and not the amplifier input parameter.38

Figure 14. Cross polarization⁷⁴ arrayed with the I4 peptide using 5 acquisitions, 1.0 s pulse delay (pd), MAS frequency = 10 kHz, and contact time = 2 ms. Cross polarization is included in the REDOR experiments to transfer magnetization from the highly abundant and polarizable ¹H nuclei to the more dilute ¹³CO nuclei to increase the ¹³CO signal to noise ratio by increasing the signal per acquisition. Additionally, the longitudinal relaxation rate is approximately 4 times faster for ¹H than for ¹³C nuclei in organic solids which allows for acquisition of ~4 times as many FIDs using cross polarization than for waiting for ¹³C repolarization⁷⁵. The cp_ramp pulse program was used, and the X cp ampl change parameter was set to 0.04. A 48 kHz ¹H cross polarization and $\pi/2$ pulse was used, and the aXcp X cp ampl start parameter was arrayed from 0.00-0.70 by 0.05 increments. Maximum signal intensities were observed between 0.25-0.30 and 0.35-0.40. Either of these regions could be used for CP since the signal intensities are nearly equivalent. Additionally, the contact time also should be arrayed to set up an optimal CP, and it is important to use a setup compound with similar nuclear magnetic relaxation rates. Typical optimal contact times for peptides in REDOR experiments are 1.5-2.0 ms for ¹³CO nuclei. Finally, array the X cp ampl change (i.e. the ¹³C ramp parameter) parameter to obtain the maximum signal where 0.04 is a typical value, but this value may vary depending upon the amplifier being used.41

Figure 15. The cp_zfilter pulse sequence (CP – $\pi/2$ – τ_z – π – acquisition) was used to set the ¹³CO π pulse with 10 acquisitions and a 1.0 sec pulse delay using the i4 peptide and a MAS = 10 kHz. When the pulse lengths are set correctly, precession of the ¹³CO magnetization can be followed using the Bloch Equation, Eq (9) where (1) equilibrium, M = +z; (2) CP, rotates to the xy plane; (3) $\pi/2$, rotates to the –z axis; (4) τ_z - remains along –z axis; (4) π , flips back to +z axis; (5) acquisition with consequent zero signal detection in the rotating frame when pulses are set accurately.42

Figure 16. REDOR spectra (32.2 ms dephasing time) of the I4 peptide where each set of black and red spectra have a corresponding ¹³CO peak in the S₀ and S₁ spectra for each arrayed points, respectively. Each S₀ and S₁ spectrum was the sum of 120 acquisitions with pd = 2.0 s. The aY180 Y 180 ampl parameter was arrayed from 0.07 to 0.21 by 0.01 increments and the pw180Y Y 180 parameter was set to 20.0 sec (w_R = 25 kHz as determined by Eq (13)-(16)). Maximum dephasing was observed with the pw 180Y Y 180 ampl = 0.13 where $\Delta S/S_0$ = 0.82 where maximum dephasing corresponds to a ¹⁵N π pulse. For further understanding of the REDOR experiment, see 2.4 MAS Solid-State NMR (REDOR).43

Figure 17. The $\Delta S/S_0$ of the i4 peptide is plotted against the ^1H decoupling w_R to ^{15}N w_R ratio for $\tau = 16.2, 24.2, 32.2, 40.2$, and 48.2 ms dephasing times where the ^{15}N π pulse was $25\ \mu\text{sec}$ ($w_R = 20\ \text{kHz}$), the ^{13}C π pulse was $11\ \mu\text{sec}$ ($w_R = 45\ \text{kHz}$), the $\text{pd} = 2.0\ \text{s}$, and 350 scans were acquired for S_0 and S_1 of each data point. These HCN REDOR experiments appear to require a ^1H decoupling w_R to be at least 3.5 times greater than the ^{15}N w_R to obtain maximum $\Delta S/S_0$. The ^{13}CO nuclei were decoupled from ^1H nuclei at a ^1H decoupling to ^{13}C w_R ratios approximately ≥ 1.5 . This is evident since equivalent error bars were obtained for data within a dephasing time period for ^1H decoupling fields of $\geq 60\ \text{kHz}$. The ^{13}CO atoms are not directly bonded to ^1H atoms which results in weaker ^1H - ^{13}CO heteronuclear dipolar couplings ($\sim 3.8\ \text{kHz}$ which was determined using Eq. (19) for a r_{HC} of $2.0\ \text{\AA}$ between the carbonyl carbon and the adjacent residue's amide proton (See 1K09.pdb) relative to amide heteronuclear ^1H - ^{15}N dipolar couplings ($11.648\ \text{kHz}$ corresponds to a $r_{\text{HN}} = 1.015\ \text{\AA}$) which is approximately the width of the dipolar powder pattern for amide ^{15}N in proteins⁷⁶. Based upon the current set of experiments, it is unclear whether the ^1H decoupling to ^{15}N ratio is a causation or correlation relationship. There may be a ^1H - ^{15}N dipolar interaction that results in lower $\Delta S/S_0$ values where ^1H decoupling fields of $>80\ \text{kHz}$ may be required to average out effects due to ^1H - ^{15}N dipolar couplings.44

Figure 18. Adamantane ^{13}C spectrum prior to chemical shift referencing. The transmitter was set near the ^{13}CO Larmor frequencies to increase ^{13}CO signal intensity in REDOR experiments as described in Figure 10. The chemical shift for the left adamantane peak is $40.5\ \text{ppm}$ downfield from the tetramethyl silane (TMS) internal standard reference, but is observed at $-113.7\ \text{ppm}$ in this figure. Therefore, $154.2\ \text{ppm}$ should be added to the chemical shift to correctly reference chemical shifts. Chemical shift referencing is important since referenced chemical shifts of carbon nuclei in peptides provide information about local secondary structure⁷⁷46

Figure 19. (a) HFPs where red and blue correspond to ^{13}CO and ^{15}N labeled residues, respectively. (b) HFP-NC, HFP-P, HFP-A, and HFP-AP were SSNMR samples which each contained a mixture of ^{13}CO and ^{15}N labeled peptides in 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 β sheet structure with a molecular mixture of ^{13}CO and ^{15}N labeled peptides in the sample. (c) Registries probed by the SSNMR REDOR experiments and labeled ^{13}CO /labeled ^{15}N proximities for the membrane-associated HFPs in these registries. Consideration of residue $1 \rightarrow 16$ or $1 \rightarrow 17$ registries is based on the fully extended conformation in HFP. For parallel sheets, there is CO (residue h) – HN (residue $h+1$) hydrogen bonding of adjacent strands.61

Figure 20. REDOR S_0 and S_1 ^{13}C SSNMR spectra at $32.2\ \text{ms}$ dephasing time for (a) HFP-NC, (b) HFP-P, (c) HFP-A, or (d) HFP-AP. Each spectrum was processed with $200\ \text{Hz}$ line broadening and baseline correction and was the sum of: (a) 38624; (b) 23488; (c) 24914; or (d)

14240 scans. Relatively narrow ^{13}C O signals were observed in the HFP-P, HFP-A, and HFP-AP samples because the HFPs were membrane-associated with predominant β sheet conformation at the labeled ^{13}C O site. A broader ^{13}C O signal was observed in the HFP-NC sample because there was no membrane and there were populations of lyophilized HFP with either α helical or β sheet conformation at the labeled ^{13}C O site.68

Figure 21. (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 nad models with fractional populations: α helical, 0.5; min β sheet, 0.25; max β 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 σ^{exp} is ± 0.02 . Variation of ± 0.02 in $(\Delta S/S_0)^{exp}$ was also observed between two different preparations of the same sample type, e.g. HFP-A.69

Figure 22. (a, b) Schematic diagrams of the HFP-F8 region of the HFP-NC sample in antiparallel β sheet structure with labeled ^{13}C O 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) β sheet backbone representations of the respective boxed regions of panels a and b with labeled ^{13}C O in red and possible na ^{15}N sites in blue, i.e. sites for which a na ^{15}N is within 7 Å of a labeled ^{13}C O. A particular spin geometry will have only one ^{15}N . The min nad model is shown in panel c and each spin geometry will have either one labeled ^{13}C O and one na ^{15}N (#1, 2 or 3) or two labeled ^{13}C O and one na ^{15}N (#4, 5, or 6). The max nad model is shown in panel d and each spin geometry will have one labeled ^{13}C O and one na ^{15}N72

Figure 23. Schematics of three adjacent HFPs for HFP-A, i.e. $u = 2$, in (a-d) fully constrained or80

Figure 24. Contour plots of χ^2 vs f_a parallel and f_a antiparallel fractional populations for (a) fully constrained and (b) unconstrained models. In each plot, f_a is the sum of populations of 1→17/1→17 and 2→17/1→16 parallel registries and f_b is the sum of populations of 16→1/1→16 and 17→1/1→17 antiparallel registries.84

Figure 25. Pictorial model of HFP (red lines) binding to membranes followed by antiparallel β 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 β 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.90

Figure 26. A summary of the gp41 sequence and regions defined from literature where FP = fusion peptide, FPPR = fusion peptide proximal region, NTH = N-terminal helix, CTH = C-terminal helix, MPER = membrane proximal region, and TMR = transmembrane region. This figure was adapted from literature¹⁵96

Figure 27. REDOR S_0 and S_1 ^{13}C SSNMR spectra at 48.2 ms dephasing time for (a) mHFP-A6CG3N, (b) HFP-L12CG5N, (c) HFP-F8CL12N, or (d) HFP-L9CG16N. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 23766; (b) 21454; (c) 40331; or (d) 39133 scans. (e) $(\Delta S/S_0)^{exp}$ ($\tau = 48.2$ ms) for all samples, u . (f) Plots of $(\Delta S/S_0)^{exp}$ vs dephasing time with the rms error. Isotopic labeling of each mHFP is displayed in the legend that correspond to HFP-A6CG3N (black, square), HFP-L12CG5N (red, circle), HFP-F8CL12N (cyan, triangle), and (d) HFP-L9CG16N (orange, inverted triangle). Variation less than ± 0.02 in $(\Delta S/S_0)^{exp}$ was also observed between two different preparations of the same sample type, e.g. HFP-F8CG13N (not displayed here)..... 109

Figure 28. Sample indices, u , with the corresponding labeling schemes are displayed along with the registries, t , that result in labeled r_{CN} of ~ 4.1 Å and ~ 5.5 Å that respectively correspond to hydrogen-bonded and non-hydrogen bonded $^{13}\text{CO} - \text{H}^{15}\text{N}$. Membrane inserted regions are highlighted in yellow, and the corresponding n and i values are listed. 115

Figure 29. Double-y plot where f_t populations (black) and ΔG_t^{min} (red) are plotted for each registry for mHFP (a) and mV2E-HFP (b)..... 116

Figure 30. (a) The $(\Delta S/S_0)^{exp}$ ($\tau = 48.2$ ms) for mHFP (black) and mV2E-HFP (green). (b) The f_t for mHFP (black) and mV2E-HFP (green) for the fully constrained model using the 3 registry fitting method. Both (a) and (b) demonstrate that mV2E-HFP has a smaller population of shorter registries ($t < 16$) and that mV2E-HFP has a larger population of longer registries ($t > 17$)..... 120

Figure 31. REDOR S_0 and S_1 ^{13}C SSNMR spectra at 48.2 ms dephasing time for (a) HFP-F8CA21N, (b) HFP-F8CG13N, (c) HFPtr-F8CG13N, (d) V2E-F8CG13N, (f) HFP-L12CA6N . (g) V2E-L12CA6N, (h) HFP-L9CG5N, or (i) V2E-L9CG5N. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 46816; (b) 36665; (c) 19372; (d) 34271; (f) 44931; (g) 46231; (h) 40272; or (i) 46809 scans. (e), (j) Plots of $(\Delta S/S_0)^{exp}$ vs dephasing time with the rms error. Isotopic labeling of each mHFP is displayed in the legend that correspond to HFP-F8CA21N, (black, square), HFP-F8CG13N (orange, inverted triangle), HFPtr-F8CG13N (cyan, triangle), (d) V2E-F8CG13N (red, circle), (f) HFP-L12CA6N (dark yellow, square), (g) V2E-L12CA6N (purple, circle), (h) HFP-L9CG5N (green, triangle), or (i) V2E-L9CG5N (wine, inverted triangle). Variation less than ± 0.02 in $(\Delta S/S_0)^{exp}$ was also observed between two different preparations of the same sample type, e.g. HFP-F8CG13N.... 124

Figure 32. Membrane insertion depth model as described in the text. Previous work has demonstrated that the Ala-1 carbonyl carbon is ~ 5 Å from the lipid phosphorus, and this phosphate region is referred to as the water/bilayer interface. 126

Figure 33. The calculated average membrane insertion depth of each residue's ^{13}CO , D_{h,i_t,n_t}^{avg} , is plotted for HFP (red) and V2E-HFP (blue).	128
Figure 34. Calculated membrane insertion depth of the Ala-6 ^{13}CO for HFP (red) and V2E-HFP (blue) is plotted for each registry. Additionally, f_t for mHFP (black) and mV2E-HFP (green) are plotted for each registry. The mV2E-HFP registry populations are shifted toward longer registries ($t > 17$) relative to mHFP and the calculated membrane insertion depth of the Ala-6 ^{13}CO is ≥ 0 for these registries. This is consistent with previous work where the Ala-6 ^{13}CO is inserted deeper in mHFP relative to mV2E-HFP ²⁶	130
Figure 35. The membrane insertion energies were derived from the Hessa biological hydrophobicity scale for the HFP, V2E-HFP and L9R-HFP by the methods described in Chapter IV. The L9R-HFP has predominantly positive ΔG_t^{\min} whereas both HFP and V2E-HFP have many registries with negative ΔG_t^{\min} which suggests that the distribution of registries should be different between constructs. Additionally, it is not obvious that mL9R-HFP should form membrane inserted β sheets since $t < 12$ registries were minimally populated in mHFP, and $t > 12$ registries have positive ΔG_t^{\min} in mL9R-HFP. Unlike mHFP and mV2E-HFP, mL9R does not predominantly form membrane inserted sheets, Appendix XIII.	138
Figure 36. Cholesterol molecules with carbon atoms numbered (a) Cholesterol (b) Cholesterol-2,2,3,4,4,6-d ₆ and (c) Cholesterol-25,26,26,26,27,27,27-d ₇	144
Figure 37. V2E-L9CI4N resin bound (prior to cleavage) with a MAS speed of 10 kHz.	147
Figure 38. V2E-L9CI4N resin bound (prior to cleavage) with a MAS speed of 6 kHz.	148
Figure 39. Chart of commercially available anti HIV drugs. This chart was last updated 12\14\2010 and was taken from www.aidsmeds.com	155
Figure 40. (a) HFP purification with a “large” C18 column (10-15 μm pore size). (b) HFP purification with a “small” C4 column (10 x 250 mm and 5 μm pore size). Better peak resolution was obtained with the C4 column. (c) Typical MALDI-TOF mass spectroscopy of peak 2 HFP from a purification similar to (b) where the expected mass was 3151 +2 g/mol where the +2 refers to the mass gain from the ^{13}C and ^{15}N isotopes.	161
Figure 41. Preliminary gradient of 15% to 80% solvent B over 40 minutes. In developing purification protocols, small amounts of crude peptide were used to make product peaks narrow. By mass spec, Peak 2 is the confirmed product peak. From Eq (62), the variables have the following values: $C_i = 15\%$; $G_s = 1.625 \text{ \%/min}$; $t_e \sim 26.5 \text{ min}$; and $P_e = 45\%$	162

Figure 42. A linear 39-48% gradient was run over 18.5 minutes for purifying the product peak. From Eq (62), the variables have the following values: $C_i = 39\%$; $G_s = 0.5 \text{ \%}/\text{min}$; $t_e \sim 15.5 \text{ min}$; and $P_e = 46.5\%$. The initial starting concentration was chosen to make the elution time around 15 minutes which was calculated by Eq (62) and a more gradual gradient was used to better separate peaks 1 and 2 from Figure 41. Also, at the end of the program, the gradient was ramped up to 80% solvent B over 0.5 minutes and the flow rate was increased to 9 mL/min to clean the column after each run. After 5 minutes, the gradient concentration of solvent B was returned to 39% over 0.5 minutes and the column equilibrated at this concentration for 3 minutes to prepare for the next run. 163

Figure 43. The program from Figure 42 was used, but a higher loading volume of the crude peptide was used which resulted in poor separation of our product peak. 164

Figure 44. A linear 37-47% gradient was run over 20 minutes. The peaks were separated better with minimal peak broadening. 164

Figure 45. Nonlinear gradients can be used to separate peak 1 from peak 2. The gradient broadened peak 1 using a more gradual slope initially while the gradient was steeper from 15 to 21 minutes to retain the sharpness of peak 2. To optimize the time of the program, it's best to have your product elute during the period where the ramp is up to 75% solvent B since nothing is achieved during this time in Figure 41-Figure 44. Recall, the elution time of a peak is 8 minutes. Therefore, in this figure, peak 2 was collected during 75% use of solvent B, but peak 2 actually began coming off the column at time $t_e - 8$ or $\sim 16-18$ minutes. 165

Figure 46. The ramp was modified to separate peak 2 from peak 3. This program was created because peak 1 also contains peptide with our products molecular weight which was collected for potential future use. Peak 3 should also be collected if the product peak is low relative to other syntheses. Peak 3 can contain HFP with N-terminal or sidechain protecting groups. 165

Figure 47. MALDI-TOF mass spectroscopy of purified HFP-L9G10. In MALDI-TOF experiments, increasing the laser power can increase the signal to noise, but it can also lead to peptide fragmentation where fragmentation can occur C-terminal of amino acids with basic sidechain groups^{65,66}. Alternatively, gas phase degradation of the peptide may be unlikely. The peptide degradation may result from hydrolytic cleavage in the matrix or possibly during isolation⁶⁴. Fragmentation of the HFP-L9G19 peptide appeared to occur C-terminal of the Arg-22, Lys-29, and Lys-30 where the respective fragments detected were likely
 AVGIGALFLGFLGAAGSTMGAR (2038 +2 g/mol),
 AVGIGALFLGFLGAAGSTMGARSWKKKKK (2952 +2 g/mol), and
 AVGIGALFLGFLGAAGSTMGARSWKKKKKKK (3080 +2 g/mol), and the HFP product had an expected mass of 3151 + 2 g/mol. 166

Figure 48. REDOR data for HFP (black boxes) and HFPdm (red circles) are displayed with error bars that are associated with rms deviation and labeling corresponds to (a) L9CG5N and (b) L12CG5N. 169

Figure 49. Chapter III spin geometries and simulated data. 265

Figure 50. Flow chart for unconstrained iterative fitting. Each iteration is denoted by the variable κ , and the χ^2_u calculations are found in Chapter IV, Eq (50).....	275
Figure 51. The ΔG_t^{\min} are plotted for registries $t = 8-24$ for HFP, V2E-HFP, F11G-HFP, and F11V-HFP. For each registry, the F11G-HFP ΔG_t^{\min} are greater than or equal to the HFP ΔG_t^{\min} which may contribute toward F11G-HFP's lower fusion activity. The F11V-HFP ΔG_t^{\min} is approximately equal to the HFP ΔG_t^{\min} relative to mHFP for each registry.	276
Figure 52. The membrane insertion energies were derived from the Hessa biological hydrophobicity scale for the HFP and L9R-HFP by the methods described in Chapter IV. The L9R-HFP has predominantly positive $\Delta G_{n_t}^{\min}$ whereas HFP has many registries with negative $\Delta G_{n_t}^{\min}$ which suggests that the distribution of registries should be different between constructs. Additionally, it is not obvious that mL9R-HFP should form membrane inserted β sheets since $t < 12$ registries were minimally populated in mHFP, and $t > 12$ registries have positive $\Delta G_{n_t}^{\min}$ in mL9R-HFP.	284
Figure 53. NMR sample of mV2E-HFP (left) compared to a water standard (right) prior to centrifugation and after mixing overnight. Aggregation of LUV's is evident in mV2E-HFP under our sample preparation conditions.	285
Figure 54. NMR samples of mL9R-HFP (L9R), mHFP (WT), mV2E-HFP (V2E) prior to centrifugation and after mixing overnight. The mL9R-HFP sample appeared to be more transparent than both mHFP and mV2E-HFP, but LUV aggregation was evident.	286
Figure 55. mL9R-HFP with F8CG13N labeling. The chemical shift of 174.6 ppm and 7.5 ppm line full-width at half maximum height indicate the presence of a distribution of secondary structures since the peak spans chemical shifts of α helical, random coil and β sheet structures.	287
Figure 56. The model systems were composed of two transmembrane domains (TM1 and TM2), two luminal domains (P1 and P2), and two glycosylation acceptor sites (G1 and G2). A third helical transmembrane domain (H) is illustrated in red. Translocation of the H segment from the membrane allows for glycosylation of both G1 and G2 while membrane insertion of the H segment only allows for glycosylation of G1. This figure was modified from literature ⁹³	290
Figure 57. The Hessa biological hydrophobicity scale. This figure was taken from literature ⁹³	291
Figure 58. The positional dependence of amino acids. Key points from these figures are discussed below. These figures were taken from literature ⁹³	292

Figure 60. Additional spin geometries and simulated data for Chapter IV five registry fittings.	300
---	-----

LIST OF ABBREVIATIONS

a: $t = 16$ or 17 antiparallel registry

AIDS: Acquired Immune Deficiency Syndrome

CHOL: Cholesterol

CTH: C-terminal helix

CO: carbonyl

CP: Cross polarization

DIEA: N, N-diisopropylethylamine

DMAP: dimethylaminopyridine

DMPC: 1, 2-dimyristoyl-sn-glycerol-3-phosphocholine

DTPC: 1, 2-di-O-tetradecyl-sn-glycerol-3-phosphocholine

DTPG: 1, 2-di-O-tetradecyl-sn-glycerol-3-[phosphor-rac-(1-glycerol)]

FID: free induction decay

Fmoc: 9-fluorenylmethoxycarbonyl

FT: Fourier Transform

FTIR: Fourier Transform Infrared

FP: Fusion Peptide

HBTU: O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate

HEPES: N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid

HFP: HIV Fusion Peptide

HFPdm: HFP dimer

HFPmn: HFP monomer

HFPtr: HFP trimer

HIV: Human Immunodeficiency Virus

HOBt: 1-hydroxybenzotriazole

HPLC: high-performance liquid chromatography

IR: Infrared

I4: Ac- AEAAAKEAAAKEAAKA-NH₂ peptide with ¹³C and ¹⁵N labels at A9 and A13, respectively.

LUVs: Large Unilamellar Vesicles

MALDI: Matrix-Assisted Laser Desorption/Ionization

MAS: magic angle spinning

max nad: maximum natural abundance dephasing

MD: Molecular Dynamics

mHFP: HFP associated with membranes containing approximately 30 mol% cholesterol

mL9R-HFP: L9R-HFP associated with membranes containing approximately 30 mol% cholesterol

mV2E-HFP: V2E-HFP associated with membranes containing approximately 30 mol% cholesterol

mHFPdm: HFPdm associated with membranes containing approximately 30 mol% cholesterol

mHFPtr: HFPtr associated with membranes containing approximately 30 mol% cholesterol

min nad: minimum natural abundance dephasing

na: natural abundance

nad: natural abundance dephasing

NTH: N-terminal helix

NMR: Nuclear Magnetic Resonance

pd: pulse delay

PDB: Protein Data Bank

PHI: pre-hairpin intermediate

POPC: 1, 2-dimyristoyl-sn-glycerol-3-phosphocholine

POPG: 1, 2-dimyristoyl-sn-glycerol-3-[phosphor-rac-(1-glycerol)]

REDOR: Rotational-echo Double Resonance

RMSD: Root-Mean Squared Deviation

RP-HPLC: reversed-phase high-performance liquid chromatography

rf: Rabi frequency

r_{XY} : X-Y internuclear distance

SDS: sodium dodecyl sulfate

SIMPSON: simulation program for solid-state NMR spectroscopy

TFA: trifluoroacetic acid

TM: Transmembrane

TOF: time-of-flight

TPPM: Two-pulse phase modulation

WT: wild type

Chapter I. Introduction

Viral replication is initiated by infection of a host cell where infection requires membrane fusion of the viral and host cell membranes¹, **Figure 1**. Vaccines have been developed to build resistance to viral infections and minimize the effects of diseases such as measles, mumps, and small pox to name a few². Relative to other viruses, the human immunodeficiency virus (HIV) strains have higher mutation rates³, and in the absence of a vaccine, new types of drugs will be needed. HIV infects T helper cells, regulatory T cells, monocytes, macrophages and dendrite cells⁴. Depletion of T cells can result from uncontrolled HIV infection and consequent development of acquired immunodeficiency syndrome (AIDS)⁵ which often results in fatality. The consequences of this disease have inspired efforts to develop antiviral therapeutic drugs that target enzymatic activity and protein-protein interactions at various stages of the HIV life cycle⁴. These efforts have decreased the rate of HIV infection within infected patients⁴, and the “death sentence” disease of the late 80’s and early 90’s can be viable to live with as high profile MSU alumni and NBA Hall of Famer Earvin “Magic” Johnson has demonstrated. However, the World Health Organization estimates that there are currently ~33 million people living with HIV and ~2 million deaths per year due to HIV infection world wide (2009 statistics). Thus, development of a HIV vaccine to prevent infection is critical, and continued development of HIV inhibitor drugs is equally important for therapeutic treatment of infected patients. Small molecule candidates for HIV inhibitory drug design have been identified using computational and high throughput screening methodologies^{4,6,7}. In general, the success rate of discovery and implementation of

small molecule candidates for anti-HIV drugs is enhanced by understanding the HIV life cycle and identifying key interactions within the HIV lifecycle to inhibit. Anti-HIV drugs have targeted different stages of the HIV life cycle (**Appendix II**), and only two molecules are commercially available that inhibit HIV entry, Enfuvirtide (discussed below, targets gp41) and Maraviroc (binds to chemokine co-receptor CCR5)⁴. Thus, HIV entry inhibitor drugs appear to be an underdeveloped area relative to other stages of the HIV life cycle (**Appendix II**), and the efficiency of drug design could be enhanced by further knowledge of important protein-protein and protein-membrane interactions that are necessary for HIV entry.

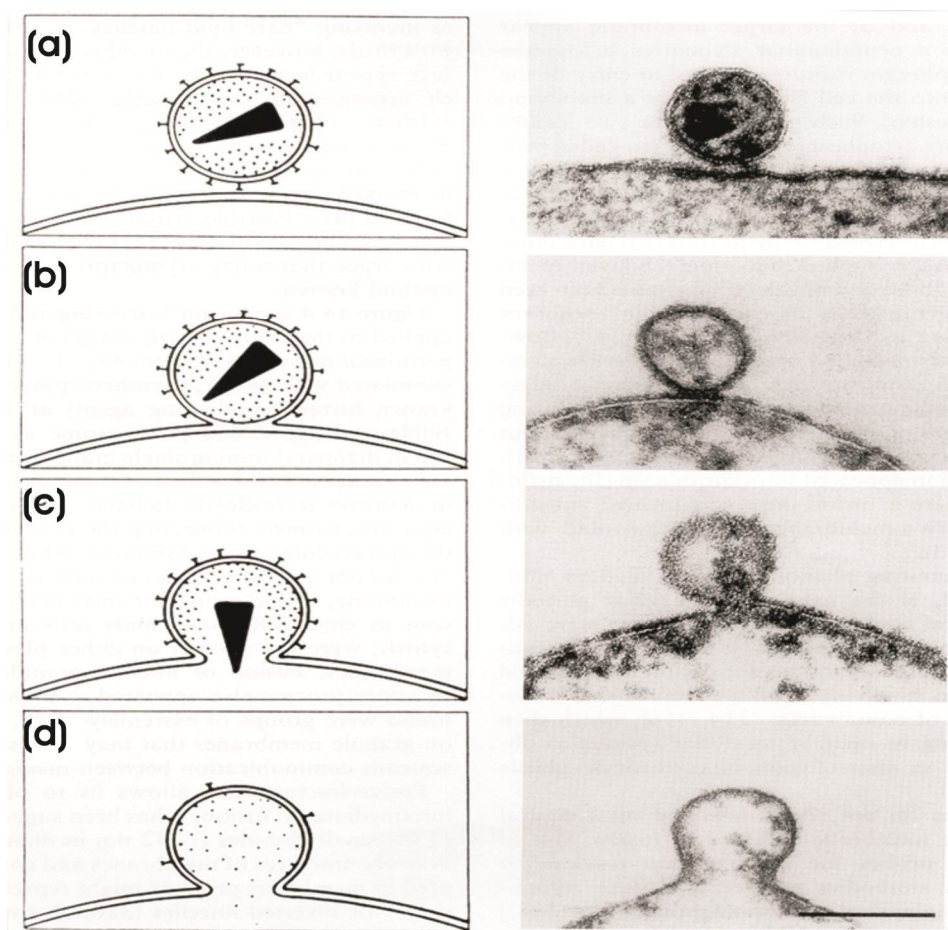


Figure 1. HIV infection model (left) and freeze fracture electron microscopy (right) of (a) binding (b) hemi-fused viral and host cell membranes (c, d) pore formation with HIV infection of the host cell (modified from literature⁸).

HIV is an enveloped virus, and its membrane is derived from its infected host cells. HIV entry and infection are initiated by the noncovalently associated glycoproteins gp120 and gp41 where gp120 is located on the exterior of the transmembrane protein gp41^{1,9}. Membrane fusion is initiated after gp120 is bound to a CD4 receptor and an additional co-receptor from the chemokine family of a host cell, and binding to these receptors results in removal of gp120 from gp41¹⁰⁻¹². The previously “covered” gp41 is exposed to aqueous solution and is thought to undergo a series of structural transitions required for infection (see below)¹³. The gp41 protein is composed of ~356 residues¹⁴ and is subdivided into regions from the N-terminus: fusion peptide (FP) (~16 residues), FP proximal region (FPPR) (~13 residues), N-terminal helix (NTH) (~40 residues), loop (~47 residues), C-terminal helix (CTH) (~37 residues), pre-transmembrane region (~18 residues), transmembrane region (~28 residues)¹⁵, and cytoplasmic endodomain (~160 residues)¹⁴. The ~175-residue N-terminal ectodomain of gp41 lies outside the virus, and X-ray crystal and liquid-state nuclear magnetic resonance (LSNMR) structures have shown organized molecular trimers for constructs that lacked the fusion peptide, transmembrane and endodomain regions^{1,15-19}. These ectodomain crystal structures showed protein trimers with three interior parallel α helical NTH segments and three exterior α helical CTH segments packed antiparallel to the NTHs. The overall structure of each monomer was a hairpin, and the trimer formed a six-helix bundle^{15-17,20}. The above domains have been defined by crystal structures of gp41 based constructs with varying interpretations of the number of residues incorporated into each domain. To my knowledge, the largest HIV gp41 crystallized construct to date has shown that the helicity extends beyond the traditionally defined NTH and CTH, **Figure 2**. The helicity of the NTH and

CTH approximately extends from the respective residues Ala-532 to Ile-580 and Asp-627 to Asn-677. Additionally, the NTH and CTH in the SIV gp41 crystal structure span residues Arg-30 to Ala-86 and Thr-104 to Lys-146¹⁸, respectively, which is analogous to HIV gp41 crystal structure residues Arg-542 to Ser-598 and Ser-616 to Glu-662. One interpretation of these combined results is that under crystallization conditions, the helicity of the NTH and CTH regions are terminated due to the length of the gp41 construct rather than the length of the NTH and CTH of the full length gp41. Of note, these gp41 structures are for gp41 without the presence of a membrane, and these structures also lack the hydrophobic fusion peptide.

For HIV gp41, mutations within the FP and FPPR have been shown to inhibit membrane fusion which suggests that both the FP and FPPR are important for membrane fusion²¹. Of special interest, transdominant inhibition of the V2E mutated gp41 (a FP mutation) has demonstrated that more than three gp41 or multiple gp41 trimers are needed to initiate membrane fusion²¹. Whatever the structure/s of the FP and FPPR are, the structure/s must allow for aggregation of FPs between more than 3 gp41 or multiple gp41 trimers. While the structures of the FP region of gp41 are the focal point of this dissertation, it should also be noted that regions that are C-terminal of FP can also be effective targets for fusion inhibitor drugs. The AIDS drug Enfuvirtide²² is a fusion inhibitor and is a 36-residue peptide containing parts of the C-helix and the pre-transmembrane regions. Enfuvirtide likely binds to exposed N-helical regions in pre-hairpin intermediate (PHI) gp41 and acts as a competitive inhibitor to the native C-helix regions with consequent prevention of the transition to the final hairpin structure. Cell-cell fusion mediated by gp41 includes in sequence: (1) lipid mixing between the membranes; (2) fusion pore formation; and (3) pore enlargement. After addition of Enfuvirtide, small pores rather than large

pores can be closed which indicates that gp41 in the PHI state mediates lipid mixing and initial fusion pore formation while the final hairpin state (or possibly the PHI → hairpin transition) mediates pore stabilization and enlargement^{13,22,23}. While Enfuvirtide is generally effective and possesses minimal side effects, Enfuvirtide is not commonly prescribed because it is a peptide that requires two daily injections²², and it is not cost-efficient. Continued development of HIV inhibitory drugs is highly valuable to reduce the costs of current HIV treatments²⁴ and enhance the accessibility to treatment of people in all social classes worldwide.



Figure 2. A summary of the gp41 sequence and regions defined from literature where FP = fusion peptide, FPPR = fusion peptide proximal region, NTH = N-terminal helix, CTH = C-terminal helix, MPER = membrane proximal region, and TMR = transmembrane region. This dissertation primarily focuses on the role of FP in membrane fusion. The ~16 residue FP sequence followed by C-terminal amino acids (HFP) was studied in this dissertation. This figure was adapted from literature¹⁵. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

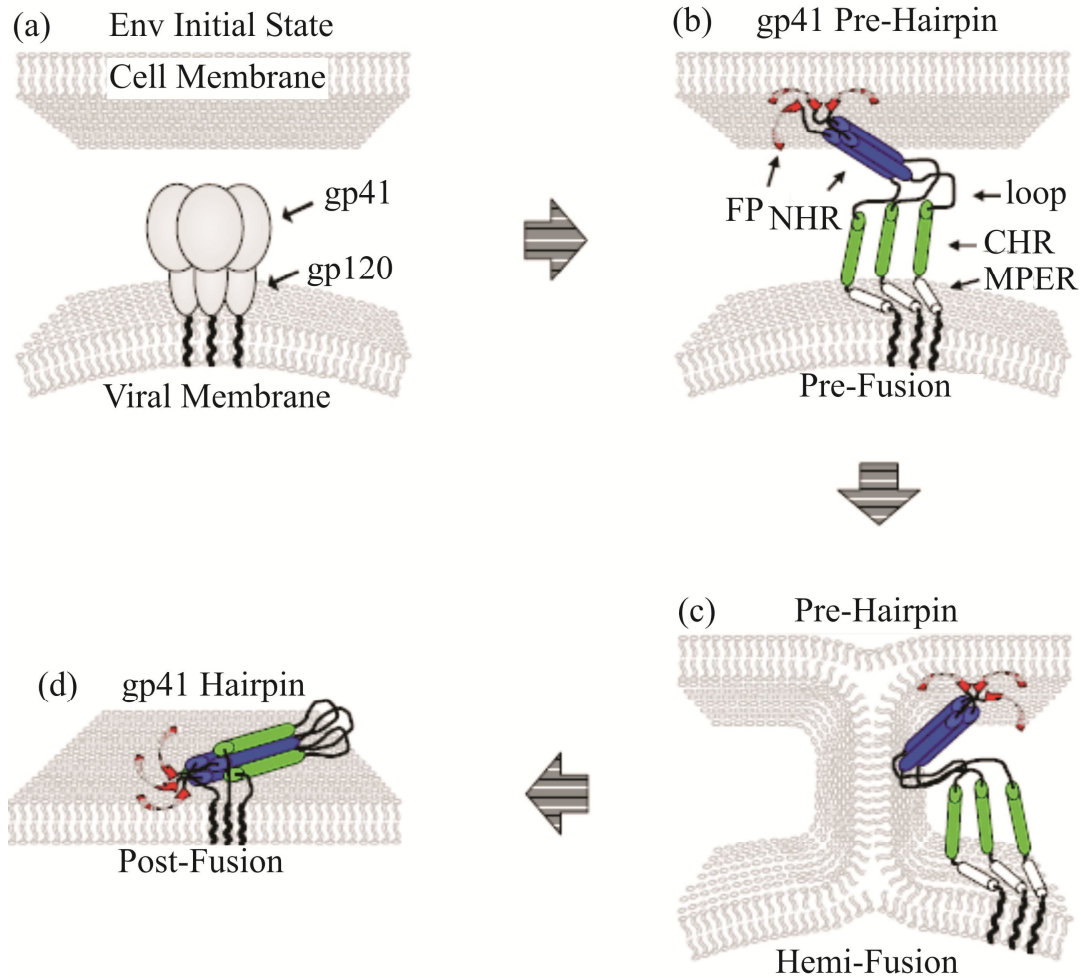


Figure 3. Proposed fusion mechanism (Figure provided by Dr. Kelly Sackett). This model for membrane fusion suggests that sequential events occur during viral and host cell membrane fusion: (A \rightarrow B) The ectodomain of gp41 is fusion inactive prior to contact with the host cell membrane. Contact between the host cell and viral membrane is initiated when gp120 binds to CD4 and a co-receptor¹². This induces a conformational change in gp41 to the proposed pre-hairpin intermediate structure and FP binds to the host cell membrane. To my knowledge, there is currently no direct evidence that FP binds to the host cell, but synthetic fusion peptides bind and insert into model membranes^{25,26}. Additionally, no structural data have been collected to elucidate the proposed PHI structure; (B \rightarrow C) The host cell and viral membrane undergo Hemi-Fusion; (C \rightarrow D) Prior to completion of membrane fusion, a fusion pore is formed within the host cell membrane. HFP constructs have been shown to insert into model membranes²⁶ which appears to be driven by local hydrophobicity of HFP secondary and tertiary structures (See Chapter IV). Similarly, FP may insert into host cell membranes which may enable FP to be involved in pore formation. The membrane fusion pore is enlarged after or during the formation of hairpin structure²⁷. Additionally, peptides with the CTH sequence inhibit membrane fusion by a proposed mechanism where these peptides bind to the NTHs in the Pre-Hairpin Intermediate structure (B and C) which prevents the formation of the 6 helix bundle structure (D). This lead to development of the membrane fusion inhibitor drug Enfuvirtide²².

The FP is a potential target for inhibitory drug design because it is a highly conserved region that is essential for HIV infection. While the FP has numerous sequential variations in different strains of HIV, the FP is composed of hydrophobic amino acids mixed with alanine and glycine residues where the AVGIGALFLGFLGAAG sequence was used in this study. The FP is essential for HIV infection which has been demonstrated by *in vivo* fusion and infection studies where mutations within the FP region of gp41 inhibited fusion and infection^{21,28}. The ~16 residue FP sequence followed by C-terminal amino acids are generally referred to as HFP constructs. HFP constructs induce membrane vesicle fusion in the absence of other regions of gp41 or other proteins^{25,29,30}. There are good correlations between the effects on gp41-mediated fusion by specific FP mutations and the effects on vesicle fusion by the corresponding mutations in HFP^{21,31}. These similar trends support investigation of HFP and elucidation of its membrane-associated structure to better understand the role of FP in fusion and as a fusion inhibitor target.

The HFP structure-function literature includes NMR data showing random coil structure for HFP in aqueous solution^{32,33}. A fluorescence and infrared (IR) study further elucidated HFP structure and reported the time-resolved courses of HFP structural changes and the intervesicle lipid mixing following addition of a HFP solution to a membrane vesicle solution³⁴. The experimental rates (R_s) were ordered $R_{HFP \text{ membrane binding}} > R_{HFP \beta \text{ sheet formation}} \sim R_{lipid \text{ mixing}}$ and were consistent with the sequence: (1) random coil HFPs bind to a membrane vesicle; and (2) HFP form oligomeric β sheets and vesicle fusion occurs. Similar structures were observed by solid-state NMR (SSNMR): (1) HFP lyophilized from aqueous solution without vesicles had a distribution of secondary structures as indicated by single site backbone ¹³CO

signals whose ~8 ppm linewidths spanned typical helical and β strand chemical shifts³⁵; and (2) HFP bound to hydrated membranes containing ~30 mol% cholesterol (i.e. the approximate cholesterol content in the membranes of HIV and its host cells³⁶⁻³⁹) formed oligomers/aggregates with β sheet structure^{35,36}. The biological relevance of HFP oligomers is further supported by the molecular trimer structure of soluble regions of the gp41 ectodomain¹⁷⁻¹⁹. The region approximately between residues Thr-25 and Gly-85 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 Thr-25. A C-terminally cross-linked HFP trimer (HFPtr) was therefore synthesized to mimic the close proximity of the three Thr-25 residues in the hairpin structure. Relative to HFP monomer, HFPtr induced membrane vesicle fusion at >15 fold faster rate which supported the functional significance of the trimer^{25,30}. Although both HFP and HFPtr formed β sheet oligomers in membranes with cholesterol, HFPtr is more deeply inserted which correlates with greater membrane perturbation and reduction of the vesicle fusion activation energy⁴⁰. 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^{21,41}. Analyses of these data supported the involvement of multiple gp41 trimers and fusion peptides in fusion⁴² and a model for this dominant inhibition is provided in **Appendix III**. Electron micrographs of virus-cell contacts have also been interpreted

to show multiple gp41 trimers at the contact site⁴³, **Figure 4**. Functional importance of fusion peptide trimers has also been demonstrated for fusion peptides of other viruses^{44,45}.

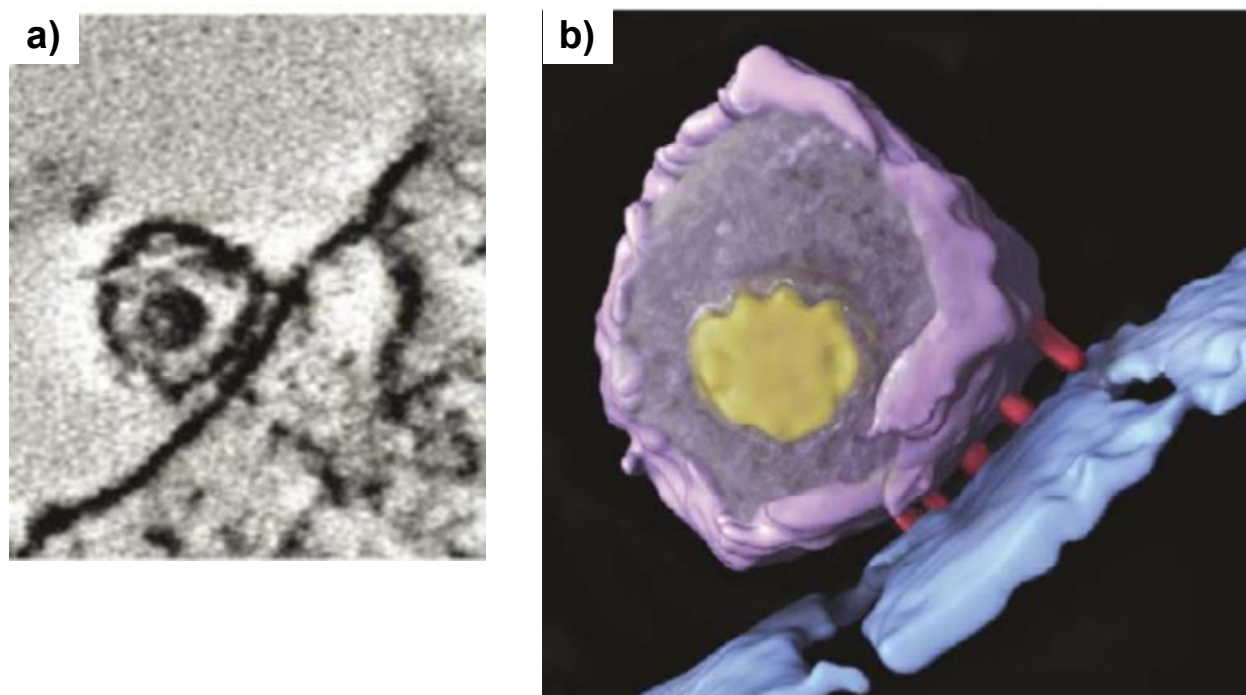


Figure 4. (a) Electron micrograph of HIV and a T cell where the width of the HIV membrane is ~100 nm. The increased electron density between the HIV and host cell membranes suggests that multiple proteins from the HIV membrane are in contact with the T cell membrane. (b) Model of HIV contact with a T cell based upon electron micrographs where multiple proteins are believed to be between the HIV and T cell membranes. This figure was modified from literature⁴⁶.

Because of the aforementioned functional significance of HIV fusion peptide oligomers, there has been effort to elucidate the membrane-associated HFP structure(s). Efforts to identify the predominant β sheet registry of membrane-associated HFP have lead to conflicting β sheet structural models such as in-register parallel⁴⁷, mixed in-register parallel and antiparallel²⁹, and β hairpin⁴⁸. However, the data from these studies were not interpreted quantitatively which left room for interpretational ambiguity that, in my opinion, explains the inconsistency of the

interpretations of these experiments. The membrane-associated HFP samples (mHFP) that are presented in this dissertation were prepared in a manner similar to that of fusion assays with fusion peptide in aqueous solution added to a membrane vesicle solution³⁷. Previous experiments have shown that 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^{30,33,49}. Additionally, SSNMR experiments are capable of detecting structure in membrane bilayers and have greatly contributed to the mHFP literature that has inspired the work of this dissertation. HFP/lipid binding has been supported by SSNMR detection of a HFP Ala-1 ¹³CO(carbonyl)-lipid ³¹P distance of ~5 Å⁴⁰ and the Ala-1 ¹³CO-lipid ³¹P contact as well as other data suggest that the number of molecules in the oligomer is small^{40,50,51}. Intermolecular ¹³C-¹³C and ¹³C-¹⁵N distances of ~4-5 Å have been detected in mHFP and supported β sheet oligomers/aggregates. Additionally, ¹³C chemical shifts of residues between Ala-1 and Gly-16 in mHFP were consistent with a fully extended β strand conformation⁵⁰.

This dissertation has contributed to the β sheet structure literature by quantitative determination of populations of specific β sheet registries. Prior to my work, the clearest information to-date on this topic has been a SSNMR experiment on membrane-associated HFP with an Ala-14 ¹³CO label and a Gly-3 ¹⁵N label whose separation (r_{CN}) was >20 Å along a single β strand⁵⁰. SSNMR can detect labeled ¹³CO-¹⁵N dipolar coupling (d_{CN}) where $d_{CN} = 3080/r_{CN}^3$ with d in Hz and r in Å. The minimum detectable $d_{CN} \sim 10$ Hz correlates with $r_{CN} \sim 7$

Å so that detectable d_{CN} in this sample were necessarily ascribed to inter- rather than intramolecular ^{13}CO - ^{15}N proximity since labeled intrastrand ^{13}CO - ^{15}N distances were > 7 Å for residues in β sheet structure. SSNMR detection of $d > 30$ Hz strongly supported a significant fraction of molecules with intermolecular Ala-14-Gly-3 hydrogen bonding and labeled r_{CN} of ~ 4.1 and ~ 5.5 Å, i.e. $16 \rightarrow 1/1 \rightarrow 16$ antiparallel β sheet registry, **Figure 5**. In general, antiparallel registries will be described using a registry index t where the $t = 16$ registry describes the $16 \rightarrow 1/1 \rightarrow 16$ antiparallel β sheet registry. Detection of similarly large d_{CN} in an Ala-14 ^{13}CO /Ile-4 ^{15}N HFP sample also supported a fraction of the $t = 17$ registry, **Figure 5**. At most half of the membrane-associated HFP molecules were in the $t = 16$ and $t = 17$ registries, i.e. a large fraction of the molecules were in registries not detected in either the Ala-14 ^{13}CO /Gly-3 ^{15}N or Ala-14 ^{13}CO /Ile-4 ^{15}N labeled samples. Because of the close proximity of the Thr-25 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. $1 \rightarrow 17/1 \rightarrow 17$ in **Figure 5**. 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}CO labels and the other with three sequential backbone ^{15}N labels²⁹. Detection of an average $d_{CN} > 10$ Hz for a Gly-5-Leu-7 ^{13}CO /Gly-5-Leu-7 ^{15}N sample and a Phe-11-Gly-13 ^{13}CO /Phe-11-Gly-13 ^{15}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}C - ^{13}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}CO label and in-register parallel structure, the labeled interstrand $r_{CC} \sim 5$ Å with $d_{CC} \sim 70$ Hz^{52,53}. These parameters will be independent of the residue that is ^{13}CO labeled. For mHFP with Phe-8 ^{13}CO , a best-fit $d_{CC} \sim 70$ Hz was detected whereas for mHFPtr, d_{CC} depended on the position of the labeled ^{13}CO residue with a range of 10-60 Hz^{54,55}. This residue dependence argued against a major fraction of in-register parallel structure in HFPtr.

$t = 16$	$t = 17$	In-Register Parallel
AV <u>G</u> IGALFLGFLG <u>A</u> AGS	AVG <u>I</u> GALFLGFLG <u>A</u> AGS	AVGIGALFLGFL <u>L</u> GAA <u>G</u> S
SGA <u>A</u> GLFGLFLAGI <u>G</u> VA	SGA <u>A</u> GLFGLFLAGI <u>G</u> VA	AVGIGALFLGFL <u>L</u> GAA <u>G</u> S

Figure 5. The $t = 16$, $t = 17$, and in-register parallel β sheet registries are illustrated. Isotopic labeling schemes to detect these registries are illustrated where red underlined atoms indicate ^{13}CO labeled residues and blue underlined atoms indicate ^{15}N labeled residues. These specific labeling schemes are effective because they have ^{13}CO - ^{15}N proximities of ~ 4 Å.

There was also an IR spectroscopy effort to distinguish between the $1 \rightarrow 17/1 \rightarrow 17$ parallel and the $t = 16$ registries using samples that contained backbone ^{13}CO labeling at either: (1) Ala-1 to Val-3, Gly-5 to Ile-9; (2) Phe-8 to Gly-16; or (3) Aal-1 to Val-3, Gly-5 to Gly-16⁴⁷. The IR wavenumbers and intensities of different samples were interpreted to support a large fraction of in-register parallel structure and little antiparallel structure. However, in my view, the extensive labeling of the IR samples precluded quantitation of specific registries and greater support for this argument is provided in Chapter III. These in-register parallel studies motivated the work for

Chapter III where more accurate quantification of the natural abundance dephasing (*nad*) in HCN rotational-echo double-resonance (REDOR) experiments was achieved. This established a more quantitative method for detection of the fraction of parallel structure in membrane-associated HFP oligomers. The lack of in-register parallel β sheets motivated the work described in Chapter IV where the complete distribution of antiparallel β sheet registries was quantified in mHFP, and the potential functional significance is discussed. The registry distribution was also quantified in the less active V2E mutant HFP (V2E-HFP). Clear differences were evident in the registry distribution between mHFP and mV2E-HFP. These observed structural differences were the basis for experimentally based structure-function models that are presented in Chapter IV.

Chapter II. Materials and Methods

2.1 HFP Synthesis and Sample Preparation

1. General HFP Preparation

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 methods⁵⁶. Standard Fmoc chemistry⁵⁷ was used to synthesize the gp41 fusion peptide (HFP) with the 23 N-terminal residues (AVGIGALFLGFLGAAGSTMGARS). A WK₆A or WK₆A β tag was added to the native sequence to for the following purposes: (1) non-native W24 was incorporated as an A₂₈₀ chromophore; (2) non-native lysines were added to reduce HFP aggregation in aqueous solution prior to membrane binding^{30,33}. This ensured that membrane-associated β sheet oligomers/aggregates were formed after membrane binding; (3) Wang Resin was preloaded with Ala or β -Ala (~0.27-0.74 mmol/g). Alternatively, Wang Resin preloaded with Lys would have likely been effective for HFP synthesis.

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 (RP-HPLC) generally with a C₄ column and a water-acetonitrile gradient containing 0.1% TFA (See **Appendix IV** for setup). Acetonitrile, TFA, and residual solvents were removed with nitrogen gas and subsequent lyophilization. HFP purity was >95% as

determined by mass spectrometry. HFP amounts were quantified using A_{280} of aqueous solutions of HFP with $\varepsilon = 5600 \text{ M}^{-1} \text{ cm}^{-1}$.

The membrane composition of membrane-associated samples 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 HIV⁵⁸. Ether- rather than more physiologically abundant ester-linked lipids were used because the latter have two COs/molecule that would contribute substantial natural abundance (na) ^{13}C signal. Bilayer phase is retained for ether-linked lipids with cholesterol and with HFP⁵⁹⁻⁶¹. In addition, membrane-associated HFP has predominant β sheet structure in either ester-linked lipid + cholesterol or ether-linked lipid + cholesterol compositions⁵⁴.

Samples were prepared by first dissolving DTPC (20-40 μmol), DTPG (5-10 μmol), and cholesterol (12.5-25 μmol) in chloroform (lipid and cholesterol amounts depended upon the amount of peptide used) and the chloroform was removed 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% NaN_3 preservative. The suspension was homogenized with ten freeze-thaw cycles and large unilamellar vesicles were formed by extrusion through a polycarbonate filter with 100 nm diameter pores (Avestin, Ottawa, ON). 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 supernatant³⁷. The pellet was lyophilized, transferred to the SSNMR rotor, and rehydrated with $\sim 30 \mu\text{L}$ of 5 mM

HEPES buffer at pH 7.0 for every 50 μmol of lipid⁶². The validity of the lyophilization/rehydration approach was supported by peak ^{13}C O chemical shifts that were within 0.6 ppm of those of samples that were not lyophilized^{50,63}.

Table 1. HFP construct labeling schemes.

Peptide	Labeled residues	Peptide	Labeled residues
HFP-F8	Phe-8 ^{13}C O	HFP-F8CA21N	Phe-8 ^{13}C O and Ala-21 ^{15}N
HFP-L12	Leu-12 ^{13}C O	V2E-A6CG3N	Ala-6 ^{13}C O and Gly-3 ^{15}N
HFP-G5A6	Gly-5 and Ala-6 ^{15}N	V2E-L7CG3N	Leu-7 ^{13}C O and Gly-3 ^{15}N
HFP-A6L7	Ala-6 and Leu-7 ^{15}N	V2E-F8CG3N	Phe-8 ^{13}C O and Gly-3 ^{15}N
HFP-L9G10	Leu-9 and Gly-10 ^{15}N	V2E-L9CG3N	Leu-9 ^{13}C O and Gly-3 ^{15}N
HFP-L12G13	Leu-12 and Gly-13 ^{15}N	V2E-L9CI4N	Leu-9 ^{13}C O and Ile-4 ^{15}N
HFP-G13A14	Gly-13 and Ala-14 ^{15}N	V2E-L9CG5N	Leu-9 ^{13}C O and Gly-5 ^{15}N
HFP-A6CG3N	Ala-6 ^{13}C O and Gly-3 ^{15}N	V2E-L12CG3N	Leu-12 ^{13}C O and Gly-3 ^{15}N
HFP-L7CG3N	Leu-7 ^{13}C O and Gly-3 ^{15}N	V2E-L12CI4N	Leu-12 ^{13}C O and Ile-4 ^{15}N
HFP-F8CG3N	Phe-8 ^{13}C O and Gly-3 ^{15}N	V2E-L12CG5N	Leu-12 ^{13}C O and Gly-5 ^{15}N
HFP-L9CG3N	Leu-9 ^{13}C O and Gly-3 ^{15}N	V2E-L12CA6N	Leu-12 ^{13}C O and Ala-6 ^{15}N
HFP-L9CI4N	Leu-9 ^{13}C O and Ile-4 ^{15}N	V2E-L12CL7N	Leu-12 ^{13}C O and Leu-7 ^{15}N
HFP-L9CG5N	Leu-9 ^{13}C O and Gly-5 ^{15}N	V2E-F8CL12N	Phe-8 ^{13}C O and Leu-12 ^{15}N
HFP-L12CG3N	Leu-12 ^{13}C O and Gly-3 ^{15}N	V2E-F8CG13N	Phe-8 ^{13}C O and Gly-13 ^{15}N
HFP-L12CI4N	Leu-12 ^{13}C O and Ile-4 ^{15}N	V2E-F8CA14N	Phe-8 ^{13}C O and Ala-14 ^{15}N
HFP-L12CG5N	Leu-12 ^{13}C O and Gly-5 ^{15}N	V2E-F8CA15N	Phe-8 ^{13}C O and Ala-15 ^{15}N
HFP-L12CA6N	Leu-12 ^{13}C O and Ala-6 ^{15}N	V2E-F8CG16N	Phe-8 ^{13}C O and Gly-16 ^{15}N
HFP-L12CL7N	Leu-12 ^{13}C O and Leu-7 ^{15}N	V2E-L9CG16N	Leu-9 ^{13}C O and Gly-16 ^{15}N
HFP-F8CL12N	Phe-8 ^{13}C O and Leu-12 ^{15}N	L9R-F8CG13N	Phe-8 ^{13}C O and Gly-13 ^{15}N
HFP-F8CG13N	Phe-8 ^{13}C O and Gly-13 ^{15}N	HFPtr-F8CG13N	Phe-8 ^{13}C O and Gly-13 ^{15}N
HFP-F8CA14N	Phe-8 ^{13}C O and Ala-14 ^{15}N	V2E-L12CA6N	Leu-12 ^{13}C O and Ala-6 ^{15}N
HFP-F8CA15N	Phe-8 ^{13}C O and Ala-15 ^{15}N	V2E-L12CL7N	Leu-12 ^{13}C O and Leu-7 ^{15}N
HFP-F8CG16N	Phe-8 ^{13}C O and Gly-16 ^{15}N	V2E-F8CL12N	Phe-8 ^{13}C O and Leu-12 ^{15}N
HFP-L9CG16N	Leu-9 ^{13}C O and Gly-16 ^{15}N	V2E-F8CG13N	Phe-8 ^{13}C O and Gly-13 ^{15}N

2. HFP Solutions

Chapter III. The samples that were prepared to detect in-register parallel and antiparallel β sheets (HFP-P, HFP-A, HFP-AP) used HFP solutions that were prepared with ^{13}C O-labeled HFP (3.0 mg) and ^{15}N labeled HFP (6.0 mg) in HEPES buffer (32 mL) before it was added drop wise to the ~2 mL lipid solution.

Chapter III. The negative control sample (HFP-NC) 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 ^{13}CO s and ^{15}N s remained much farther apart than 7 Å which is the approximate REDOR detection limit.

All other HFP constructs. The HFP had a single ^{13}CO label and a single ^{15}N label. HFP (4.0-6.5 mg) was suspended in HEPES buffer (32 mL) before it was added dropwise to the ~2 mL lipid solution.

3. Trimeric HFP (HFPtr) Synthesis

Literature methods²⁵ were used to synthesize the dimeric and trimeric peptides. Cross linking reactions between two peptides, denoted A and B, are listed as well as the product in **Table 2**. The HFPdm(Lys) (0.10 mmol) was synthesized by adding Fmoc-Lys(Mtt)-OH, Fmoc-Lys-OH with a Mtt sidechain protecting group, at position Lys-30. After addition of the Lys-30 residue, the resin was washed and “capped” with acetic anhydride solution. The Mtt sidechain protecting group was cleaved with an acid solution (5 mL of dichloromethane (DCM) with 1% v/v TFA) using 6 x 6 minute cycles. Fmoc-Cys(Trt)-OH was added to the unprotected Lys sidechain. Standard Fmoc chemistry was used to synthesize the rest of HFPdm(Lys) where HFPdm(Lys) contained HFPWK₅**K**A and HFPWK₅**C** peptides that were cross-linked at the K and C residues indicated in bold font.

The HFPdm(Cys) was synthesized by dimerization of the HFPWK₅CA peptide. Formation of the disulfide bond between peptides works well at a pH of ~8, and the solutions can

be adjusted by addition of acid and base. However, when the pH of the solution becomes too acidic, the peptides become insoluble and precipitate out of solution. Precipitated HFPWK₅CA peptides are difficult to dissolve. To avoid this problem, I changed the DMAP concentration from 10 mM (literature values²⁵) to 20 mM. In a 15 mL conical vial, 2.5 mmol of HFPWK₅CA was dissolved in 490 μ L of 20 mM DMAP solution and gently vortexed in air. After approximately 2 hours, 100 μ L of DMAP solution was used to rinse the sides of the 15 mL conical vial and the mixture was then vortexed in air overnight. The sample was additionally sonicated every 30 minutes for the first 3 hours to redissolve any precipitated peptide. The effect of sonication on reaction yield was not measured. The pH was ~8 under these conditions and precipitation of the peptide was not observed. The reaction yielded 5.7 mg of HFPdm(Cys) from 7.7 mg of HFPWK₅CA after purification (~74% yield). The same peptide and DMAP concentration was used for the cross-linking reaction of HFPWK₅CA and HFPdm(Lys) to yield HFPtr. The ratio of HFPWK₅CA:HFPdm(Lys) peptides should affect the molar ratio of reaction products (HFPWK₅CA:HFPdm(Lys):HFPdm(Cys):HFPtr:HFPte)²⁵ where HFPte is tetrameric HFP. A 2:1 mol ratio of HFPWK₅CA to HFPdm(Lys) was used and 13.2 mg HFPWK₅CA was reacted with 13.2 mg of HFPdm(Lys). This reaction yielded 4.4 mg of HFPdm(Lys)+HFPdm(Cys), 5.0 mg of HFPtr, and 2.9 mg of HFPte after HPLC purification. This reaction was only run once, and a representative HPLC chromatogram is displayed in **Figure 6**. HFPtr was used in Chapter IV while HFPdm data are displayed in **Appendix V**.

Table 2. Oligomeric HFP synthesis summary.

Peptide A	Peptide B	Product
HFPWK ₅ CA	None	HFPWK ₅ CA
HFPWK ₅ CA	HFPWK ₅ CA	HFPdm(Cys)
HFPdm(Lys)	None	HFPdm(Lys)
HFPdm(Lys)	HFPWK ₅ CA	HFPtr
HFPdm(Lys)	HFPdm(Lys)	HFPte

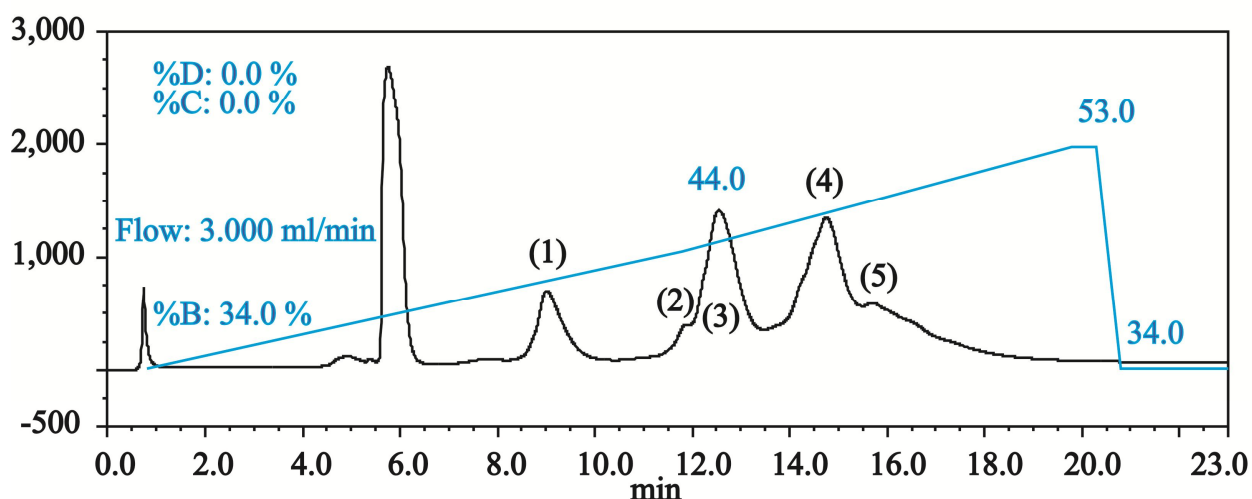


Figure 6. Purification of HFPtr synthesis. The product of each peak was identified by MALDI-TOF mass spectrometry, **Figure 7**, and each peak corresponded to the following: (1) HFPWK₅CA; (2) HFPdm(Lys) and HFPdm(Cys); (3) HFPdm(Lys) and HFPdm(Cys); (4) HFPtr; and (5) HFPte. Additionally, peaks (4) and (5) were not present in HFPdm(Cys) which further supports that these peaks result from HFPtr and/or HFPte. Peak (5) was not used for any experiments in this dissertation while Peak (4) was collected for the one HFPtr experiment in Chapter IV. The identity of HFPte would have to be confirmed before meaningful experimental data is collected for this peptide.

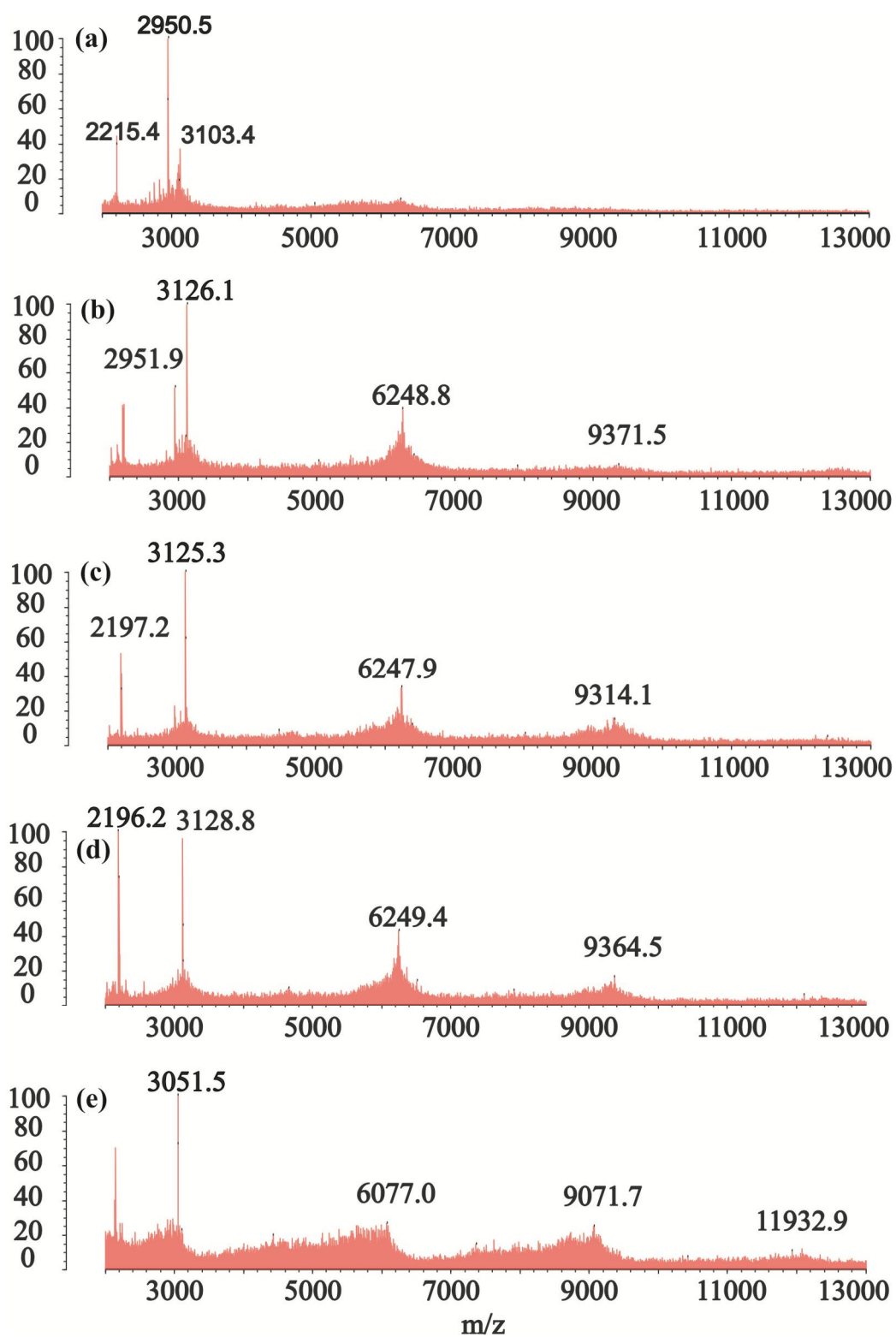


Figure 7. MALDI-TOF mass spectroscopy from the HFPtr purification displayed in **Figure 6**.

Figure 7 (cont'd). The fractions collected from the numbered peaks from HPLC are correlated to the chromatograms and peptides (expected mass) as follows: (1)-(a)-HFPWK₅CA (3126 +2 g/mol); (2)-(b)-HPFdm (6250 +4 g/mol) and HFPdm(Cys) (6257 +4 g/mol); (3)-(c)-HPFdm (6250 +4 g/mol) and HFPdm(Cys) (6257 +4 g/mol); (4)-(d)-HFPtr (9381 +6 g/mol); and (5)-(e)-HFPte (12512 +8 g/mol) where an ¹⁵N or ¹³CO label adds +1 g/mol each. In (a)-(e), expected product peaks and expected product peaks missing an Ala residue are observed. The missing Ala residue(s) is likely due to laser induced C-terminal fragmentation or hydrolytic cleavage⁶⁴ rather than synthetic impurities since acetic anhydride was used to “cap” failed syntheses. The mass of a failed synthesis should correspond to a peak with mass equal to the failed synthesis’s native sequence plus 43 (acyl group). Although not likely here, fragmentation of C-terminal of amino acids with basic sidechain groups can occur^{65,66}. Additionally, fragmentation at a disulfide bond may occur⁶⁷, and peptides may contain multiple charges. Either scenario could result in detection of a mass to charge ratio that is a fraction of the expected mass to charge ratio of the dimer, trimer, or tetramer product peaks.

2.2 General NMR Background

1. NMR Spectroscopy

NMR experiments have an external magnetic field aligned along the z-axis of the laboratory frame where **B₀** represents external magnetic field. For this chapter, bold characters are used to represent vector quantities while non-bold characters are used to represent scalars. NMR experiments collect nuclear resolvable data where the vector quantity of the nuclear magnetic moment, **μ**, for an atom can be described below:

$$\mathbf{\mu} = \gamma \mathbf{I} \quad (1)$$

where γ is the gyromagnetic ratio of the nucleus and **I** is the spin angular momentum of the nucleus. **I** can be written in terms of scalars and unit vectors by:

$$\mathbf{I} = I_x \hat{x} + I_y \hat{y} + I_z \hat{z} \quad (2)$$

When placed in an external magnetic field, **B₀**, **μ** has a potential energy, *V*.

$$V = -\mathbf{\mu} \cdot \mathbf{B}_0 = -\gamma \mathbf{I} \cdot \mathbf{B}_0 \quad (3)$$

where

$$\mathbf{B}_0 = B_x \hat{x} + B_y \hat{y} + B_z \hat{z}$$

Because external magnetic fields are typically aligned in the z direction, **Eq. (3)** can be rewritten.

$$V = -\gamma I_z B_z \quad (4)$$

By replacing I_z with the spin operator \hat{I}_z , the Schrodinger equation for a nuclear spin can be written using the nuclear spin Hamiltonian.

$$\hat{H}\psi = -\gamma B_z \hat{I}_z \psi = E\psi \quad (5)$$

The selection rules in NMR spectroscopy demonstrate that the wave functions are spin eigenfunctions in that $\hat{I}_z \psi = \hbar m_I \psi$ where $m_I = I, I-1, \dots, -I$. Therefore, from **Eq. (5)**, energy can be calculated as:

$$E = -\gamma \hbar m_I B_z \quad (6)$$

Nuclei have $2I+1$ energy levels, and $m_I = \pm 1/2$ nuclei, such as ^1H and ^{13}C , have two energy levels are separated by ΔE .

$$\Delta E = \gamma \hbar B_z \quad (7)$$

The two energy levels correspond to the $m_I = +1/2$ and $m_I = -1/2$. The equilibrium populations of these spins in a magnetic field are described by a Boltzmann distribution. In NMR experiments, an ensemble of spins are observed and both the $m_I = +1/2$ and $m_I = -1/2$ spin states are populated at equilibrium in the presence of \mathbf{B}_0 where $m_I = +1/2$ is the lower energy state.

2. Rotating Frame and Bloch Equations

Simplistically speaking, NMR spectroscopy measurements are made after the magnetization is transferred from a z-axis orientation to the xy-plane. In general, NMR experiments detect coherence of the nuclear magnetic moments as magnetization in the xy-plane. A nuclear magnetic moment can be represented by the vector quantity, μ , and the vector sum of the nuclear magnetic moments is the magnetization, \mathbf{M} .

$$\mathbf{M} = \sum_{a=1}^n \mu_a \quad (8)$$

where a is an index for nuclear magnetic moments being detected and n represents the number of nuclear magnetic moments detected in a NMR sample. As discussed in **2.4 MAS Solid-State NMR (REDOR)**, structural information can be obtained by observing decoherence of the nuclear magnetic moments which results in a smaller \mathbf{M} . Before discussing coherence and decoherence of the nuclear magnetic moments, it is essential to first understand how nuclear magnetic moments or magnetization interact with the external magnetic field, \mathbf{B}_0 , and radiative magnetic field, \mathbf{B}_1 . First, consider the interaction of \mathbf{M} with a magnetic field, \mathbf{B} , by the Bloch equation:

$$\frac{d\mathbf{M}}{dt} = \mathbf{M} \times (\gamma \mathbf{B}) \quad (9)$$

where \mathbf{B} is the vector sum of the external magnetic field, \mathbf{B}_0 , and the radiative magnetic field \mathbf{B}_1 .

Before considering the interactions in pulsed experiments, it is important to expand upon **Eq. (9)**.

Consider the specific example below:

$$\left| \frac{d\mathbf{M}}{dt} \right| = |\mathbf{M}_z \times \gamma \mathbf{B}_x| = \gamma |\mathbf{M}_z| |\mathbf{B}_x| \sin\phi \quad (10)$$

Eq. (10) evaluates the magnitude of a vector quantity where φ is the angle between the z and x axes. In general, directional dependence for evaluating the direction of a cross product can be determined by using the right hand rule where $\mathbf{M}_z \times \mathbf{B}_x$ would have a y direction.

Before introducing the concept of a pulsed experiment, it is important to understand the interaction of a dipole moment with the external magnetic field. In this brief discussion, consider a single dipole moment as depicted in **Figure 8**. Due to the interaction with the external magnetic field, \mathbf{B}_0 , μ precesses about the z axis with an angular frequency of ω_0 , the Larmor frequency.

$$\omega_0 = -\gamma B_0 \quad (11)$$

As depicted in **Figure 8**, a rotating frame with the Cartesian coordinates $(\cos(\omega t), \sin(\omega t), z)$, where z has a constant value, can be thought to precess at the same frequency as the Larmor frequency only if $\omega_1 = \omega_0$. A rotating frame of reference is created by imagining \mathbf{x}' and \mathbf{y}' axes rotating about the z axis. This is a useful concept since B_1 fields are applied in the xy plane at the transmitter frequency ω_1 . The B_1 field precesses at the same frequency as the rotating \mathbf{x}' and \mathbf{y}' axes, and the \mathbf{x}' and \mathbf{y}' axes precess at the Larmor frequency only when $\omega_0 = \omega_1$. The rotating frame (\mathbf{x}' and \mathbf{y}' plane) is defined by ω_1 because signal detection is performed relative to the transmitter frequency. Thus, the B_1 field can arbitrarily be thought to be applied along the \mathbf{x}' axis. This rotates \mathbf{M}_z about the \mathbf{x}' axis in the $\mathbf{z}-\mathbf{y}'$ plane. Applying a 90° pulse along the \mathbf{x}' axis is equivalent to the mathematical expression $\mathbf{M}_z \times \mathbf{B}_{x'}$, which reorients the magnetization in the \mathbf{y}' direction where the resultant magnetization is denoted $\mathbf{M}_{y'}$. These concepts build a foundation

for understanding how to set up a pulsed NMR experiment and also establish a foundation for understanding coherence of the nuclear magnetic moments in the rotating frame as discussed in section 2.4 MAS Solid-State NMR (REDOR).

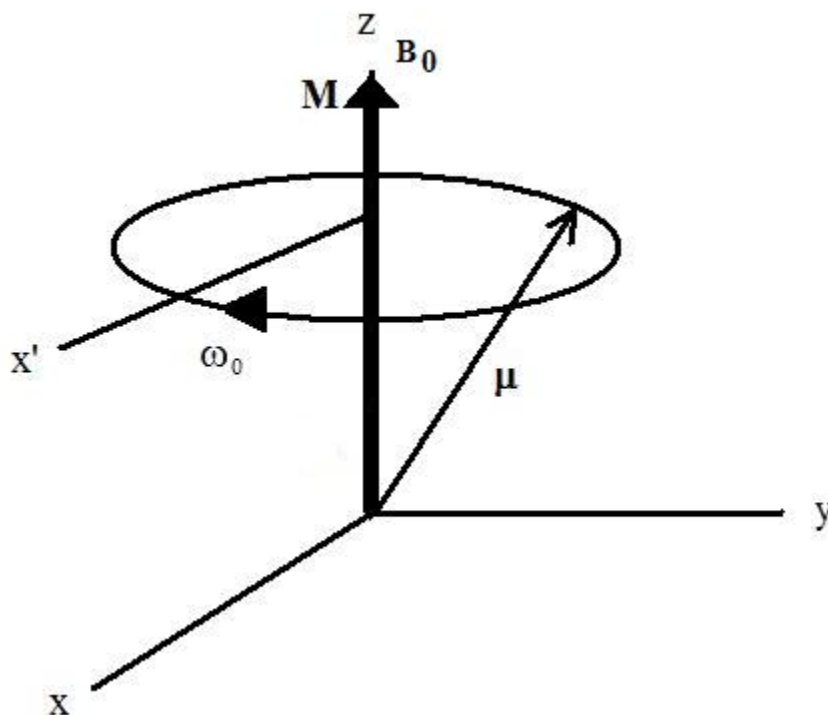


Figure 8. A dipole moment in the laboratory frame (x , y , and z axes) where the B_0 external magnetic field is aligned along the z -axis. A new axis, x' , rotates at the Larmor frequency only if $\omega_0 = \omega_1$.

2.3 Experimental Setup

1. General Concepts for Setting the Pulses

Many NMR experiments begin by applying a $\pi/2$ pulse using an external magnetic field, B_1 , to change the magnetization from the z -orientation to the xy plane. In a classical picture and for simplicity, consider the x' , y' , and z' axes to be the axes of the rotating frame where the axes rotate at the transmitter frequency, ω_1 . In a complex system with many nuclear magnetic moments, a sample placed in an external magnetic field has a sum of the nuclear magnetic

moments that yield an observable net magnetization oriented along the **z**-axis, \mathbf{M}_z . In NMR spectroscopy, the transverse magnetization (magnetization in the **xy** plane) is detected in the time domain and is then typically transformed to the frequency domain by a Fourier transformation. In order to detect the magnetization, \mathbf{M}_z must be reoriented to the **xy** plane of the laboratory frame. For simplicity, first consider the scenario where the $\omega_0 = \omega_1$ (i.e. the precession frequency of \mathbf{B}_0 = the precession frequency of \mathbf{B}_1). Because \mathbf{B}_1 is always stationary in the rotating frame, we can arbitrary assign \mathbf{B}_1 to the “ \mathbf{B}_x ,” direction. If the pulse duration of \mathbf{B}_1 is short relative to both T_1 and T_2 ($\sim < 50 \mu\text{s}$), we can ignore the T_1 and T_2 effects on the angle that magnetization rotates at the time when \mathbf{B}_1 is applied. T_1 and T_2 represent the longitudinal (**z**-axis) and transverse (**xy**-plane) magnetic relaxation time constants. Typical relaxation times for membrane-associated peptides studied in this dissertation were on the millisecond timescale for T_2 ’s and seconds time scale for T_1 ’s. Additionally, recall that in an NMR experiment for $m_z = \frac{1}{2}$ nuclei, we are inducing a population change in the spin up and spin down states and the energy for this transition can be denoted ΔE , Eq (7). The energy for this transition is supplied by the \mathbf{B}_1 radiative field and is related to power, P , and the time, τ , for which the power is applied by Eq (12).

$$E = \int_0^{\tau} P(\tau) d\tau \quad (12)$$

The power, $P(\tau)$, can be related to voltage $V(\tau)$ and current $I(\tau)$ by Eq (13).

$$P(\tau) = V(\tau) \bullet I(\tau) \quad (13)$$

Therefore, from **Eq (12)** and **Eq (13)**,

$$E = \int_0^{\tau} V(\tau) \bullet I(\tau) d\tau \quad (14)$$

Experimentally, we can measure the average voltage over a period of time using an oscilloscope.

In NMR the Rabi frequency can be thought of as the frequency of oscillation between the spin up and spin down states due to the applied magnetic field, B_1 , where

$$\omega_R = \gamma B_1 \quad (15)$$

In the rotating frame, ϑ is the angle that the magnetization rotates about the axis of the radiative field as written below:

$$\vartheta = \omega_R \tau_p \quad (16)$$

where ϑ is expressed in radians, τ_p is the pulse length used to achieved a ϑ pulse, and the interaction between the magnetization and the radiative field was previously described by **Eq (9)**. This is a general convention that is used so that a simple expression can be written for calculating the Rabi frequency for a pulse where the Rabi frequency for a 2π pulse can be calculated by **Eq 19**.

$$\frac{1}{\tau_{2\pi}} = \frac{\omega_R}{2\pi} \quad (17)$$

Additionally, the $\omega_R/2\pi$ can be calculated for a $\pi/2$ pulse by **Eq (18)** assuming that a $\pi/2$ pulse takes 1/4 times as long as a 2π pulse.

$$\frac{\omega_R}{2\pi} = \frac{1}{4(\tau_{\pi/2})} \quad (18)$$

where $\tau_{\pi/2}$ is the time required to achieve a $\pi/2$ pulse. By varying the strength of the \mathbf{B}_1 radiative field, there are many variations of ω_R and τ_p that can be used to achieve a $\pi/2$ pulse, but different

combinations of w_R and τ_p used to rotate θ to a fixed angle, such as $\pi/2$, affect the bandwidth of excitation where the bandwidth of excitation is approximately proportional to the inverse of the pulse duration⁶⁸, **Figure 10**. Shorter $\pi/2$ pulses have a larger bandwidth of excitation since there is greater uncertainty associated with the approximation of using a square wave to represent a $\pi/2$ pulse. In a ^{13}C NMR experiment where ^{13}CO nuclei are the primary focus of data analysis, it may be less important to have a large bandwidth of excitation since the range of ^{13}CO nuclei resonant frequencies are narrow (anisotropic chemical shift range of ~ 140 ppm^{69,70} or ~ 14 kHz) relative to the spectral width (50 kHz). However, shorter ^{13}C π pulses are still important for REDOR experiments. By Eq (13)-(16), shorter pulses that require more power (i.e. higher voltage measured on the oscilloscope) excite a broader range of resonant frequencies, and the measured voltage required to achieve a $\pi/2$ pulse is inversely proportional to the τ_p . Typical pulse lengths used in the HCN REDOR experiments are 5 μs for $\pi/2$ pulses and 10 μs for π pulses for the ^1H and ^{13}CO respective nuclei while longer pulses (~ 20 -25 μs) were typically used for ^{15}N (see **9. Proton TPPM Decoupling to Nitrogen Rabi Frequency Ratio in HCN REDOR**). As displayed in **Figure 10** the largest amplitude of excitation is at the transmitter frequency, w_m , but it is important not to set the transmitter frequency to the frequency of the observed nuclei, generally ^{13}CO in HCN REDOR experiments. Detection is performed in the rotating frame which is at the transmitter frequency of the laboratory frame. This commonly leads to detection of the DC voltage at the transmitter frequency which results in a nonzero “baseline” of the free induction decay (FID), **Figure 9**, and a signal at 0 Hz in a Fourier

transformed spectrum. The DC offset correction can be applied prior to Fourier transforming the FID so that the baseline of the FID is correctly offset to zero. However, this processing method is not perfect and often does not entirely eliminate the “blip” at 0 Hz. Thus, it is best to move the transmitter frequency $\sim 15\text{-}20$ ppm ($\sim 1.5\text{-}2$ kHz) when detecting ^{13}CO nuclei on a 400 MHz spectrometer) away from the observed nuclei’s resonant frequency so that the “blip” does not overlap and interfere with the ^{13}CO signal. Additionally, it is important that the observed nuclei have resonant frequencies that are near the maximum excitation energy. This is important since the energy due to the radiative field changes less near transmitter frequency as evident by **Figure 10** where the slope of curve is approximately equal to zero at the transmitter frequency, ω_1 .

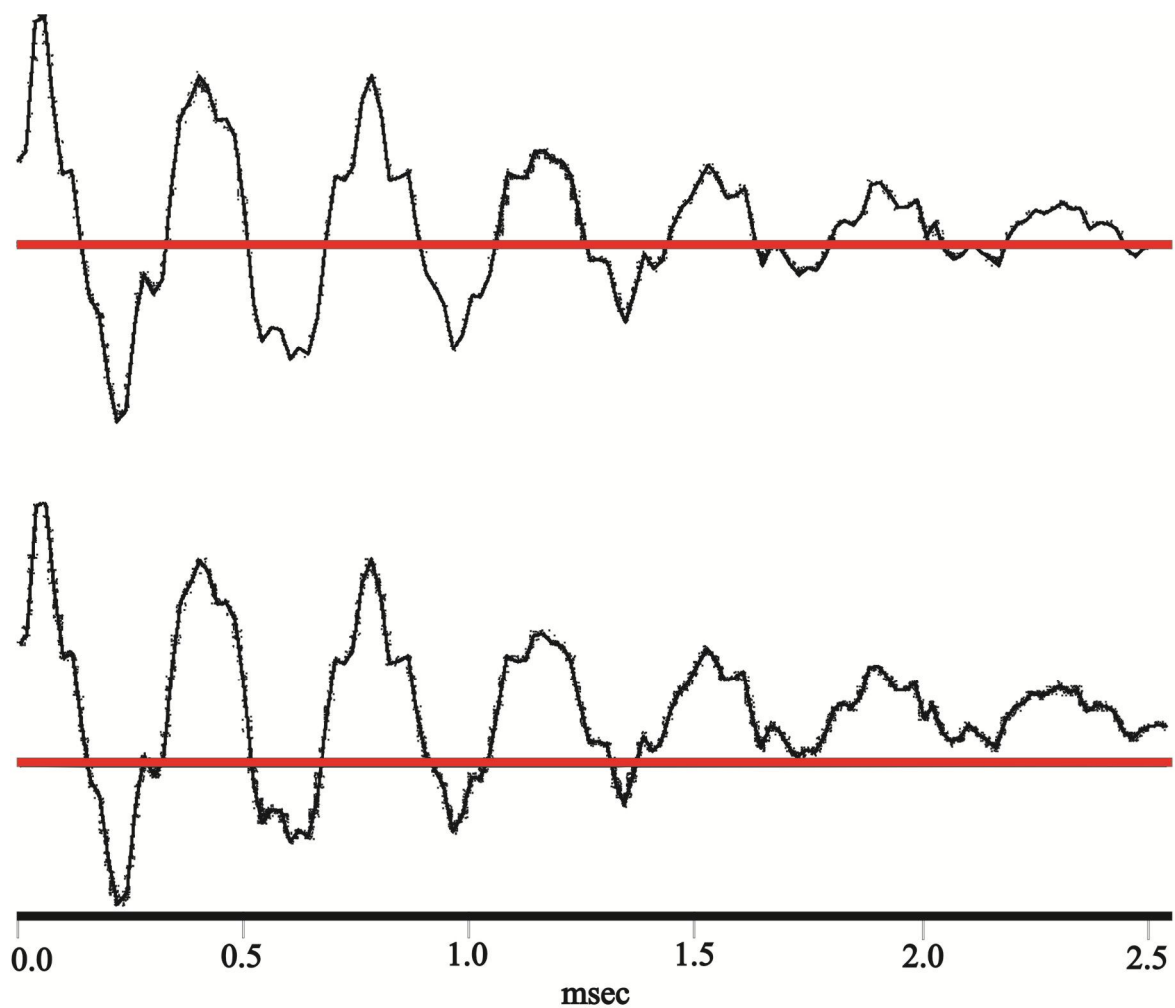


Figure 9. FID's for cross polarization acquisitions of the ^{13}C labeled i4 peptide where the red horizontal line signifies the zero value or baseline for the y-axis. (a) The FID decays to the baseline. (b) The FID decays to a positive y-value as evident by the zero point value of the FID (~2.5 ms) being above the red line.

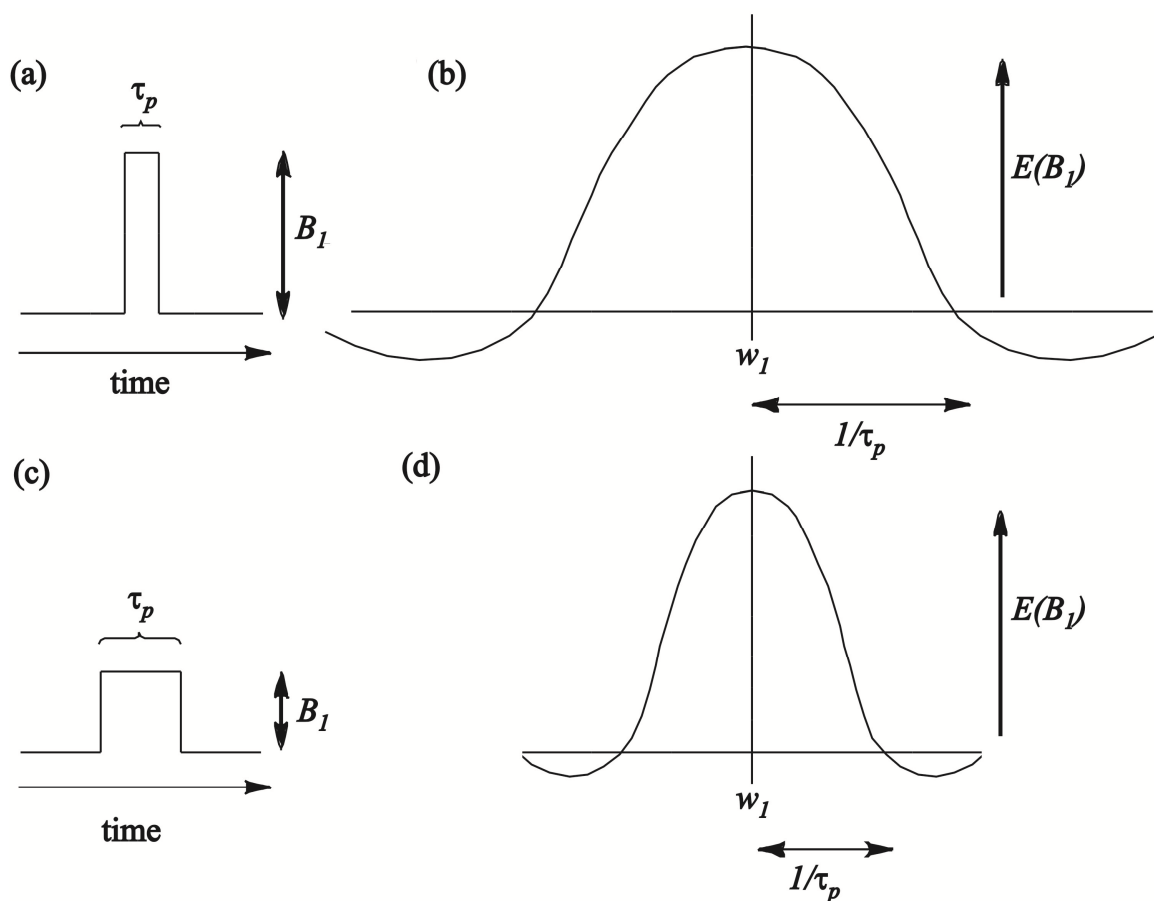


Figure 10. A pulse of a single radiofrequency is illustrated in the time domain (a,c) and the frequency domain (b,d). The relationship between pulse length (τ_p), radiative field (B_1), and the energy of the B_1 field are illustrated. The excitation frequency is centered about the transmitter frequency, w_1 . The figure was made using concepts from the literature where the frequency domain is represented as the Fourier transform of the time domain where the Fourier transform of a step function can be represented by a sinc function (b), (d)⁶⁸.

2. Setup Compounds

I4 peptide. A 17-residue acetylated and amidated peptide with the sequence Ac-AEAAAKEAAAKEAAKA-NH₂, and I4 had a ¹³CO label at Ala-9 and a ¹⁵N label at Ala-13.

The peptide was lyophilized from aqueous solution and is predominantly α helical ($83 \pm 6\%$) at Ala-9⁷¹. These labeled nuclei should have a ¹³C-¹⁵N internuclear distance of 4.1 Å in an α helix.

Adamantane. Adamantane was used for chemical shift referencing (see below). The chemical structure of adamantane is much different than the chemical structure of backbone ^{13}C O in HFP, and adamantane should not be used for optimizing cross polarization (CP) parameters for the REDOR pulse sequence (see below).

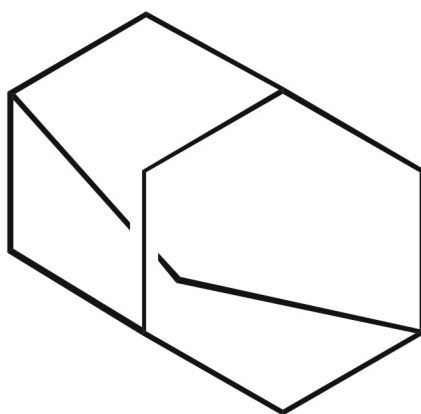


Figure 11. The chemical structure of adamantane.

3. Setting the Magic Angle

In the Fourier transformed spectrum, quadrupolar nuclei have many spinning sidebands where the intensity of the sidebands is sensitive to the magic angle, and the magic angle is 54.7° ⁷². Thus, quadrupolar nuclei such as ^{79}Br can be used to set the magic angle. We currently have a rotor packed with adamantane and KBr. Additionally, it is convenient to detect ^{79}Br when setting the magic angle because (1) Detection of ^{79}Br in KBr does not require ^1H decoupling which reduces the power applied to the probe and allows for shorter pulse delays without damaging the probe; and (2) The ^{79}Br nuclei have a gyromagnetic ratio that is close to ^{13}C and therefore can easily be used without having to change hardware to detect ^{79}Br signal (i.e. tuning from ^{13}C to ^{79}Br can be accomplished only by adjusting tuning rods). In a simple one pulse

experiment (“1pulse” on the Infinity Plus spectrometers), a $\pi/2$ pulse is applied (say in the $-y$ direction) followed by acquisition. Consequently, the magnetization is transferred to the x axis and magic angle spinning (MAS) during acquisition results in no net evolution of the magnetization over each rotor period due to quadrupolar and chemical shift anisotropy (See **2.4 MAS Solid-State NMR (REDOR)** for brief discussion of anisotropy). At the end of each rotor period, the magnetization is realigned on the x -axis which results in a series of “spikes” in the FID that are separated by the spinning frequency. With a 4.0 kHz spinning frequency, the spikes typically extend out to 10-15 ms for ~64-128 scans acquired. It is not critical to set the pulse nutation angle exactly to 90° (see Ernst angle for more detail⁶⁸). Using KBr, signal averaging of ^{79}Br requires ~5-10 seconds to resolve spin echoes as observed in **Figure 12**, and it is sufficient to use the pulse length and amplifier input parameters that would be used for ^{13}C . Additionally, another helpful pdf file that demonstrates how to setup magic angle spinning from scratch can be found in the folder @..home/mb4b/data/Scott/REDOR/Setup/KBr_Magic_angle.pdf. When preliminarily setting the magic angle, the continuously updated 1-acquisition FID can be used to get near the magic angle whereas more scans should be acquired to finely adjust the magic angle

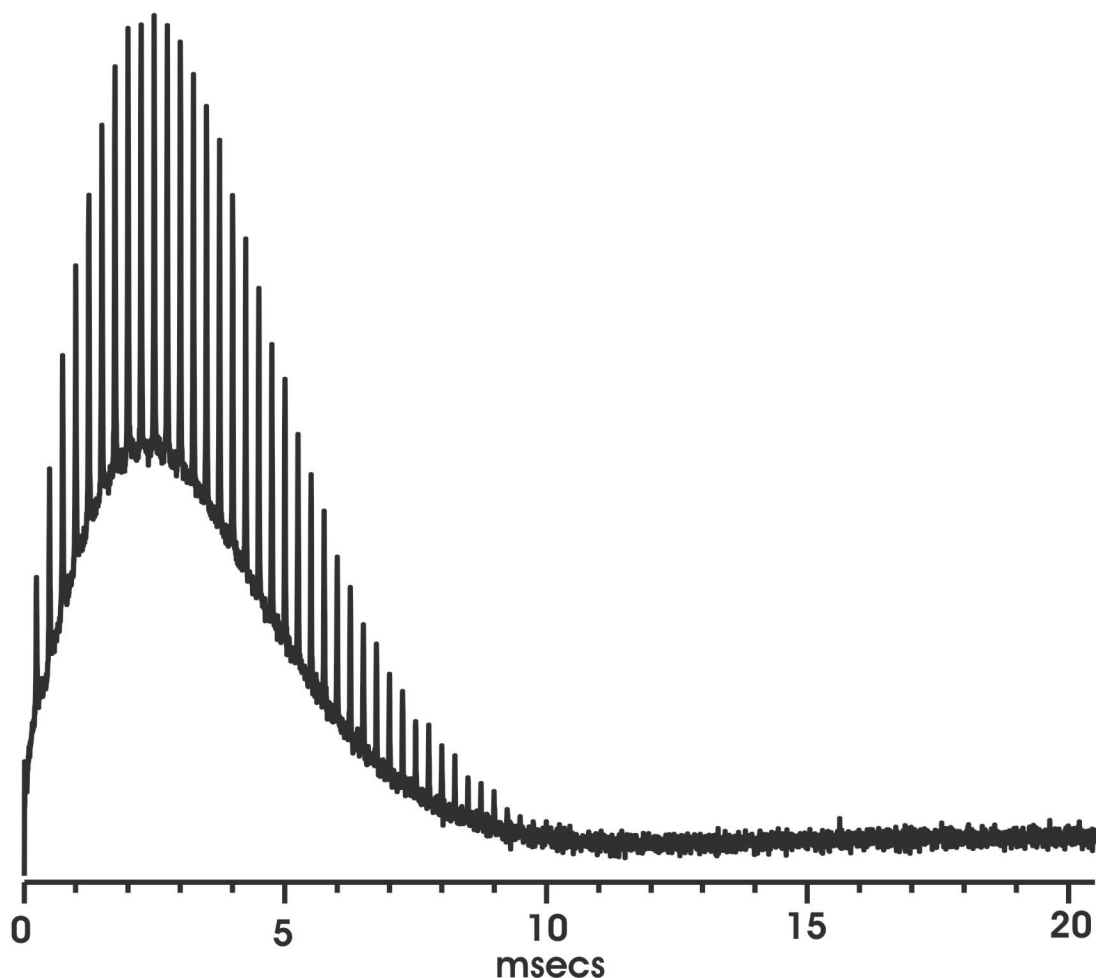


Figure 12. Magic angle spectrum FID of KBr with 64 acquisitions and a $pd = 0.10$ s. The transmitter was moved to the ^{79}Br resonant frequency of KBr, and an exponential decay is observed in the free induction decay (FID).

4. Setting Proton $\pi/2$ Pulse

When setting up experiments, equivalent parameter sets between experiments are essential to obtaining reproducible results. One way that we monitor the consistency between experiments is by measuring the Rabi frequencies, ω_R , of our pulses by **Eq. (16)**. In general, the pulses are used to induce a radiative transition between the spin up and spin down energy levels. As previously described, the time that the B_1 field is applied must be shorter than the relaxation

transition between the spin up and spin down states so that rotation of the magnetization is instantaneous relative the longitudinal relaxation time. In **Figure 17**, it can be observed that the data acquisitions with a lower decoupling field had lower signal to noise (i.e. larger error bars). For these points, this suggests that there was a larger observable ^1H - ^{13}CO coupling. As the decoupling field was increased, the proton decoupling Rabi frequency became greater than the ^1H - ^{13}CO dipolar coupling frequency, w_d , which results in a non-observable ^1H - ^{13}CO dipolar coupling interaction. A general equation for dipolar coupling frequency between two nuclei is listed below where “a” and “b” identify two nuclei.

$$w_d = \frac{\mu_0 \gamma_a \gamma_b \hbar}{8\pi^2 r_{ab}^3} \quad (19)$$

Additionally, the dipolar coupling energy is described by Eq. (20).

$$E = \frac{\mu_0 \gamma_a \gamma_b \hbar^2}{4\pi r_{ab}^3} (1 - 3 \cos^2 \phi) m_{aI} m_{bI} \quad (20)$$

where ϕ is the angle between \mathbf{B}_0 and the internuclear vector and μ_0 is the magnetic permeability.

The process of becoming an NMR spectroscopist is an ongoing learning experience. Because altering Rabi frequencies can potentially alter experimental data (see **9. Proton TPPM Decoupling to Nitrogen Rabi Frequency Ratio in HCN REDOR** for example), it is best to try to keep Rabi frequencies consistent and equivalent between experiments. From my experience as a graduate student, the power output/efficiency of the amplifiers/NMR circuitry can vary over time when users are continually altering experimental setup conditions such as probe soldering, ^1H amplifier tuning, and cable or capacitor swapping to name a few examples. Because of this, it is best to set the pulse length by varying the amplifier input parameter (see **Carbonyl π pulse**

below) than by varying the pulse length so that equivalent radiative fields are used for all experimental data sets that will be compared to each other. Variations in the amplifier input parameters between experiments result in no interpretational ambiguity between experimental data sets as long as equivalent Rabi frequencies are used. Alternatively, variation of pulse lengths used between experiments result in different Rabi frequencies for fixed θ , Eq. (16), that can yield different experimental data for a sample (see **9. Proton TPPM Decoupling to Nitrogen Rabi Frequency Ratio in HCN REDOR**). Alternatively, you can set the amplifier amplitude and vary the pulse length (see **4. Setting Proton $\pi/2$ Pulse**), but I would recommend doing the former since the Rabi frequencies affect simulated and experimental data while variation of the amplifier input parameters between experiments does not. Quite simply put, over time, there may be variation of the amplifier input parameters used to achieve an equivalent output voltage. Since the REDOR experiments are affected by Rabi frequencies, keeping the nutation angles and pulse lengths the same for all REDOR experiments will simplify comparisons between experimental data. Below, I have listed the setup protocols that will be helpful to new users. The protocols are listed in the order that I would generally perform setup experiments before running HCN REDOR experiments.

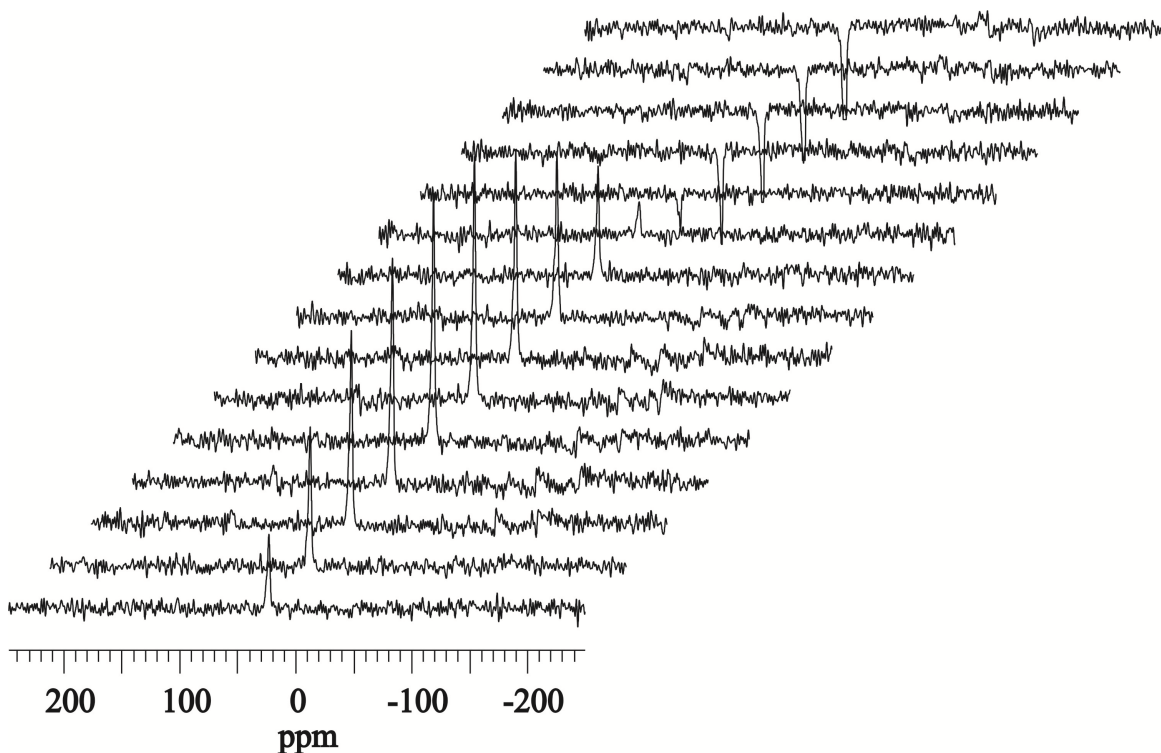


Figure 13. Proton pulse length arrayed using the I4 peptide, a $pd = 1.0$ sec and 10 acquisitions per spectrum. The H rf ampl parameter was set to 0.2300 and the H 90 pulse parameter was arrayed from 1.0 to 15.0 μs by 1.0 μs increments. Maximum signal was observed at $\sim 5 \mu s$ which approximately corresponds to a $\pi/2$ pulse. Zero signal should be observed between 10.0 and 11.0 μs which would correspond to a π pulse. The change in signal intensity is the greatest at points surrounding the π pulse so it is more accurate to set the $\pi/2$ 1H pulse length by identifying the π pulse length and dividing pulse length time by 2. To do this for the above data, the H 90 pulse parameter could be arrayed from 10.0 to 11.0 μs by increments of 0.1 μs and increase the number of scans per spectrum to enhance the signal to noise ratio. Alternatively, a better approach would be to set the H 90 pulse parameter to 10.0 μs and array the H rf ampl parameter. Upon determining which H rf ampl that yields zero signal, change the 10.0 μs to 5.0 μs to set the $\pi/2$ pulse. While the proton amplifier output voltage is approximately linear with respect to the H rf ampl parameter, it is not exactly linear so you need to divide the pulse length by 2 and not the amplifier input parameter.

5. Proton Decoupling Field

In the HCN REDOR experiments, TPPM decoupling⁷³ was used which is a windowless series of π pulses on the 1H channel as reflected by the input parameters TPPM aHdec dec. ampl and pw TPPM dcpl pulse. Proton decoupling prevents the loss of ^{13}CO M due ^{13}CO - 1H dipolar

couplings when the proton decoupling Rabi frequency is greater than the frequency of the ^{13}CO - ^1H dipolar couplings. However, in practice this is only an approximation for REDOR experiments. This statement is supported by **Figure 17**, where decoupling ω_R 's > 50 kHz did not improve the signal to noise for the 16.2 ms dephasing time points where as decoupling ω_R 's > 50 kHz provided better signal to noise per acquisition for the 48.2 ms dephasing time points. For the operational range of our probes (input ω_R 's < 95 kHz), higher decoupling fields reduce ^{13}CO transverse magnetization loss due to ^{13}CO - ^1H dipolar interactions which results in ^{13}CO signal enhancement. However, high decoupling fields can lead to sample heating and arcing of the probe. Arcing occurs when current travels across an unintended path (i.e. path of least impedance is not the intended path) which shorts the circuit and often results in excessive heating. Excessive heating or arcing points can generally be identified by spotting black soot in the LC circuit. The REDOR experiments work with the MAS probes with 4 mm diameter rotors with decoupling fields of 80-90 kHz. Currently, the proton amplifier output is approximately linear for the input parameters in Spinsight that yield ~ 40 -100 kHz ^1H decoupling fields (I have experimental data for one of the amplifiers in my 2nd NMR notebook on page 23 to support this statement while actual input parameter values depend upon the amplifier). The ^1H Rabi frequency of a pulse can be explicitly quantified by **Eq (16)**, and the ^1H decoupling energy can be quantified by measuring the forward voltage of the ^1H $\pi/2$ pulse and comparing it to the forward voltage applied for ^1H decoupling by **Eq (21)**.

$$\frac{wR1}{V_1} \propto \frac{wR2}{V_2} \quad (21)$$

where V represents the measured voltage going into the probe (use oscilloscope). The forward voltage for a ^1H $\pi/2$ pulse is easily measured if the H rf ampl is set equal to the aHcp H CP ampl. Otherwise, it is difficult to resolve the ^1H $\pi/2$ pulse voltage from the ^1H cross polarization voltage.

6. Proton-Carbon Cross Polarization

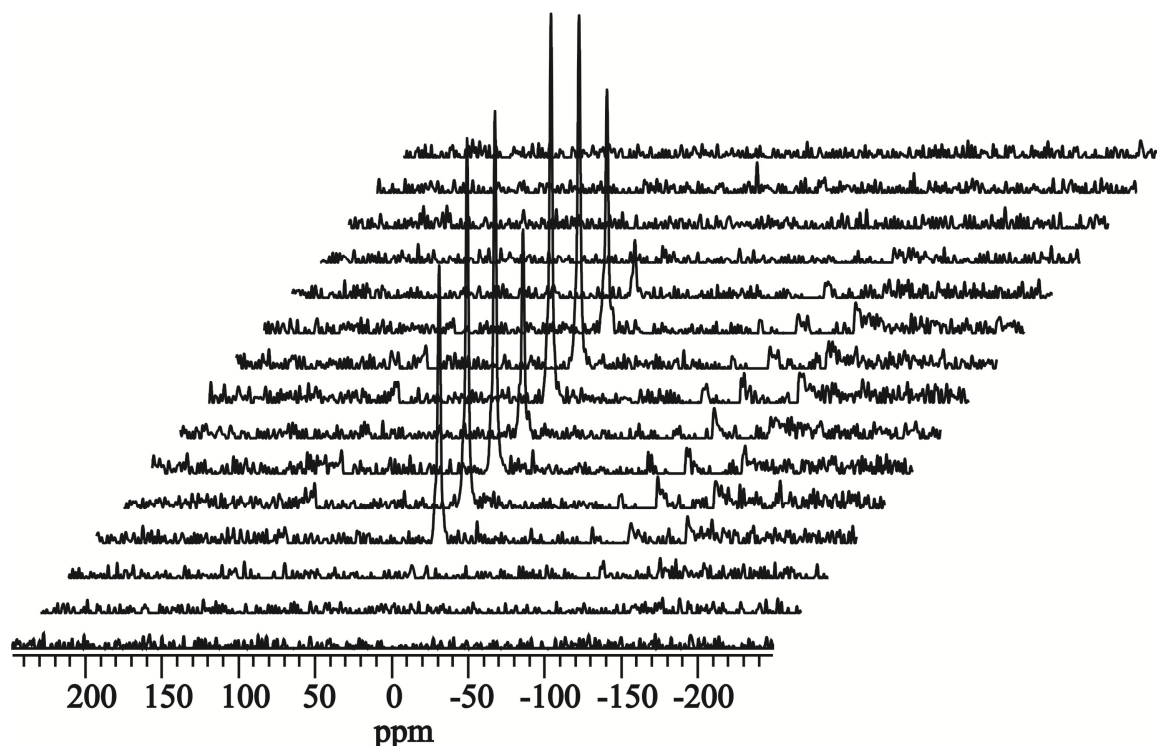


Figure 14. Cross polarization⁷⁴ arrayed with the I4 peptide using 5 acquisitions, 1.0 s pulse delay (pd), MAS frequency = 10 kHz, and contact time = 2 ms. Cross polarization is included in the REDOR experiments to transfer magnetization from the highly abundant and polarizable ^1H nuclei to the more dilute ^{13}C CO nuclei to increase the ^{13}C CO signal to noise ratio by increasing the signal per acquisition. Additionally, the longitudinal relaxation rate is approximately 4 times faster for ^1H than for ^{13}C nuclei in organic solids which allows for acquisition of ~4 times as many FIDs using cross polarization than for waiting for ^{13}C repolarization⁷⁵. The cp_ramp pulse program was used, and the X cp ampl change parameter was set to 0.04. A 48 kHz ^1H cross polarization and $\pi/2$ pulse was used, and the aXcp X cp ampl start parameter was arrayed from 0.00-0.70 by 0.05 increments. Maximum signal intensities were observed between 0.25-0.30 and 0.35-0.40. Either of these regions could be used for CP since the signal intensities are nearly equivalent. Additionally, the contact time also should be arrayed to set up an optimal CP, and it is important to use a setup compound with similar nuclear magnetic relaxation rates. Typical optimal contact times for peptides in REDOR experiments are 1.5-2.0 ms for ^{13}C CO nuclei. Finally, array the X cp ampl change (i.e. the ^{13}C ramp parameter) parameter to obtain the maximum signal where 0.04 is a typical value, but this value may vary depending upon the amplifier being used.

7. Carbonyl π Pulse

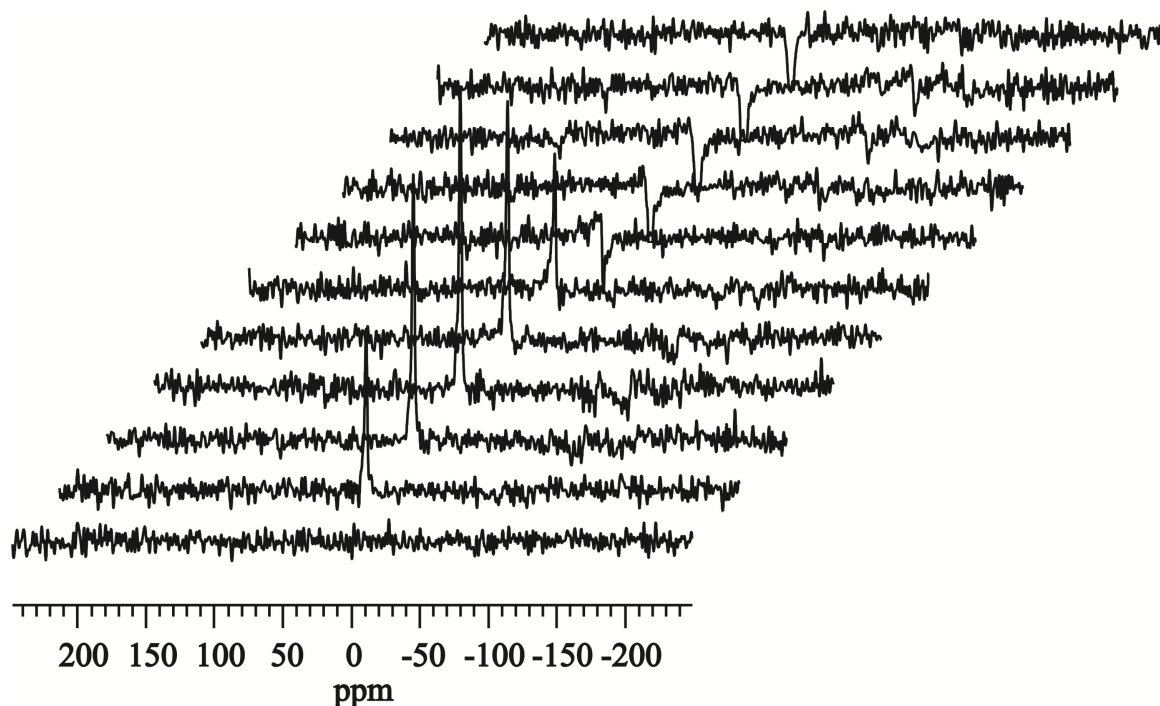


Figure 15. The cp_zfilter pulse sequence ($CP - \pi/2 - \tau_z - \pi$ - acquisition) was used to set the ^{13}C π pulse with 10 acquisitions and a 1.0 sec pulse delay using the i4 peptide and a MAS = 10 kHz. When the pulse lengths are set correctly, precession of the ^{13}C magnetization can be followed using the Bloch Equation, **Eq (9)** where (1) equilibrium, $M = +z$; (2) CP, rotates to the xy plane; (3) $\pi/2$, rotates to the $-z$ axis; (4) τ_z - remains along $-z$ axis; (4) π , flips back to $+z$ axis; (5) acquisition with consequent zero signal detection in the rotating frame when pulses are set accurately.

In **Figure 15**, the aX X 180 ampl parameter was arrayed from 0.00 – 0.50 by 0.05 increments, the pw90X X 90 pulse parameter was set to 5.50 μs , and the pw180X X 180 pulse set to 11.0 μs . The ^{13}C π pulse was set correctly when zero signal intensity was observed. As evident by **Figure 15**, the correct amplitude parameter for zero signal is between 0.25-0.30. To more accurately set correct ^{13}C pulses, array the aX X 180 ampl parameter from 0.25 to 0.30 by 0.01 increments and increase the number of scans to obtain better signal to noise.

8. Nitrogen π Pulse

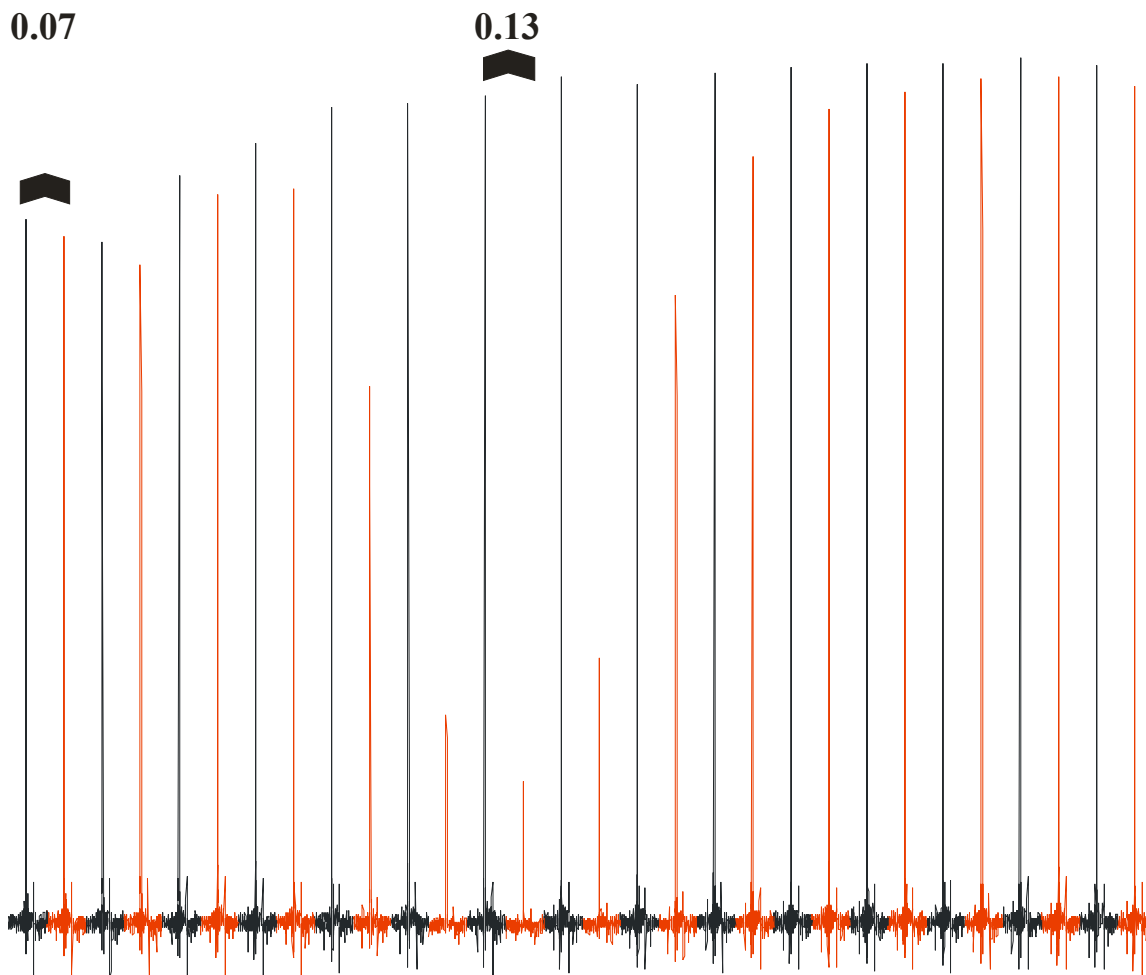


Figure 16. REDOR spectra (32.2 ms dephasing time) of the I4 peptide where each set of black and red spectra have a corresponding ^{13}C O peak in the S_0 and S_1 spectra for each arrayed points, respectively. Each S_0 and S_1 spectrum was the sum of 120 acquisitions with $\text{pd} = 2.0$ s. The aY180 Y 180 ampl parameter was arrayed from 0.07 to 0.21 by 0.01 increments and the pw180Y Y 180 parameter was set to 20.0 sec ($\omega_R = 25$ kHz as determined by Eq (13)-(16)). Maximum dephasing was observed with the pw 180Y Y 180 ampl = 0.13 where $\Delta S/S_0 = 0.82$ where maximum dephasing corresponds to a ^{15}N π pulse. For further understanding of the REDOR experiment, see **2.4 MAS Solid-State NMR (REDOR)**.

9. Proton TPPM Decoupling to Nitrogen Rabi Frequency Ratio in HCN REDOR

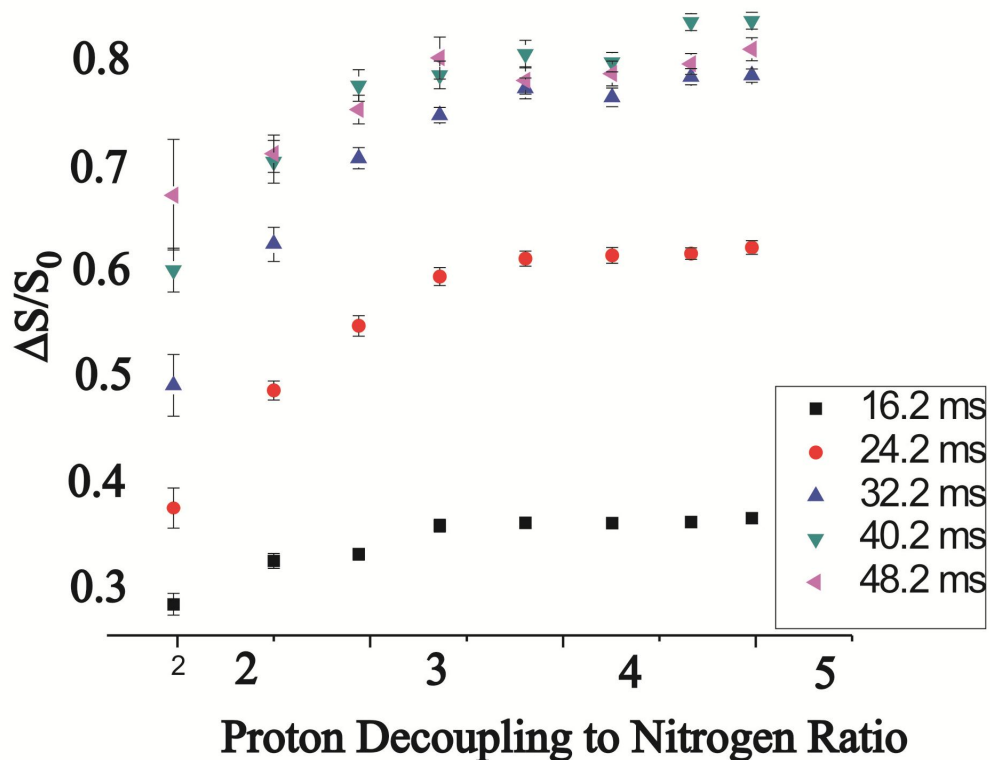


Figure 17. The $\Delta S/S_0$ of the i4 peptide is plotted against the ^1H decoupling w_R to ^{15}N w_R ratio for $\tau = 16.2, 24.2, 32.2, 40.2$, and 48.2 ms dephasing times where the ^{15}N π pulse was $25 \mu\text{sec}$ ($w_R = 20$ kHz), the ^{13}C π pulse was $11 \mu\text{sec}$ ($w_R = 45$ kHz), the $\text{pd} = 2.0$ s, and 350 scans were acquired for S_0 and S_1 of each data point. These HCN REDOR experiments appear to require a ^1H decoupling w_R to be at least 3.5 times greater than the ^{15}N w_R to obtain maximum $\Delta S/S_0$. The ^{13}CO nuclei were decoupled from ^1H nuclei at a ^1H decoupling to ^{13}C w_R ratios approximately ≥ 1.5 . This is evident since equivalent error bars were obtained for data within a dephasing time period for ^1H decoupling fields of ≥ 60 kHz. The ^{13}CO atoms are not directly bonded to ^1H atoms which results in weaker ^1H - ^{13}CO heteronuclear dipolar couplings (~ 3.8 kHz which was determined using Eq. (19) for a r_{HC} of 2.0 \AA between the carbonyl carbon and the adjacent residue's amide proton (See 1K09.pdb) relative to amide heteronuclear ^1H - ^{15}N dipolar couplings (11.648 kHz corresponds to a $r_{\text{HN}} = 1.015 \text{ \AA}$) which is approximately the width of the dipolar powder pattern for amide ^{15}N in proteins⁷⁶. Based upon the current set of experiments, it is unclear whether the ^1H decoupling to ^{15}N ratio is a causation or correlation relationship. There may be a ^1H - ^{15}N dipolar interaction that results in lower $\Delta S/S_0$ values where ^1H decoupling fields of >80 kHz may be required to average out effects due to ^1H - ^{15}N dipolar couplings.

Figure 17 (cont'd). However, the causation for the lower $\Delta S/S_0$ with lower ^1H decoupling w_R to ^{15}N w_R ratios is not understood. The SIMPSON simulations used in this dissertation did not include ^1H atoms so this affect was not simulated. Additionally, as evident by the larger error bars, less signal per scan was acquired for lower ^1H decoupling w_R to ^{15}N w_R ratios because ^1H decoupling varied while the ^{15}N w_R was held constant. The lower signal per scan implies that ^{13}C - ^1H dipolar coupling interactions were not averaged to zero. When the ^1H decoupling frequency is greater than the ^1H - ^{13}C dipolar coupling frequency, the effective or average ^1H spin quantum state is “0”. This results in no observable ^1H - ^{13}C CO dipolar couplings.

10. Chemical Shift Referencing

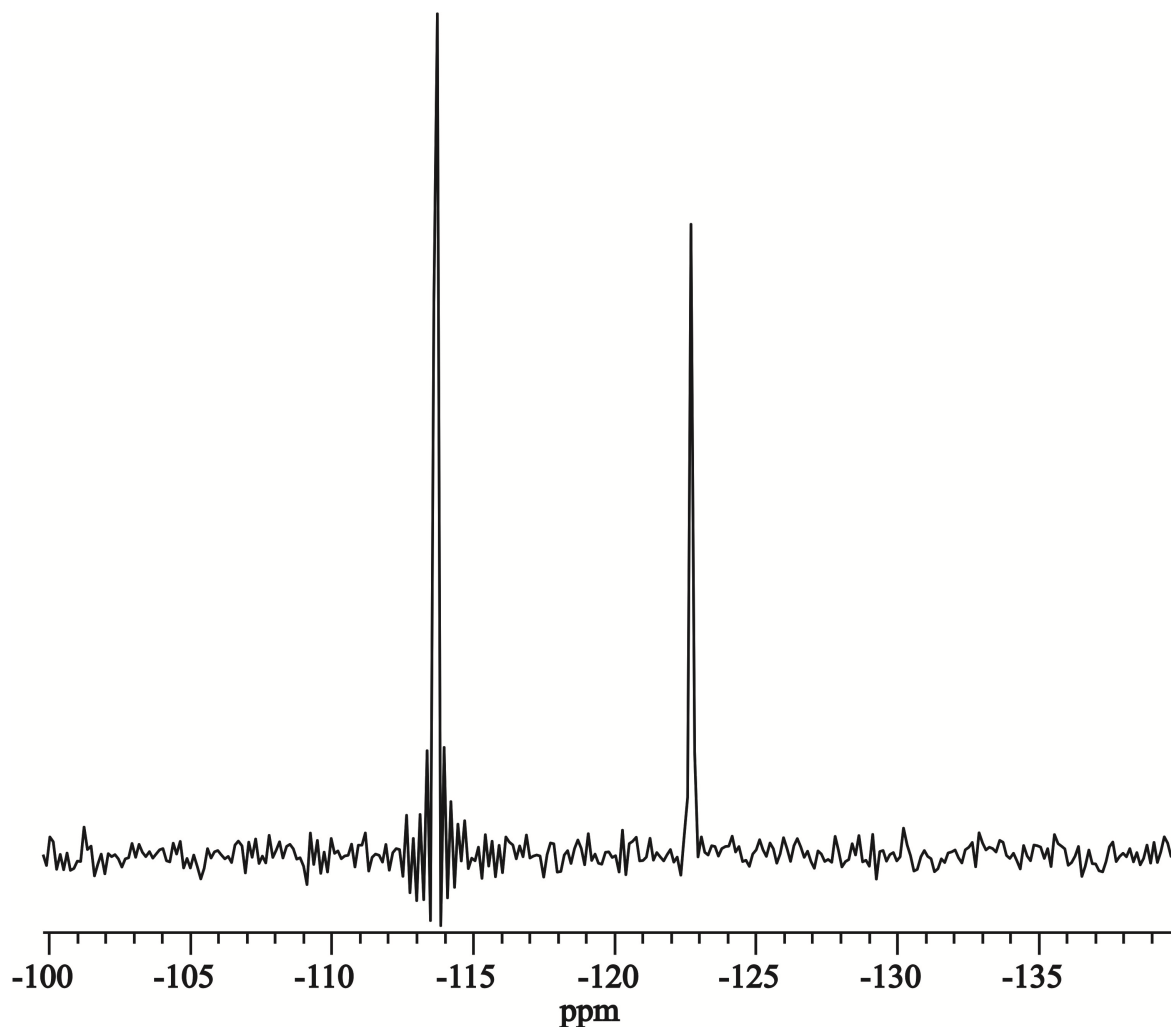


Figure 18. Adamantane ^{13}C spectrum prior to chemical shift referencing. The transmitter was set near the ^{13}CO Larmor frequencies to increase ^{13}CO signal intensity in REDOR experiments as described in **Figure 10**. The chemical shift for the left adamantane peak is 40.5 ppm downfield from the tetramethyl silane (TMS) internal standard reference, but is observed at -113.7 ppm in this figure. Therefore, 154.2 ppm should be added to the chemical shift to correctly reference chemical shifts. Chemical shift referencing is important since referenced chemical shifts of carbon nuclei in peptides provide information about local secondary structure⁷⁷.

2.4 MAS Solid-State NMR (REDOR)

Data were collected on a 9.4 T spectrometer (Varian Infinity Plus, Palo Alto, CA) using triple and quadruple resonance MAS probes equipped for 4.0 mm rotors and tuned to ^{13}C , ^1H ,

and ^{15}N nuclei at respective frequencies of 100.8, 400.8, and 40.6 MHz. The ^{13}C chemical shift was externally referenced to the methylene resonance of adamantane at 40.5 ppm, and the ^{13}C transmitter was set to ~ 153 ppm. The ^{13}CO - ^{15}N dipolar coupling (d_{CN}) was probed with REDOR experiments with typical parameters: (1) 50 kHz ^1H $\pi/2$ pulse; (2) 1.6 ms cross-polarization with 50 kHz ^1H field and 56-63 kHz ramped ^{13}C field; (3) dephasing period of duration τ for which the “ S_0 ” and “ S_1 ” acquisitions had 45 kHz ^{13}C π pulses at the end of each rotor cycle except the last cycle, and the S_1 acquisitions additionally had 25 kHz ^{15}N π pulses in the middle of each rotor cycle; and (4) ^{13}C detection^{54,78}. The MAS frequency was 10 kHz, the recycle delay was 2 s, 85 kHz TPPM ^1H decoupling was applied during the dephasing period ($\tau = 2.2, 8.2, 16.2, 24.2, 32.2, 40.2, \text{ and } 48.2$ ms) and detection period. The XY-8 phase cycling (x, y, x, y, y, x, y, x) was used for the heteronuclear train of π pulses during τ except for the last ^{13}CO π pulse^{73,79}. Samples were typically cooled by nitrogen gas at -50 °C to enhance ^{13}CO signal and reduce motional averaging of d_{CN} ⁸⁰. The typical difference between ^{13}C shift in cooled and uncooled membrane-associated HFP samples is ≤ 0.5 ppm and indicates little variation in secondary structure with temperature⁵¹.

NMR pulse sequences are designed to distinguish observable interactions between nuclei in an NMR sample. In solid state NMR experiments, this can be challenging because a functional groups, such as a ^{13}CO , have a unique chemical shift that depends upon the ^{13}CO bond orientation relative to the $\mathbf{B_0}$ field. Peptide/protein NMR samples contain many ^{13}CO bonds

which results in a broad distribution of ^{13}CO bond orientations and therefore a broad distribution of chemical shifts (also commonly referred to as a powder pattern). By spinning a sample at a high frequency along an axis, the average orientation of a ^{13}CO bond over one rotor period lies along the axis of rotation which makes the average orientation of each ^{13}CO bond equivalent for all ^{13}CO functional groups over a rotor period. When the axis of rotation set to be the magic angle, 54.7° , the dipolar coupling energy has an average value of zero over each rotor period as per **Eq (20)**. This results in much narrower line widths.

Peptides and proteins contain many carbonyl and amide functional groups and solid state NMR experiments often detect ^{13}CO signals to obtain protein and peptide structural information. The ^{13}CO REDOR experiments collect S_0 and S_1 spectra where individual ^{13}CO nuclear magnetic moments precess in the **xy** plane of the rotating frame at different rates due to different shielding from the **B₀** field and dipolar couplings. These different rates of precession lead to decoherence of the ^{13}CO nuclear magnetic moments. Because magnetization is the sum of nuclear magnetic moments, decoherence of the ^{13}CO nuclear magnetic moments in the **xy** plane leads to a smaller magnitude vector sum and detection of a smaller observable ^{13}CO magnetization (i.e. smaller ^{13}CO signal). Consider the different precession rates of the ^{13}CO nuclear magnetic moments due different shielding from the **B₀** field. Ignoring longitudinal relaxing contributions that result in the loss of transverse magnetization, coherence of the ^{13}CO nuclear magnetic moments can be reestablished by introducing time synchronized ^{13}CO π pulses

as accomplished in the Hartman-Hahn Spin-Echo pulse sequence⁸¹. Thus, one function of the ^{13}CO π pulses in a REDOR experiment is to establish coherence of the nuclear magnetic moments in the rotating frame at the time of signal acquisition to enhance the signal acquired per acquisition. Additionally, the REDOR experiments in this dissertation introduced ^{13}CO and ^{15}N labeled residues making ^{13}CO - ^{15}N dipolar couplings (d_{CN} s) an observable property. The d_{CN} can be related to distance between ^{13}CO and ^{15}N nuclei (r_{CN}) by $d_{\text{CN}} = 3080/r_{\text{CN}}^3$ where d and r are expressed in Hz and Å, respectively (See **2.5 SIMPSON Simulations** for discussion regarding isolating observable NMR interactions). A dipolar coupling interaction between two nuclei lead to precession of the ^{13}C magnetization where ^{13}C magnetization is detected in the **xy**-plane of the rotating frame. The rate of precession of ^{13}CO magnetization that results from this for ^{13}C - ^{15}N dipolar coupling interaction can be described by **Eq (22)**

$$w_{d_{\text{CN}}}(\phi) = \frac{d_{\text{CN}}}{2} \left(3 \cos^2 \phi - 1 \right) \quad (22)$$

where ϕ is the angle between \mathbf{B}_0 and the internuclear vector. This observable property can be negated by magic angle spinning (MAS) which makes the average value for ϕ equal to 54.7° . Experimentally, this is achieved by making a 54.7° angle between the rotor axis and \mathbf{B}_0 . Spinning the rotor at a high frequency makes the average orientation for any internuclear vector equal to 54.7° resulting in an average dipolar coupling equal to zero. Thus, MAS results in no net evolution of ^{13}C transverse magnetization from a dipolar coupling interaction as observed in the S_0 spectra in REDOR experiments. In the S_1 spectra, a ^{13}C or ^{15}N π pulse is applied at the

middle and end of each rotor period. The direction of the precession of the ^{13}CO nuclear magnetic moments in the rotating frame that results from this for ^{13}C - ^{15}N dipolar coupling interaction can be reversed by either a ^{13}C or ^{15}N π pulse⁸². By applying a ^{13}C or ^{15}N π pulse at the middle and end of each rotor period, the precession of the ^{13}CO nuclear magnetic moments in the rotating frame is nonzero. This results in average dipolar coupling interaction making d_{CN} an observable property in the S_1 spectra. The precession of the ^{13}CO nuclear magnetic moments due to ^{13}C - ^{15}N dipolar coupling interactions results in decoherence of the ^{13}CO nuclear magnetic moments and consequent loss of ^{13}CO magnetization (i.e. ^{13}CO signal). This decoherence can be quantified in terms of the average d_{CN} in the S_1 spectra relative to the S_0 spectra⁷⁸. The d_{CN} were quantified by comparing the ^{13}CO signal intensities of the S_0 spectra (no average ^{13}CO evolution due to d_{CN}) to the reduced ^{13}CO signal intensities of the S_1 spectra (evolution from d_{CN}) where $d_{\text{CN}} \propto r_{\text{CN}}^{-3}$. Greater reduction of the ^{13}CO signal intensity in the S_1 spectrum relative to the S_0 spectrum generally occurred for longer τ . These ^{13}CO peak intensities were denoted S_1 and S_0 and were determined from integration over a shift range that encompassed most of the ^{13}CO signal. The width of the integrated region was different for each sample type, but the normalized dephasing was quantified by:

$$\left(\frac{\Delta S}{S_0} \right)^{\text{exp}} = \frac{S_0^{\text{exp}} - S_1^{\text{exp}}}{S_0^{\text{exp}}} = 1 - \frac{S_1^{\text{exp}}}{S_0^{\text{exp}}} \quad (23)$$

and its standard deviation:

$$\sigma^{exp} = \left(\frac{\Delta S}{S_0} \right) \sqrt{ \frac{ \left(\frac{\sigma_{S_0}^2}{S_0^2} \right) + \left(\frac{\sigma_{S_1}^2}{S_1^2} \right) }{ \left(\frac{S_0 - S_1}{S_0 S_1} \right)^2 } + \left(\frac{\sigma_{S_0}}{S_0} \right)^2 } \quad (24)$$

where σ_{S_0} and σ_{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 features⁸³. The calculation in Eq (24) is derived from error analysis of $(\Delta S/S_0)$ and did not consider a covariance term. Alternatively, error analysis could be derived from calculating $(1-(S_1/S_0))$.

$$\sigma^{exp} = \left(\frac{S_1}{S_0} \right) \sqrt{ \left(\frac{\sigma_{S_1}}{S_1} \right)^2 + \left(\frac{\sigma_{S_0}}{S_0} \right)^2 } \quad (25)$$

Eq (24) was used to quantify error in this dissertation and comparisons between these two methods for calculating error are found in **Table 3** and **Table 4**. Typical variation in σ^{exp} between these error analysis methods appears to be 0.000 to 0.003 for our experiments.

Table 3. Error analysis for V2E-F8CG13N.

τ	$(\Delta S/S_0)$	S_0	S_1	σ_{S_0}	σ_{S_1}	Error (Eq (24))	Error (Eq (25))
48.2	0.400	100.2861	60.4066	0.767	1.435	0.017	0.015
40.2	0.379	94.9187	58.9761	0.980	0.808	0.014	0.011
32.2	0.330	106.9589	71.6677	0.777	0.421	0.009	0.006
24.2	0.262	124.2104	91.6186	0.836	0.878	0.010	0.009
16.2	0.146	127.5113	108.9537	0.637	0.612	0.007	0.006
8.2	0.056	70.7462	66.8136	0.329	0.316	0.006	0.006
2.2	0.009	81.0649	80.3265	0.549	0.303	0.008	0.008

Table 4. Error analysis for HFP-L12CA6N.

τ	$(\Delta S/S_0)$	S_0	S_1	σ_{S_0}	σ_{S_1}	Error (Eq (24))	Error (Eq (25))
48.2	0.275	135.2543	98.0556	1.658	2.267	0.021	0.019
40.2	0.247	109.808	82.7338	1.292	1.127	0.016	0.014
32.2	0.192	137.7211	111.2584	1.193	0.988	0.011	0.010
24.2	0.155	126.5162	106.8944	1.231	1.07	0.013	0.012
16.2	0.099	194.515	175.3549	1.136	0.853	0.007	0.007
8.2	0.058	86.3175	81.3029	0.363	0.758	0.010	0.010
2.2	0.004	74.1142	73.7822	0.323	0.57	0.009	0.009

2.5 SIMPSON Simulations

MAS solid state NMR pulse sequences are typically designed to limit the number of observable spin interactions that affect signal intensities and line shapes in order to simplify experimental data interpretations. REDOR experiments incorporate heteronuclear trains of π pulses that are rotor synchronized and result in time dependent signal attenuation in the S_1 spectra relative to the S_0 spectra that results from a loss of transverse magnetization due to dipolar couplings. Quantification of these dipolar couplings can be achieved by comparing the experimental data to simulated data. REDOR experiments can be simulated with the SIMPSON program⁸⁴ which was used as a tool to evaluate the evolution of the time-dependent Schrödinger equation. The SIMPSON program is an accessible program that is capable of simulating many

different pulse sequences, but I will provide a brief introduction from the literature⁸⁴ with key concepts relevant to the REDOR pulse sequence highlighted.

The SIMPSON simulations use the Liouville Von Neumann equation

$$\frac{d}{dt} \rho(t) = -i[H(t), \rho(t)] \quad (26)$$

where specific observables can be quantified using time dependent wavefunction/s ($\psi(t)$) time-dependent Hamiltonian/s ($H(t)$) and the density matrix operator ($\rho(t)$). In general, a time dependent wavefunction, $\psi(t)$, can be used to describe quantum mechanics. While $\psi(t)$ can be a linear combination of many basis states, $\psi(t)$ is defined below for a simple wavefunction with two basis states:

$$\psi(t) = \sum_{p=1,2} c_p(t) \psi_p \quad (27)$$

where ψ_p represents two time independent basis states, and the time dependent terms are accumulated in the coefficients, $c_p(t)$. Wavefunctions can be incorporated into density matrix theory to follow the time evolution of any operator using the time dependent density matrix operator, $\rho(t)$, where

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| \quad (28)$$

For a simple two-state quantum mechanics system, this can be represented in matrix form by:

$$\rho(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} \begin{pmatrix} c_1^*(t) & c_2^*(t) \end{pmatrix} = \begin{pmatrix} |c_1(t)|^2 & c_1(t)c_2^*(t) \\ c_2(t)c_1^*(t) & |c_2(t)|^2 \end{pmatrix} \quad (29)$$

When normalized, the trace($\rho(t)$) = 1.

The expectation value for an operator, such as $H(t)$, can be calculated as follows:

$$\begin{aligned}
\langle H(t) \rangle &= \langle \psi(t) | H(t) | \psi(t) \rangle = \begin{pmatrix} c_1^*(t) & c_2^*(t) \end{pmatrix} \begin{pmatrix} H(t)_{11} & H(t)_{12} \\ H(t)_{21} & H(t)_{22} \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} \\
&= |c_1(t)|^2 H(t)_{11} + c_1(t) c_2^*(t) H(t)_{21} + c_2(t) c_1^*(t) H(t)_{12} + |c_2(t)|^2 H(t)_{22} \quad (30) \\
&= \text{trace} \begin{pmatrix} |c_1(t)|^2 & c_1(t) c_2^*(t) \\ c_2(t) c_1^*(t) & |c_2(t)|^2 \end{pmatrix} \begin{pmatrix} H(t)_{11} & H(t)_{12} \\ H(t)_{21} & H(t)_{22} \end{pmatrix} = \text{trace}(\rho(t) H(t))
\end{aligned}$$

A detailed step by step description describing how to use density matrix theory to simulate NMR experiments for a simple spin $\frac{1}{2}$ system is described in literature⁸⁵, and further expansion of density matrix theory to a two spin system is also described in literature⁸⁶. These articles describe the basic principals that were used to build the framework for SIMPSON simulations. Currently, these two aforementioned papers can be obtained through MSU's interlibrary loan system, and in my graduate school experience, these papers provided the most clear and complete overview of density matrix theory for NMR experiments.

Furthermore, the Liouville Von Neumann equation can be rewritten to evaluate the density operator at a given time by:

$$\rho(t) = U(t,0) \rho(0) U^\dagger(t,0) \quad (31)$$

where $\rho(t)$ is the density operator at a given time, $\rho(0)$ is the density operator at equilibrium, and $U(t,0)$ is a unitary operator used to evaluate the time evolution of spin system for time 0 to t .

$$U(t,0) = T \exp \left(-i \int_0^t H(t') dt' \right) \quad (32)$$

where T is the Dyson time-ordering operator that contains noncommuting terms of the Hamiltonian. This expression can be simplified to express as a time-ordered product.

$$U(t,0) = \prod_{j=0}^{n-1} \exp(-iH(j\Delta t)\Delta t) \quad (33)$$

where n is the number of time increments needed for time 0 to t for which the $H(t)$ can be approximated to be time-independent for each Δt increment. For a simple one spin system, spin I , the components of the Hamiltonian can be described by:

$$H = H_{RF} + H_{CS} + H_J + H_D + H_Q \quad (34)$$

where the Hamiltonians for the resonant frequency pulse (RF), chemical shift (CS), J spin coupling (J), dipole-dipole coupling (D), and quadrupolar coupling (Q) can be incorporated into the unitary operator. The HCN MAS REDOR experiments allow for isolation of the H_{RF} , H_{CS} , H_D terms.

$$H_{RF} = \sum_a \left| w_{rf}^a(t) \right| (I_{ax} \cos \vartheta_a + I_{ay} \sin \vartheta_a) \quad (35)$$

$$H_{CS} = \sum_a w_{CS,0}^a(t) I_{az} \quad (36)$$

$$H_D = \frac{\mu_0 \gamma_a \gamma_b}{4\pi r_{ab}^3} (1 - 3 \cos^2 \phi) \hat{I}_{az} \hat{I}_{bz} \quad (37)$$

where a and b refer to specific nuclei spins, ϑ is the phase of the applied field, ϕ is the angle between \mathbf{B}_0 and the internuclear vector, and μ_0 is the magnetic permeability. Thus, input parameters that affect these Hamiltonians are incorporated into the SIMPSON simulations and allow for comparisons between experimental data and simulated data. For HCN experiments, the ^{13}CO - ^{15}N dipolar coupling (d_{CN}) can be related to distance between by ^{13}CO and ^{15}N nuclei (r_{CN}) by $d_{CN} = 3080/r_{CN}^3$ where d and r are expressed in Hz and Å, respectively. NMR samples were modeled as a sum of S_0 and a sum of S_1 signals of different structure dependent spin geometries of ^{13}CO and ^{15}N nuclei (see Chapter III). Unless otherwise noted, $r_{CN} > 7 \text{ Å}^{63}$ were

not included in simulation and $S_1 = S_0$ for these nuclei. The notation $(\Delta S/S_0)^{sim}$ will be generally used for simulated $(\Delta S/S_0)$ and can refer to a particular generalized spin geometry or to the population weighted sum using calculations from different spin geometries over the range of experimental dephasing times τ . For the former case, the $(S_1/S_0)^{sim} \equiv \gamma$ were calculated using the SIMPSON program with input parameters that included d_{CN} s as well as Euler angles in a fixed crystal frame for each ^{13}CO - ^{15}N vector and for the ^{13}CO chemical shift anisotropy (CSA) principal axis system^{84,87}. For 5 spin simulations, these input parameters were calculated by the SIMMOL program using ^{13}CO and ^{15}N coordinates from a region of a high-resolution crystal structure with the appropriate structural motif, e.g. antiparallel β sheet⁸⁸. Similarly, the (α, β, γ) Euler angles of d_{CN} PAS were $(0, 0, 0)$ and $(0, \vartheta, 0)$ for 2 and 3 spin simulations, respectively, where ϑ was the angle between two d_{CN} vectors. Specific Protein Data Bank (PDB) files used to model spin geometries associated with a structural motif are noted in the Materials and Methods section of each Chapter that simulations were used. 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}CO CSA principal values of 247, 176, and 99 ppm were inputs to the SIMPSON calculations⁸⁹, and ^1H s and relaxation rates were not considered. The same set of chemical shift input parameters were used for all comparisons between simulated and experimental data. Using the same ^{13}CO CSA principal values as above, the isotropic chemical shift was varied between 167-180 ppm for a 5 spin system (four ^{15}N , one

^{13}CO) which more than covered the range of typical β sheet isotropic chemical shifts and the $(S_1/S_0)^{sim}$ at 48 ms was affected by less than 0.01. SIMPSON and SIMMOL script files for each spin geometry type can be found in **Appendix VI** where spin geometry type refers to the number of ^{13}CO and ^{15}N nuclei used in each simulation (not structure type).

Chapter III. Natural Abundance Calculations for Solid State NMR REDOR Experiments and Quantitative Determination of In-Register Parallel β Sheet Registries in Membrane-Associated HFP

3.1 Background.

The HFP structure-function literature includes NMR data showing random coil structure for HFP in aqueous solution^{32,33}. Solid-state nuclear magnetic resonance (SSNMR) has shown predominant β sheet structure for residues 1-16 of membrane-associated HFP (mHFP) where the membranes contained ~30 mol% cholesterol, which is comparable to the mol% cholesterol of membranes of HIV and host cells of HIV^{37,50,58,90}. 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 solution³⁴. The experimental rates (R_s) were ordered $R_{HFP \text{ membrane binding}} > R_{HFP \beta \text{ sheet formation}} > R_{lipid \text{ mixing}}$ and were consistent with the sequence: (1) random coil HFPs binds to membrane vesicles and HFP structure changes to oligomeric β sheet; (2) vesicle fusion.

The biological relevance of HFP oligomers, potentially β sheet conformation, is further supported by the molecular trimer structure of soluble regions of the gp41 ectodomain^{17,18} (~175 residues). In these structures, the residue 30-80 region of each molecule forms a continuous helix, and the helices of the different molecules form a parallel coiled-coil. The Thr-25 residue places the N-terminal helical regions of three gp41 close to one another. The Thr-25 is C-terminal of the approximate 16 residue N-terminal fusion peptide which motivated the study of a C-terminally cross-linked HFP trimer (HFPtr). Relative to HFP monomer, HFPtr induced membrane vesicle fusion with ~15-40-fold faster rate which supported the functional

significance of the trimer³⁰. Although both mHFP and mHFPtr formed β sheet oligomers, mHFPtr is more deeply inserted which correlates with greater membrane perturbation and a reduction of the vesicle fusion activation energy⁴⁰. 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^{21,41}. Analyses of these data supported the involvement of multiple gp41 trimers and fusion peptides in fusion⁴². Electron micrographs of virus-cell contacts have also been interpreted to show multiple gp41 trimers at the contact site⁴³. Functional importance of fusion peptide trimers has also been demonstrated for fusion peptides of other viruses^{44,45}.

Because of the aforementioned functional significance of HIV fusion peptide oligomers, there has been effort to elucidate the distribution of structures of mHFP 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 solution³⁷. 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^{30,33,49}. HFP/lipid binding was supported by SSNMR detection of a HFP Ala-1 ^{13}C CO(carbonyl)-lipid ^{31}P distance of $\sim 5 \text{ \AA}$ ⁴⁰. For mHFP, the ^{13}C chemical shifts derived from an unambiguous assignment were consistent with a fully extended β strand conformation for residues between Ala-1 and Gly-16⁵⁰. Detection of intermolecular ^{13}C - ^{13}C and ^{13}C - ^{15}N distances of $\sim 5 \text{ \AA}$ supported β sheet oligomer/aggregate structure, and the Ala-1

^{13}CO -lipid ^{31}P contact and other data suggest that the number of molecules in the oligomer is small^{40,50,51}.

This chapter focuses on quantitative determination of populations of in-register parallel β sheets relative to the previously identified antiparallel β sheet registries³⁶. The clearest information to-date on this topic has been a SSNMR experiment on membrane-associated HFP with an Ala-14 ^{13}CO label and a Gly-3 ^{15}N label whose separation (r_{CN}) was >20 Å along a single β strand⁵⁰. SSNMR can detect labeled ^{13}CO - ^{15}N dipolar coupling (d_{CN}) where $d_{\text{CN}} = 3080/r_{\text{CN}}^3$ with d in Hz and r in Å. The minimum detectable $d_{\text{CN}} \sim 10$ Hz correlates with $r_{\text{CN}} \sim 7$ Å so that detectable d_{CN} in this sample were necessarily ascribed to inter- rather than intramolecular ^{13}CO - ^{15}N proximity. SSNMR detection of $d_{\text{CN}} > 30$ Hz strongly supported a significant fraction of molecules with intermolecular Ala-14-Gly-3 hydrogen bonding and labeled r_{CN} of 4.1 and 5.5 Å, i.e. 16 \rightarrow 1/1 \rightarrow 16 antiparallel β sheet registry ($t = 16$). **Figure 19** displays this registry with isotopic labeling from the present study and not the earlier study. Detection of similarly large d_{CN} in an Ala-14 ^{13}CO /Ile-4 ^{15}N HFP sample supported a fraction of 17 \rightarrow 1/1 \rightarrow 17 antiparallel registry ($t = 17$). As noted above, multiple gp41 trimers are required for membrane fusion and close proximity of multiple gp41 trimers make both parallel and interleaved antiparallel β sheet oligomers potential models for HFP oligomeric structure in vivo.

(a) Peptides				
	1	10	20	30
HFP-F8	AVGIGAL	F LGFLGAAGSTMGARSWKKKKKKA		
HFP-L12	AVGIGALFLGF	L GAAGSTMGARSWKKKKKKA		
HFP-G5A6	AVGI	G ALFLGFLGAAGSTMGARSWKKKKKKA		
HFP-A6L7	AVGIGAL	A LFLGFLGAAGSTMGARSWKKKKKKA		
HFP-L12G13	AVGIGALFL	G LFLGAAGSTMGARSWKKKKKKA		
HFP-G13A14	AVGIGALFLGFL	G AAGSTMGARSWKKKKKKA		
(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		Parallel	
	1-->17/1-->17		2-->17/1-->16	
HFP-P	AVGIGALFLGFL	G AAGS	AVGIGALFLGFL	G AAGS
	AVGIGALFLGFL	L GAAGS	AVGIGALFLGFL	L GAAGS
HFP-A	AVGI	G ALFLGFLGAAGS	AVGI	G ALFLGFLGAAGS
	AVGIGALFLGFL	L GAAGS	AVGIGALFLGFL	L GAAGS
HFP-AP	AVGIGALFL	G LFLGAAGS	AVGIGALFL	G LFLGAAGS
	AVGIGAL	F LGFLGAAGS	AVGIGAL	F LGFLGAAGS
	Antiparallel		Antiparallel	
	16-->1/1-->16		17-->1/1-->17	
HFP-P	SGA	A GLFGLFLAGIGVA	SGA	A GLFGLFLAGIGVA
	AVGIGALFLGFL	L GAAGS	AVGIGALFLGFL	L GAAGS
HFP-A	SGAAGLFGLFL	A GIGVA	SGAAGLFGLFL	A GIGVA
	AVGIGALFLGFL	L GAAGS	AVGIGALFLGFL	L GAAGS
HFP-AP	SGAAGLF	G LFLAGIGVA	SGAAGLF	G LFLAGIGVA
	AVGIGAL	F LGFLGAAGS	AVGIGAL	F LGFLGAAGS

Figure 19. (a) HFPs where red and blue correspond to ^{13}C O and ^{15}N labeled residues, respectively. (b) HFP-NC, HFP-P, HFP-A, and HFP-AP were SSNMR samples which each contained a mixture of ^{13}C O and ^{15}N labeled peptides in 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 β sheet structure with a molecular mixture of ^{13}C O and ^{15}N labeled peptides in the sample. (c) Registries probed by the SSNMR REDOR experiments and labeled ^{13}C O/labeled ^{15}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 in HFP. For parallel sheets, there is CO (residue h) – HN (residue $h+1$) hydrogen bonding of adjacent strands.

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 Ala-14 ^{13}C O/Gly-3 ^{15}N or Ala-14 ^{13}C O/Ile-4 ^{15}N labeled samples. Because of the close proximity of the Thr-25 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 **Figure 19c**. 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}C O labels and the other with three sequential backbone ^{15}N labels²⁹. Detection of an average $d_{\text{CN}} > 10$ Hz for a Gly-5-Leu-7 ^{13}C O/Gly-5-Leu-7 ^{15}N sample and a Phe-11-Gly-13 ^{13}C O/Phe-11-Gly-13 ^{15}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}C - ^{13}C dipolar couplings (d_{CC} s) where $d_{\text{CC}} = 7710/r_{\text{CC}}^3$ with d_{CC} in Hz and r_{CC} in Å. For mHFP with a single ^{13}C O label and in-register parallel structure, the labeled interstrand $r_{\text{CC}} \sim 5$ Å with $d_{\text{CC}} \sim 70$ Hz^{52,53}. These parameters will be independent of the residue that is ^{13}C O labeled. For mHFP with Phe-8 ^{13}C O, a best-fit $d_{\text{CC}} \sim 70$ Hz was detected whereas for mHFPtr, d_{CC} depended on the position of the labeled ^{13}C O residue with a range of 10-60

Hz^{54,55}. 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 ¹³C labeling at either: (1) Ala-1 to Val-3, Gly-5 to Ile-9; (2) Phe-8 to Gly-16; or (3) Ala-1 to Val-3, Gly-5 to Gly-16⁴⁷. The IR wavenumbers and intensities of different samples were interpreted to support a large fraction of in-register parallel structure and little $t = 16$ antiparallel structure. However, in my 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.

This chapter reports a determination of the fraction of in-register parallel structure in mHFP oligomers. This study was motivated because of: (1) the functional significance of HIV fusion peptide oligomers; and (2) the existing undefinitive and conflicting data and interpretations relevant to this question. As part of this effort, an experimentally validated model was developed to quantify effects of natural abundance ¹³C and ¹⁵N nuclei on SSNMR measurements of d_{CN} .

3.2 Materials and Methods.

1. SSNMR samples.

As shown in Error! Reference source not found.**b**, each sample contained a ¹³CO and a ¹⁵N labeled peptide in a 1:2 mol ratio. Three samples contained mHFP β sheet oligomers/aggregates that were each a statistical mixture of ¹³CO and ¹⁵N labeled HFPs. Detection of substantial d_{CN} by SSNMR indicated proximity of the labeled ¹³CO and ¹⁵N nuclei

on adjacent strands and was used to estimate the fractional populations of specific registries as detailed below. As shown in **Figure 19c**, 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 ($t = 16$ and $t = 17$), 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 ^{13}CO and ^{15}N nuclei, there will always be proximity between labeled ^{13}CO and some natural abundance (na) ^{15}N nuclei as well as proximity between some na ^{13}CO and labeled ^{15}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 ^{13}CO , ^{15}N , and na sites as the HFP-P, HFP-A, and HFP-AP samples; and (2) labeled ^{13}CO – labeled ^{15}N r_{CN} that are much greater than the approximate REDOR detection limit of ~ 7 Å. 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 mHFP was not yet well-defined at the time of these experiments. 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 ^{13}CO s and ^{15}N s remained much farther apart than the 7 Å, the approximate REDOR

detection limit. Although there were populations of β sheet as well as α helical lyophilized peptides in the HFP-NC sample, each population yielded very similar $(\Delta S/S_0)$ – see Results section for further details.

2. Modeling.

Experimental dephasing of a mHFP sample was modeled as a sum of S_0 and a sum of S_1 signals of different spin geometries of ^{13}CO and ^{15}N nuclei where the geometries reflected statistical distributions of ^{13}CO and ^{15}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 $(S_1/S_0)^{sim} \equiv \gamma_l$ 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}CO - ^{15}N vector and for the ^{13}CO chemical shift anisotropy (CSA) principal axis system^{84,87}. These input parameters were calculated by the SIMMOL program using ^{13}CO and ^{15}N coordinates from a region of a high-resolution crystal structure with the appropriate structural motif, e.g. parallel β sheet⁸⁸. 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}C CO CSA principal values of 247, 176, and 99 ppm were inputs to the SIMPSON calculations⁸⁹ and ^1H s and relaxation were not considered.

Chemical shift and conformational distributions. **Figure 20** displays REDOR S_0 and S_1 ^{13}C SSNMR spectra for $\tau = 32.2$ ms. Each S_0 spectrum has a ~50% contribution from the labeled ^{13}C CO and ~50% contribution from natural ^{13}C COs of the unlabeled residues. The full-width at half-maximum linewidths of the mHFP samples in **Figure 20b-d** are 3-4 ppm and indicate a distinct secondary structure. For the HFP-AP sample with Phe-8 ^{13}C CO label, the peak ^{13}C CO shift of 173 ppm is the same as was observed for Phe-8 ^{13}C CO of mHFP in known β strand conformation and is very different from the 178 ppm shift observed in α helical conformation^{50,54}. For the HFP-P and HFP-A samples with Leu-12 ^{13}C CO label, the 174 ppm peak shift is also the same as β strand HFP and different from the 179 ppm shift of Leu in helical HFP^{54,91}. 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 content⁵⁰.

The linewidth of the lyophilized HFP-NC sample with F8 ^{13}C CO label is ~7 ppm and correlates with a broad distribution of secondary structures that is also evidenced by a 176 ppm peak ^{13}C CO shift that is midway between typical Phe helical and β strand shifts⁷⁷. S_0 and S_1 signals were quantified by integration of 8 ppm regions for the HFP-NC spectra and 5 ppm

regions were used for HFP-P, HFP-A, and HFP-AP spectra. These regions were chosen to encompass most of the ^{13}CO signal, and they also resulted in low σ^{exp} .

3.3 Results.

1. Qualitative Analysis of the REDOR Data.

Relative to the S_0 signals, there is attenuation in the S_1 signals of ^{13}CO s within ~ 7 Å of ^{15}N s and the associated $\Delta S/S_0$ normalized dephasing increased with dephasing time, **Figure 20** and **Figure 21**. Because of the physical separation of the ^{13}CO and ^{15}N labeled HFPs in HFP-NC, the S_1 attenuation and $(\Delta S/S_0)^{\text{exp}}$ of this sample reflected Phe-8 ^{13}CO -na ^{15}N and na ^{13}CO -Ala-6,Leu-7 ^{15}N proximities but not Phe-8 ^{13}CO -Ala-6,Leu-7 ^{15}N proximity, **Figure 20a** and **Figure 21a**. There was similar S_1 attenuation and $(\Delta S/S_0)^{\text{exp}}$ in HFP-P, **Figure 20b** and **Figure 21b**, which demonstrated that there was little Leu-12 ^{13}CO -Gly-13,Ala-14 ^{15}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 $(\Delta S/S_0)^{\text{exp}}$ for the HFP-A sample, **Figure 20c** and **Figure 21b**, which indicated significant Leu-12 ^{13}CO -Gly-5,Ala-6 ^{15}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 $(\Delta S/S_0)^{\text{exp}}$ were observed for the HFP-AP sample, **Figure 20d** and **Figure 21b**. 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 Phe-8 ^{13}CO -Leu-9,Gly-10 ^{15}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→17/1→16 registries. Detection of a substantial fraction of these antiparallel registries is consistent with earlier SSNMR data for sparsely labeled HFP⁵⁰. 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^{29,47}. These data highlight the importance of sparse labeling to reduce interpretational ambiguity for systems with a structural distribution like mHFP.

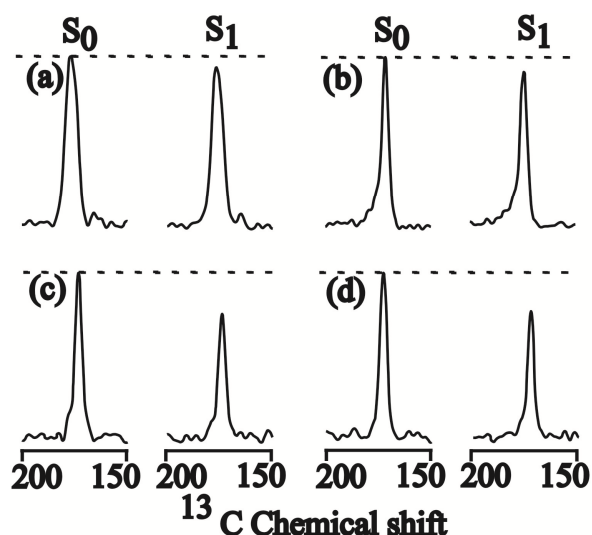


Figure 20. REDOR S_0 and S_1 ^{13}C SSNMR spectra at 32.2 ms dephasing time for (a) HFP-NC, (b) HFP-P, (c) HFP-A, or (d) HFP-AP. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 38624; (b) 23488; (c) 24914; or (d) 14240 scans. Relatively narrow ^{13}C CO signals were observed in the HFP-P, HFP-A, and HFP-AP samples because the HFPs were membrane-associated with predominant β sheet conformation at the labeled ^{13}C CO site. A broader ^{13}C CO signal was observed in the HFP-NC sample because there was no membrane and there were populations of lyophilized HFP with either α helical or β sheet conformation at the labeled ^{13}C CO site.

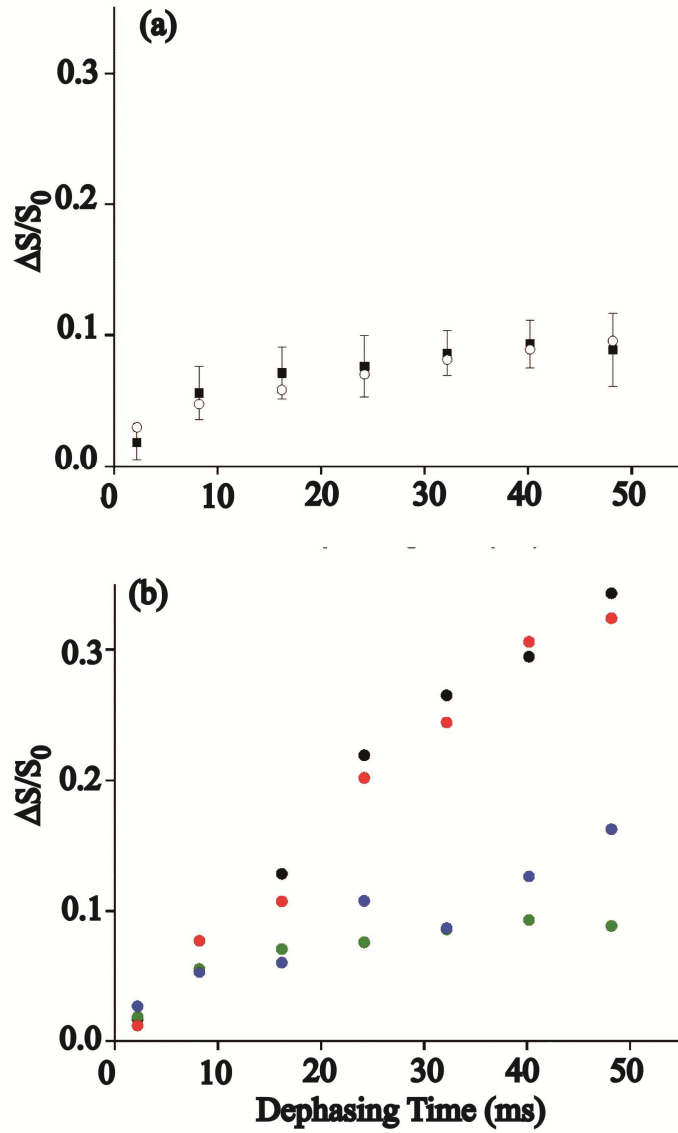


Figure 21. (a) Plot of REDOR $(\Delta S/S_0)^{exp}$ (filled squares with error bars) and $(\Delta S/S_0)^{sim}$ (open circles) vs dephasing time for the lyophilized HFP-NC sample. The $(\Delta S/S_0)^{sim}$ were calculated using a mixture of nad models with fractional populations: α helical, 0.5; min β sheet, 0.25; max β 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 σ^{exp} is ± 0.02 . Variation of ± 0.02 in $(\Delta S/S_0)^{exp}$ was also observed between two different preparations of the same sample type, e.g. HFP-A.

2. 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 (nad) model, i.e. a model that accounts for effects of labeled ^{13}CO -na ^{15}N and na ^{13}CO -labeled ^{15}N proximities. Both types of proximities were considered but for conciseness of presentation, the discussion in this chapter focuses on labeled ^{13}CO -na ^{15}N . One measure of validity of a nad 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}CO s and one na ^{15}N . Geometries with two or more na ^{15}N s were not considered because the fractional isotopic abundance of ^{15}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 τ . 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 for these experiments. We consider this a “long-range” nad model which is distinguished from a “short-range” model of earlier studies that only considered na nuclei separated by one or two bonds from a labeled nucleus, i.e. $r_{CN} < 3 \text{ \AA}$ ^{50,54}. The broad spectral linewidth of HFP-NC indicated both helical and β strand conformational populations and coordinates of spin geometries for both α helical and β sheet structures were obtained from corresponding regions of high-resolution structure PDB files. For α helical structure, the $r_{CN} < 7 \text{ \AA}$ criterion resulted in geometries with a single labeled ^{13}CO at

residue h and a single $\text{na } ^{15}\text{N}$ at a residue between $h - 3$ and $h + 5$. These nine geometries are one aspect of the α nad model.

Figure 22 illustrates relevant labeled ^{13}CO s and $\text{na } ^{15}\text{N}$ s for antiparallel β 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}CO and one $\text{na } ^{15}\text{N}$ within the same strand and three geometries had two labeled ^{13}CO s on vicinal strands and one $\text{na } ^{15}\text{N}$ in the intervening strand. In panels b and d, the strands have different registries so that the labeled ^{13}CO in the top strand was $>7 \text{ \AA}$ from the nine $\text{na } ^{15}\text{N}$ sites of the ^{13}CO of the third strand. The structure of panels b and d has nine unique spin geometries and is denoted a maximum β sheet nad (max β nad) model while the structure of panels a and c has six geometries and is denoted a minimum β sheet nad (min β nad) model. In either structure there are nine $\text{na } ^{15}\text{N}$ sites within 7 \AA of each labeled ^{13}CO but in the min nad model, some sites (e.g. 4-6 in **Figure 22c**) are “shared” between two ^{13}CO s, i.e. within 7 \AA of two ^{13}CO s. This reduces the average number of $\text{na } ^{15}\text{N}$ sites per ^{13}CO and the overall nad.

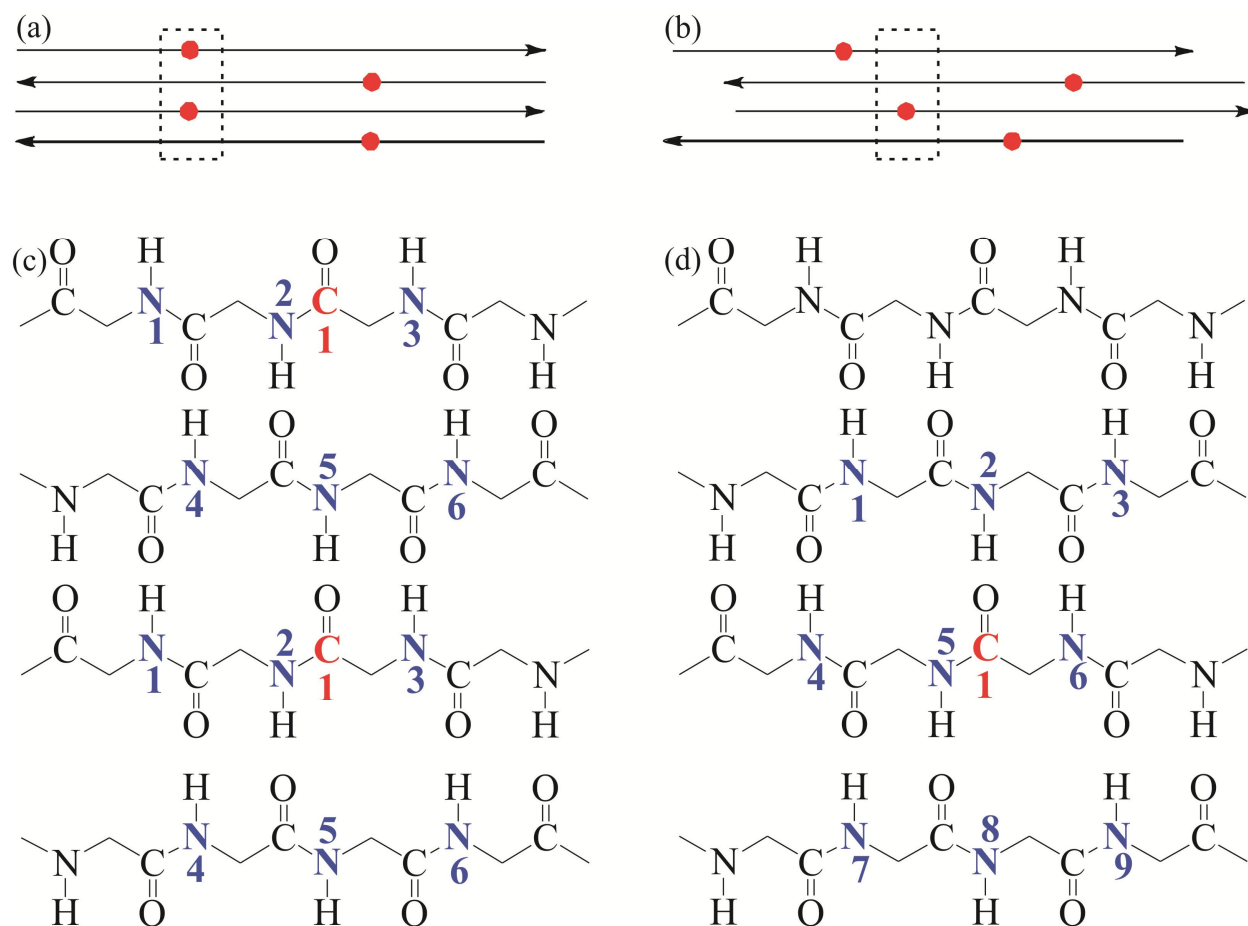


Figure 22. (a, b) Schematic diagrams of the HFP-F8 region of the HFP-NC sample in antiparallel β sheet structure with labeled ^{13}CO s represented as red circles. Panel a shows a model that is fully constrained to a single registry while panel b shows multiple registries. (c, d) β sheet backbone representations of the respective boxed regions of panels a and b with labeled ^{13}CO s in red and possible $\text{na } ^{15}\text{N}$ sites in blue, i.e. sites for which a $\text{na } ^{15}\text{N}$ is within 7 Å of a labeled ^{13}CO . A particular spin geometry will have only one ^{15}N . The min nad model is shown in panel c and each spin geometry will have either one labeled ^{13}CO and one $\text{na } ^{15}\text{N}$ (#1, 2 or 3) or two labeled ^{13}CO s and one $\text{na } ^{15}\text{N}$ (#4, 5, or 6). The max nad model is shown in panel d and each spin geometry will have one labeled ^{13}CO and one $\text{na } ^{15}\text{N}$.

Table 5. Chapter III indices and parameters

Index/parameter	Description	Values
f_t	fractional population of structure t	determined by fitting
j, k	index for n.a. site <7 Å from a labeled site: j , n.a. ^{15}N near labeled ^{13}CO ; k , n.a. ^{13}CO near labeled ^{15}N	
J, K	number of n.a. sites <7 Å from a labeled site: J , n.a. ^{15}Ns near labeled ^{13}CO ; K , n.a. ^{13}COs near labeled ^{15}N	α helical structure, $J = 9, K = 10$; min β sheet structure, $J = 6, K = 8$; max β 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
$S_l, S_{lj}, S_{lk}, S_{lu}$	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: p \equiv parallel registry a \equiv t = 16 or 17 antiparallel registry c \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}CO labeling and the ^{13}CO is hydrogen bonded to HN of top HFP	1 \equiv ^{13}CO HFP (top), ^{13}CO HFP (bottom) 2 \equiv ^{15}N HFP (top), ^{13}CO HFP (bottom) 3 \equiv ^{13}CO HFP (top), ^{15}N HFP (bottom) 4 \equiv ^{15}N HFP(top), ^{15}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: $w_1 = 1/81$, $w_2 = 2/81$, $w_3 = 2/81$, $w_4 = 4/81$
$\gamma_{IN}^{na}(\tau)$, $\gamma_{IC}^{na}(\tau)$	$S_l(\tau)/S_0$: $\gamma_{IN}(\tau)$, labeled ^{13}CO -n.a. ^{15}N pairs; $\gamma_{IC}(\tau)$, n.a. ^{13}CO -labeled ^{15}N pairs	$\gamma_{0N}(\tau) = 1$; $\gamma_{0C}(\tau) = 1$; $\gamma_{1N}(\tau)$ and $\gamma_{1C}(\tau)$ determined by calculation

Table 5 (cont'd).

$\gamma_{tuv}^{lab}(\tau)$, $\gamma_{t1t2uv}^{lab}(\tau)$	$S_{lu}^{lab}(\tau)/S_0^{lab}$ for arrangement of labeled ^{13}CO and ^{15}N nuclei: $\gamma_{tuv}^{lab}(\tau)$, fully constrained model; $\gamma_{t1t2uv}^{lab}(\tau)$, unconstrained model	$\gamma_{0tuv}^{lab}(\tau) = 1$; $\gamma_{0t1t2uv}^{lab}(\tau) = 1$; $\gamma_{1tuv}^{lab}(\tau)$ and $\gamma_{1t1t2uv}^{lab}(\tau)$ determined by calculation
τ	REDOR dephasing time	2.2, 8.2, 16.2, 24.2, 32.2, 40.2, or 48.2 ms
b	nad	1 = min β nad model 2 = max β nad model 3 = α helical nad model

There are many indices and parameters in the quantitative modeling and descriptions and possible values for them are compiled in **Table 5**. For the α , min β , or max β nad models, the average $\gamma_l = (S_l/S_0)$ for the relevant spin geometries:

$$\gamma_{lN}^{na}(\tau) = (J)^{-1} \sum_{j=1}^J \left[\frac{S_{lj}(\tau)}{S_{0j}(\tau)} \right]^{sim} \quad (38)$$

where “na” refers to natural abundance, $l = 0$ or 1 , “N” refers to na ^{15}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 α or max β nad and $J = 6$ for min β nad). The $\gamma_{0N}^{na}(\tau) = 1$ for all τ while $\gamma_{1N}^{na}(\tau)$ were calculated using $[S_{1j}(\tau)/S_{0j}(\tau)]^{sim}$ from the SIMPSON program and generally decreased with increasing τ . After setting a total labeled ^{13}CO population of 1.0 for HFP-F8 in HFP-NC, the relative population affected by na ^{15}N is $J \times 0.0037$ while the remainder population, $[1 - (J \times 0.0037)]$, has $S_1 = S_0$. There is also a na ^{13}CO contribution from unlabeled residues in HFP-F8

with a relative population 30×0.011 and with $S_1 = S_0$. Similar analysis for ^{15}N labeled HFP-A6L7 in HFP-NC results in:

$$\gamma_{lC}^{na}(\tau) = (K)^{-1} \sum_{k=1}^K \left[\frac{S_{lk}(\tau)}{S_{0k}(\tau)} \right]^{sim} \quad (39)$$

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

$$S_{l,tot}^{sim}(\tau) = \left\{ \left[J \times 0.0037 \times \gamma_{lN}^{na}(\tau) \right] + \left[2K \times 0.011 \times \gamma_{lC}^{na}(\tau) \right] \right\} + \{ 2.0 - [J \times 0.0037] - [2K \times 0.011] \} \quad (40)$$

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

For each nad model (α , min β , and max β), $(\Delta S/S_0)^{sim}$ for each τ was calculated with **Eq.**

(40) and statistical comparison was then made to $(\Delta S/S_0)^{exp}$:

$$\chi^2 = \sum_{m=1}^7 \left\{ \frac{\left(\left(\Delta S/S_0 \right)_m^{sim} - \left(\Delta S/S_0 \right)_m^{exp} \right)^2}{\sigma_m^{exp}} \right\} \quad (41)$$

where m is the index for an experimental datum, i.e. a particular τ . The respective χ^2 for the α , min β , and max β nad 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 nad calculation was supported by good fits for all models⁸³.

The broad ^{13}CO linewidth of HFP-NC in **Figure 20a** was consistent with two HFP populations, one with helical and one with β strand secondary structure. It was therefore reasonable to calculate $(\Delta S/S_0)^{sim}$ for “mixtures” with contributions from multiple models:

$$S_{l,mix}^{sim}(\tau) = \sum_{b=1}^3 \left[f_b \times S_l^{sim}(\tau) \right] \quad (42)$$

where $l = 0$ or 1 , b was the index that referred to the α , min β , or max β model, f_b was fractional population with $\sum f_b = 1$, and each $S_l^{sim}(\tau)$ was calculated using **Eq. (40)**. The HFP-NC distribution of ^{13}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 **Figure 21a** 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 nad 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 na sites within 7 Å of a labeled site which led to continually increasing $(\Delta S/S_0)$ with τ , **Figure 21a**. The $(\Delta S/S_0)^{sim}$ were also calculated using a short-range model that only considered na 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 $\chi^2 = 29$.

3. 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: (p) 1→17/1→17 and 2→17/1→16 parallel registries; (a) 16→1/1→16 and 17→1/1→17 antiparallel registries; and (c) X registries not detected by any of our labeling schemes, **Figure 19c**. Fraction p contributed to the $(\Delta S/S_0)^{exp}$ of HFP-P, fraction a contributed to the $(\Delta S/S_0)^{exp}$ of HFP-A, and fractions p and a 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, **Figure 21b**, and this analysis required calculation of the nad contribution to $(\Delta S/S_0)^{exp}$. Because a 1:2 ^{13}CO -HFP: ^{15}N -HFP ratio was used for all samples, this contribution was calculated using models developed for HFP-NC and resulted in a modified **Eq. (42)** appropriate for HFP-P, HFP-A, and HFP-AP:

$$S_{lu,tot}^{sim}(\tau) = \left\{ \left[J \times 0.0037 \times \gamma_{lN}^{na}(\tau) \right] + \left[2K \times 0.011 \times \gamma_{lC}^{na}(\tau) \right] \right\} + \left\{ 1.0 - (2K \times 0.011) \right\} + \left\{ [1.0 - (J \times 0.0037)] \times \gamma_{luv}^{lab}(\tau) \right\} \quad (43)$$

$$= S_l^{na}(\tau) + S_{lu}^{lab}(\tau) \quad (44)$$

where S_l^{na} is the sum of the first two braced terms in **Eq. (43)** 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 5** and **Figure 19b**. The first braced term in **Eq. (43)** corresponds to labeled and na ^{13}CO that experience nad, the second braced term corresponds to na ^{13}CO s that do not experience nad, and the third braced term corresponds to labeled ^{13}CO s that do not experience nad but may experience dephasing from labeled ^{15}Ns . The secondary structure of membrane-associated HFP was predominantly β sheet, **Figure 20b-d**, and the best estimates of the nad terms in the first braced term were taken to be the average of the max β and min β 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 $\gamma_{0tuv}^{lab}(\tau) = 1$, and $S_{0u}^{lab}(\tau) = 1.0 - (J \times 0.0037)$ while $\gamma_{1tuv}^{lab}(\tau)$ and therefore $S_{1u}^{lab}(\tau)$ were first calculated with a “fully constrained” model, **Figure 22a,c**, in which a β sheet region contained either: (a) $1 \rightarrow 17/1 \rightarrow 17$ or $2 \rightarrow 17/1 \rightarrow 16$ parallel registries; or (b) $16 \rightarrow 1/16 \rightarrow 1$ or $17 \rightarrow 1/1 \rightarrow 17$ antiparallel registries; or (c) X registries not directly detected by any of our labeling schemes, **Figure 19c**. A sample was considered to be a mixture of the three registry types each denoted by index $t = a, b$, or c and fractional population f_t , **Table 5**. The $S_{1u}^{lab}(\tau)$ was calculated by modified **Eq. (42)**:

$$S_{1u}^{lab}(\tau) = \sum_{t=a}^c \left[f_t \times \gamma_{1tuv}^{lab}(\tau) \right] \times S_0^{lab}(\tau) \quad (45)$$

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 , **Figure 19c**. For some combinations of t and u , all

labeled $r_{CN} > 7 \text{ \AA}$ with consequent $d_{CN} \sim 0$ and $\gamma_{1tuv}^{lab}(\tau) = 1$. Specific examples are $t = p$ and $u = 2$, $t = a$ and $u = 1$, and $t = c$ and $u = 1, 2$, or 3 . For other combinations of t and u , $\gamma_{1tuv}^{lab}(\tau)$ were determined from SIMPSON calculations and **Figure 23a-d** displays schematic examples for $t = a$, $u = 2$ with numerical values of $\gamma_{1tuv}^{lab}(\tau)$ and more details in the **Appendix VII**. Column a, b, c, or d corresponds to particular arrangements of ^{13}CO and ^{15}N labeled HFPs that are respectively denoted by the index $v = 1, 2, 3$, or 4 . For each v , the typical difference between the calculated $\gamma_{1a2v}^{lab}(\tau)$ for the $16 \rightarrow 1/1 \rightarrow 16$ or $17 \rightarrow 1/1 \rightarrow 17$ registry was ≤ 0.01 and the final $\gamma_{1a2v}^{lab}(\tau)$ were the average for the two registries. The antiparallel $\gamma_{1a3v}^{lab}(\tau)$ of HFP-AP were analogously calculated and the parallel $\gamma_{1p1v}^{lab}(\tau)$ of HFP-P and parallel $\gamma_{1p3v}^{lab}(\tau)$ of HFP-AP were calculated using the $1 \rightarrow 17/1 \rightarrow 17$ registry and had similar values to γ_s calculated using the $2 \rightarrow 17/1 \rightarrow 16$ registry. Fractional weightings w_v were based on the 1:2 ratio of ^{13}CO HFP: ^{15}N HFP with $w_1 = 1/9$, $w_2 = 2/9$, $w_3 = 2/9$, and $w_4 = 4/9$. A more complete version of **Eq. (45)**:

$$S_{1u}^{lab}(\tau) = \sum_{t=p,a,c} \left\{ f_t \times \sum_{v=1}^4 \left[w_v \times \gamma_{1tuv}^{lab}(\tau) \right] \right\} \times S_0^{lab}(\tau) \quad (46)$$

with indices and parameters summarized in **Table 5**.



Figure 23. Schematics of three adjacent HFPs for HFP-A, i.e. $u = 2$, in (a-d) fully constrained or



Figure 23 (cont'd). (e-h) unconstrained models. Red and blue correspond to ^{13}CO and ^{15}N labeled residues, respectively, and the labeled ^{13}CO in the middle strand was hydrogen bonded to the HN group of the residue in the top strand.

Figure 23 (cont'd). Panels a-d display antiparallel 16→1/1→16 (top) or 17→1/1→17 (bottom) registries while panels e-h display parallel 1→17/1→17, antiparallel 16→1/1→16, and X registries where X refers to a registry for which the labeled $r_{CN} > 7$ Å, i.e. beyond the approximate detection limit of the SSNMR experiment, and which is not 1→17/1→17, 2→17/1→16, 16→1/1→16, or 17→1/1→17. Correspondence between columns and the index ν are: a and e, $\nu = 1$; b and f, $\nu = 2$; c and g, $\nu = 3$; d and h, $\nu = 4$. Both rows of three-strand arrangements in panels a-d correspond to $t = b$ and the row, t_1 , t_2 correspondence in panels e-h is: row 1, $t_1 = p$, $t_2 = p$; row 2, $t_1 = p$, $t_2 = a$; row 3, $t_1 = p$, $t_2 = c$; row 4, $t_1 = a$, $t_2 = p$; row 5, $t_1 = a$, $t_2 = a$; row 6, $t_1 = a$, $t_2 = c$; row 7, $t_1 = c$, $t_2 = p$; row 8, $t_1 = c$, $t_2 = b$; row 9, $t_1 = c$, $t_2 = c$. 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 $t_2 = b$, fitting to experiment used γ s that were the average of those calculated with 16→1/1→16 and 17→1/1→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 $\gamma_{1t_1t_2uv}^{lab}(\tau) = 1$.

The values of f_p , f_a , and f_c were the same for the HFP-P, HFP-A, and HFP-AP samples with $f_c = 1 - f_p - f_a$. Best-fit values of f_p and f_a were obtained by calculating $\chi^2(f_p, f_a)$ using an expression analogous to **Eq. (41)**:

$$\chi^2(f_p, f_a) = \sum_{u=1}^3 \sum_{m=1}^7 \left\{ \frac{\left[\frac{\Delta S(f_p, f_a)}{S_0} \right]_{um}^{sim} - \left[\frac{\Delta S}{S_0} \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (47)$$

then selecting the f_p and f_a values which corresponded to minimum χ^2 , i.e. χ_{min}^2 . In **Eq. (47)**, m is the index for each τ , and $[\Delta S(f_p, f_a)/S_0]_{um}^{sim}$ was determined using **Eqs. (43)-(47)**. For this fully constrained model, **Figure 24a** displays a plot of χ^2 vs f_p and f_a with best-fit $f_p = 0.12$ and $f_a = 0.52$ and $\chi_{min}^2 = 11$. The model was reasonable as evidenced by χ_{min}^2 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_p fractional parallel population was small which was consistent with qualitative analysis of the data, **Figure 21b**. The f_a antiparallel population was substantially larger and also consistent with **Figure 21b**. The $f_c \approx 0.35$ indicated a substantial population of X registries not detected by the labeling of the three samples.

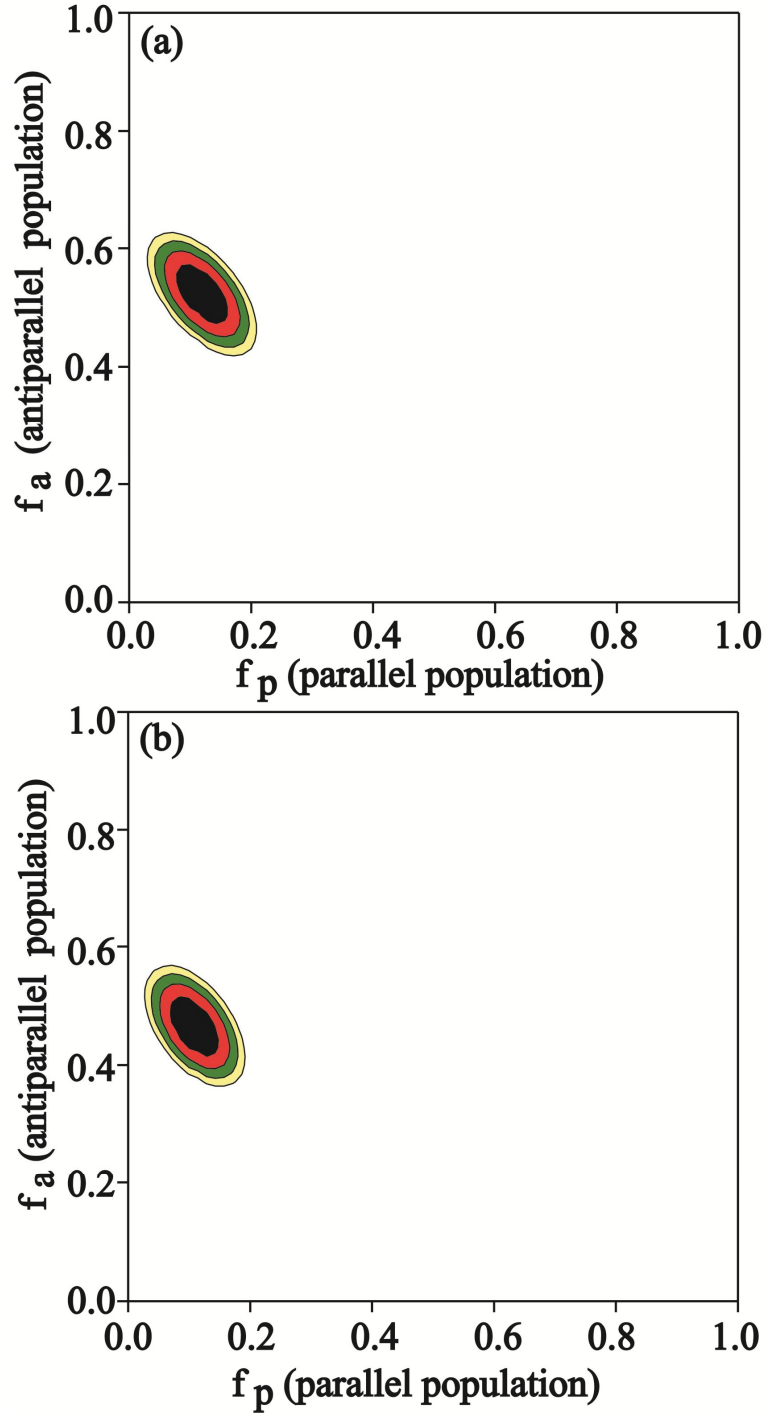


Figure 24. Contour plots of χ^2 vs f_a parallel and f_a antiparallel fractional populations for (a) fully constrained and (b) unconstrained models. In each plot, f_a is the sum of populations of $1 \rightarrow 17/1 \rightarrow 17$ and $2 \rightarrow 17/1 \rightarrow 16$ parallel registries and f_b is the sum of populations of $16 \rightarrow 1/1 \rightarrow 16$ and $17 \rightarrow 1/1 \rightarrow 17$ antiparallel registries.

Figure 24 (cont'd). For plot (a), the best-fit values were $f_a = 0.12 \pm 0.03$ and $f_b = 0.52 \pm 0.04$ with $\chi^2_{min} = 11$, and for plot (b), $f_a = 0.11 \pm 0.03$, $f_b = 0.46 \pm 0.04$, and $\chi^2_{min} = 8$. For plot (a), the best-fit values were $f_a = 0.12 \pm 0.03$ and $f_b = 0.52 \pm 0.04$ with $\chi^2_{min} = 11$, and for plot (b), $f_a = 0.11 \pm 0.03$, $f_b = 0.46 \pm 0.04$, and $\chi^2_{min} = 8$. Parameter uncertainties were determined by the populations with χ^2 within about three units of χ^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 $\chi^2 > 20$.

The above fitting was done using a long-range nad model that considered effects of na sites within 7 Å of each labeled nucleus. Fitting displayed in **Figure 24a** was based on nad calculated from half min β and half max β sheet structure, **Figure 22**, but the best-fit f_p and f_a and χ^2_{min} were not sensitive to the structural composition of the long-range nad model. For example, fitting done using nad for half α helical and half max β sheet structure yielded best-fit f_p and f_a and χ^2_{min} respectively within 0.01, 0.01, and 1 of the corresponding **Figure 24a** values. For HFP-NC fitting, nad was underestimated by a short-range model that only considered na 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 nad model and led to best-fit $f_p = 0.22$ and $f_a = 0.57$ which were significantly higher than the **Figure 24a** values. The $\chi^2_{min} = 20$ using the short-range model was also higher than the χ^2_{min} in **Figure 24a**.

4. 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: (p) 1→17/1→17 parallel registries; (a) 16→1/1→16 and 17→1/1→17 antiparallel registries; and (c) X registries not directly detected by any of our labeling schemes, **Figure 19c**. Each pairwise

registry type was labeled by $t = p, a, \text{ or } c$ with fractional population f_t . For this unconstrained model, **Figure 23e-h** displays schematics of three-strand registries with ^{13}CO 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}CO HFP/ ^{15}N HFP arrangement which respective label $v = 1, 2, 3, \text{ or } 4$. The ^{13}CO HFP: ^{15}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 $v = 2, 3, \text{ or } 4$ 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. (46)** was modified for the unconstrained model:

$$S_{lu}^{lab}(\tau) = \sum_{t_1=p,a,c} \sum_{t_2=p,a,c} \left\{ f_{t_1} \times f_{t_2} \times \sum_{v=1}^4 \left[w_v \times \gamma_{1t_1t_2uv}^{lab}(\tau) \right] \right\} \times S_0^{lab}(\tau) \quad (48)$$

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 **Figure 23e-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_p , f_a , and f_c in the unconstrained model were the same for the HFP-P, HFP-A, and HFP-AP samples with $f_c = 1 - f_p - f_a$. Best-fit values of f_p and f_a were obtained with **Eq. (47)** and **Figure 24b** displays a plot of χ^2 vs f_p and f_a with best-fit $f_p = 0.11$ and $f_a = 0.46$

and corresponding $\chi^2_{min} = 8$. The unconstrained model was reasonable as evidenced by a best-fit χ^2 which was smaller than the number of degrees of fitting, 19. The results were similar to the fully constrained model in that the f_p fractional parallel population was small, the f_a antiparallel population was large, and the $f_c \sim 0.4$ which suggest that there is a significant population of other structures.

This unconstrained model fitting was done with nad calculated with a long-range model and half min β and half max β structure. Similar to the fully constrained model, best-fit f_p, f_a , and χ^2 for the unconstrained model were: (1) negligibly affected by the structural distribution of the long-range nad model; and (2) significantly increased by use of a short-range nad model.

3.4 Discussion

Chapter III sets an upper limit of ~ 0.15 on the fraction of mHFP in in-register parallel β sheet structure, and this result is supported by both qualitative analysis of the data, **Figure 21b**, as well as quantitative analyses with fully constrained and unconstrained models, **Figure 24**. 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^{29,47,54,55}. The Chapter III 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. I believe that there was ambiguity of interpretation in the studies with 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 τ (24 ms), and with samples that contained equimolar amounts of HFP ^{13}C O labeled on three sequential residues and HFP ^{15}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 that also had some contribution from nad. 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 τ ; and (3) a “HFP-NC”-type sample was not studied and nad was therefore not quantitatively modeled.

The samples for the IR study were also extensively labeled with backbone ^{13}C O labeling at either: (1) Ala-1 to Val-3, Gly-5 to Leu-9; (2) Phe-8 to Gly-16; or (3) Ala-1 to Val-3, Gly-5 to Gly-16. The authors’ interpretation of their spectra to support predominant in-register parallel structure was based in part on expected effects of ($^{13}\text{C}=\text{}^{16}\text{O}$ electric dipole)···($^{13}\text{C}=\text{}^{16}\text{O}$ 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 it should be noted 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 (ν) of a $^{13}\text{C}=\text{}^{16}\text{O}$ vibration is sensitive to nearby (~ 5 Å away) $\text{C}=\text{O}$ vibrations and is higher with $^{12}\text{C}=\text{}^{16}\text{O}$ neighbors than with $^{13}\text{C}=\text{}^{16}\text{O}$ neighbors. If there is hydrogen bonding between ^{13}C O labeled residues of adjacent strands in an

undiluted sample, the corresponding diluted sample should have an increased fraction of $^{13}\text{C=O}/^{12}\text{C=O}$ proximities, decreased fraction of $^{13}\text{C=O}/^{13}\text{C=O}$ proximities, and $\Delta\nu = \nu_{\text{diluted}} - \nu_{\text{undiluted}} > 0$. If there were a major fraction of parallel 1 \rightarrow 17/1 \rightarrow 17 structure (as claimed by the authors), dilution of (1) Ala-1-Val-3, Gly-5-Leu-9; (2) Phe-8-Gly-16; or (3) Ala-1-Val-3, Gly-5-Gly-16 labeled HFPs would have had comparable effect on proximities and resulted in similar $\Delta\nu$. However, the experimental $\Delta\nu_{\text{AI-V3,G5-L9}} \approx \Delta\nu_{\text{F8-G16}} \approx (\Delta\nu_{\text{AI-V3,G5-G16}})/2$ 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.

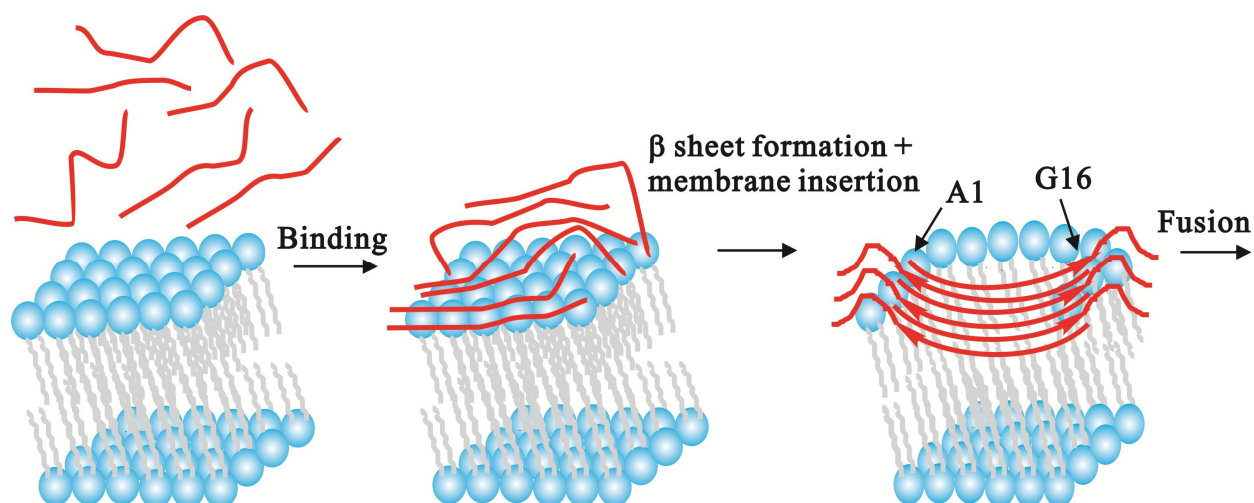


Figure 25. Pictorial model of HFP (red lines) binding to membranes followed by antiparallel β sheet formation and membrane insertion and then fusion. Time increases from left-to-right. For reasons of clarity, some lipids are not shown in the right-most picture. Although there are no data yet on fusion peptide structure during HIV/host cell fusion, the antiparallel β 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.

Figure 25 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^{30,32,33}. HFP sequentially: (1) binds to membranes; (2) forms β sheet oligomers with a significant fraction of 16 \rightarrow 1/1 \rightarrow 16 and 17 \rightarrow 1/1 \rightarrow 17 antiparallel registries; and (3) induces membrane fusion as monitored by intervesicle lipid mixing^{30,33,34,37}. It is also known that the Ala-6 and Leu-9 residues of β sheet HFP insert shallowly into the membrane with correlation between membrane insertion depth and both membrane perturbation and fusion rate^{40,61,92}. 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 **Figure 25** (right) shows a β 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 β sheet oligomers likely contain a small number of molecules because: (1) for ~90% of HFP molecules, there is an Ala-1 ^{13}C -lipid ^{31}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^{40,51}. 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 β sheet region (e.g. Ala-6 to Leu-12) 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 Leu-12→Ala-6/Ala-6→Leu-12, **Figure 19c**⁹³. A similar calculation yielded $\Delta G^{\text{insertion}} \sim -3$ kJ/mol for Leu-12→Gly-5/Gly-5→Leu-12 of the 16→1/1→16 registry. The $\Delta G^{\text{insertion}} \sim -6$ kJ/mol for Ala-6→Leu-12/Ala-6→Leu-12 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 $\epsilon_{dielectric} = 78$ and hexameric HFP, $\Delta G^{electrostatic} \approx +5.1$ kJ/mol for the 1→17/1→17 registry and +1.5 kJ/mol for the 17→1/1→17 registry.

It is expected that inclusion of the non-native C-terminal W(K)₆A 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 Leu-12 and random coil tag structure is supported by broad NMR linewidths in the C-terminal region of HFP^{37,40}. 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 ¹³CO and ¹⁵N labeled HFPs with or without the tag^{29,33}.

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 β 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^{52,53}. The width of a fibril is at most a few protein molecules and the length is >200 molecules and along the intermolecular β 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 seeds⁹⁴.

One distinctive feature of this chapter is the development of a quantitative nad 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 nad model which included effects of na nuclei <7 Å

from each labeled nucleus. For this model, the nad was approximately independent of secondary and tertiary structure. The nad was systematically underestimated by a short-range model which only considered na nuclei separated by one or two bonds from each labeled nucleus.

CHAPTER IV. Quantitative Identification of the Antiparallel Sheet Registry Distribution in Membrane Associated HFP and V2E-HFP Samples

4.1 Background.

Viral replication is initiated by infection of a host cell, and vaccines have been developed to build resistance to viral infections which has minimized the effects of diseases such as measles, mumps, and small pox to name a few^{2,93}. Relative to other viruses, sequences of HIV have higher mutation rates³, and there is not yet an HIV vaccine. Development of an HIV vaccine is critical since vaccines prevent infection, and inhibitory drug design is equally essentially for therapeutic treatment of infected patients. Development of inhibitory drugs would be enhanced by obtaining a deeper understanding of the HIV life cycle. The HIV life cycle's initial step in infection is membrane fusion, a joining of the HIV and the host cell membrane. The glycoproteins on the ectodomain of the HIV cell membrane are responsible for catalyzing membrane fusion where gp120 binds to a host cell receptor which is shed from gp41^{10,12}. Removal of gp120 from gp41 is believed to catalyze structural transitions within gp41 from a unknown structure to a hypothesized extended coiled-coil pre-hairpin intermediate structure (PHI) which eventually folds into a hairpin structure (see below)¹³. The gp41 protein is composed of ~356 residues¹⁴ and is subdivided into regions from the N-terminus: fusion peptide (FP) (~16 residues), FP proximal region (FPPR) (~13 residues), N-terminal helix (NTH) (~40 residues), loop (~47 residues), C-terminal helix (CTH) (~37 residues), pre-transmembrane region (~18 residues), transmembrane region (~28 residues)¹⁵, and cytoplasmic endodomain (~160 residues)¹⁴. To date, high-resolution crystal structures of gp41 were only for ectodomain

constructs that lacked the fusion peptide, transmembrane and endodomain regions. These ectodomain crystal structures showed protein trimers with three interior parallel α helical NTH segments and three exterior α helical CTH segments packed antiparallel to the NTHs. The overall structure of each monomer was a hairpin, and the trimer formed a six-helix bundle^{15-17,20}. The above domains have been defined by crystal structures of gp41 based constructs with varying interpretations of the number of residues incorporated into each domain. To my knowledge, the largest HIV gp41 crystallized construct to date has shown that the helicity extends beyond the traditionally defined NTH and CTH, **Figure 26**. The helicity of the NTH and CTH approximately extends from the respective residues Ala-532 to Ile-580 and Asp-627 to Asn-677. Additionally, the NTH and CTH in the SIV gp41 crystal structure span residues Arg-30 to Ala-86 and Thr-104 to Lys-146¹⁸, respectively, which is analogous to HIV gp41 crystal structure residues Arg-542 to Ser-598 and Ser-616 to Glu-662. One interpretation of these combined results is that under crystallization conditions, the helicity of the NTH and CTH regions are terminated due to the length of the gp41 construct rather than the length of the NTH and CTH of the full length gp41. Of note, these gp41 structures are for gp41 without the presence of a membrane, and these structures also lack the hydrophobic fusion peptide. For HIV gp41, mutations within the FP and FPPR have been shown to inhibit membrane fusion which suggests that both the FP and FPPR are important for membrane fusion²¹. Of special interest, transdominant inhibition of the V2E mutated gp41 (a FP mutation) has demonstrated that more than three gp41 or multiple gp41 trimers are needed to initiate membrane fusion²¹. Whatever the structure/s of the FP and FPPR are, the structure/s must allow for aggregation of FPs between more than 3 gp41 or multiple gp41 trimers. While the structures of the FP region of gp41 are the

focal point of this chapter and dissertation, it should also be noted that regions that are C-terminal of FP can also be effective targets for fusion inhibitor drugs. The AIDS drug Enfuvirtide²² is a fusion inhibitor and is a 36-residue peptide containing parts of the C-helix and the pre-transmembrane regions. Enfuvirtide likely binds to exposed N-helical regions in PHI gp41 and acts as a competitive inhibitor to the native C-helix regions with consequent prevention of the transition to the final hairpin structure. Cell-cell fusion mediated by gp41 includes in sequence: (1) lipid mixing between the membranes; (2) fusion pore formation; and (3) pore enlargement. After addition of Enfuvirtide, small pores rather than large pores can be closed which indicates that gp41 in the PHI state mediates lipid mixing and initial fusion pore formation while the final hairpin state (or possibly the PHI → hairpin transition) mediates pore stabilization and enlargement^{13,22,23}. Despite its efficacy, Enfuvirtide is not widely used because it is a peptide and must be administered by injection rather than orally²².



Figure 26. A summary of the gp41 sequence and regions defined from literature where FP = fusion peptide, FPPR = fusion peptide proximal region, NTH = N-terminal helix, CTH = C-terminal helix, MPER = membrane proximal region, and TMR = transmembrane region. This figure was adapted from literature¹⁵

The FP of gp41 is an alternative fusion inhibitor target and its membrane-associated structure is the subject of the present study. HIV and cellular studies have shown that point mutations within the FP lead to inhibition of fusion and infection^{21,28}. Various “HFP” constructs containing the ~16 residue FP have also been made and often catalyze fusion between membrane vesicles. There are good correlations between the effects on gp41-mediated fusion by specific FP

mutations and the effects on vesicle fusion by the corresponding mutations in HFP³¹. These collective data support FP (HFP) binding to the cell (vesicle) membrane early in the fusion process with fusion catalysis correlated to the membrane perturbation induced by the FP (HFP). The present study of the membrane-associated structures of HFP is important because knowledge of these structures will aid: (1) understanding of the FP as a fusion catalyst; and (2) the potential of the FP as a fusion inhibitor target.

Infrared (IR) spectroscopy and liquid-state NMR (LSNMR) have shown that HFP in aqueous solution adopts random coil structure^{34,37}. A fluorescence and IR study provided rates of: (1) HFP binding to membranes; (2) the transition to HFP β sheet structure; and (3) vesicle fusion. The study showed that binding to vesicles and β sheet formation occurred before vesicle fusion³⁴. Similar structures were observed by solid-state NMR (SSNMR): (1) HFP lyophilized from aqueous solution without vesicles had a distribution of secondary structures as indicated by single site backbone ¹³CO signals whose ~8 ppm linewidths spanned typical helical and β strand chemical shifts; and (2) HFP bound to hydrated membranes with ~30 mol% cholesterol was peptide oligomers/aggregates with β sheet structure. The cholesterol content of the membranes approximated that of the membranes of HIV and its host cells³⁶⁻³⁹. These data support β sheet HFP oligomers as an important structure in fusion.

Initial SSNMR and IR experiments with extensively labeled HFP respectively suggested that ~50% and ~100% of the membrane-associated oligomers had in-register parallel β sheet structure^{47,54}. In some contrast, Chapter III experiments with sparsely labeled HFP showed that in-register parallel structure accounted for $\leq 15\%$ of the molecules⁶³. The REDOR experiments

from Chapter III and literature³⁶ also identified significant populations for at least two distinct and specific antiparallel registries. The “HFPtr” construct has also been studied and contained three C-terminally cross-linked peptides that modeled the close contact of Thr-25 residues in the hairpin structure of the N-helix + C-helix regions of gp41. Relative to non-cross-linked HFP, HFPtr induced vesicle fusion with >15-fold faster rate^{25,30}. SSNMR showed that membrane-associated HFPtr had β sheet structure similar to that of HFP including lack of in-register parallel structure⁵⁴. The main structural distinction observed to-date between mHFP and mHFPtr was deeper membrane insertion of mHFPtr^{25,54,95}. Other data have also supported small β sheet oligomers as a structural model for FP in gp41: (1) electron micrographs of virus-cell contacts showed multiple nearby gp41 trimers in contact with the cell membrane and (2) fusion and HIV infection were dominantly inhibited in samples containing a small fraction of gp41 with a V2E point mutation and a large fraction of wild-type gp41^{21,96,97}.

Efforts to identify the predominant β sheet registry of mHFP have led to conflicting β sheet structural models such as in-register parallel⁴⁷, mixed in-register parallel and antiparallel²⁹, and β hairpin⁴⁸. As described above, Chapter III showed that in-register parallel β sheets are minimally populated while the present Chapter quantitatively describes the distribution of antiparallel registries and accounts for nearly all of the mHFPs. Each registry was considered as a β sheet composed of residues 1 to t of the HFP molecule. The $t = 22, 18$ and 13 registries are illustrated in **Figure 28**. In the present Chapter, many HFPs were prepared that differed in the residue numbers of their single ¹³CO and single ¹⁵N backbone labels (See Chapter II, **Table 1**). The sample index $u = \zeta + \eta - 1$ and ζ and η were respectively the residue

numbers of the ^{13}CO and ^{15}N labels. The intermolecular labeled ^{13}CO - ^{15}N distance was denoted r_{CN} and for sample u , the smallest r_{CN} of $\sim 4 \text{ \AA}$ was observed for molecules with registry $t = u$ and an interpeptide $^{13}\text{CO}\cdots\text{H}^{15}\text{N}$ hydrogen bond. In the solid-state NMR (SSNMR) rotational-echo double-resonance (REDOR) experiments of the present study, this $^{13}\text{CO}\cdots^{15}\text{N}$ proximity was qualitatively and quantitatively understood by the extent of ^{13}CO signal attenuation. The most important and novel result was significant ^{13}CO signal attenuation from many different samples; i.e. there was a broad distribution of antiparallel registries. Additionally, similar experiments were run for mV2E-HFP. The mV2E-HFP samples displayed a distribution of antiparallel registries, but the distribution of registries was distinctly different than that of mHFP. These distinct differences in the registry distributions have led to development of new hypotheses to explain the previously observed transdominant inhibition of fusion that was observed when wild type gp41 was co-expressed with dilute amounts of V2E mutant gp41²¹.

4.2 Materials and Methods

Table 6. Chapter IV indices and parameter

Index/parameter	Description	Values
f_t	fractional population of structure t	determined by fitting
j, k	index for n.a. site <7 Å from a labeled site: j , n.a. ^{15}N near labeled ^{13}CO ; k , n.a. ^{13}CO near labeled ^{15}N	
J, K	number of n.a. sites <7 Å from a labeled site: J , n.a. ^{15}Ns near labeled ^{13}CO ; K , n.a. ^{13}COs near labeled ^{15}N	min β sheet structure, $J = 6, K = 6$; max β sheet structure, $J = 9, K = 9$
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 correspond to 2.2, 8.2, 16.2, 24.2, 32.2, 40.2 and 48.2 ms dephasing time, respectively.
$S_l, S_{lj}, S_{lk}, S_{lu}$	REDOR signal intensity	determined by experiment or calculation
u	membrane associated sample index $u = \zeta + \eta - 1$ ζ = residue number for ^{13}CO labeled η = residue number for ^{15}N labeled	$u = \zeta + \eta - 1$ ζ = residue number for ^{13}CO labeled η = residue number for ^{15}N labeled 8 \equiv HFP-A6CG3N, 9 \equiv HFP-L7CG3N, etc.
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	7-25, and X where 8 \equiv 8 residue antiparallel registry, X $\equiv t \neq (u-1), u$, or $(u-1)$
$\gamma_{IN}^{na}(\tau), \gamma_{IC}^{na}(\tau)$	$S_l(\tau)/S_0$: $\gamma_{IN}(\tau)$, labeled ^{13}CO -n.a. ^{15}N pairs; $\gamma_{IC}(\tau)$, n.a. ^{13}CO -labeled ^{15}N pairs	$\gamma_{0N}(\tau) = 1; \gamma_{0C}(\tau) = 1; \gamma_{1N}(\tau)$ and $\gamma_{1C}(\tau)$ determined by calculation
$\gamma_{ltu}^{lab}(\tau), \gamma_{lt_1t_2u}^{lab}(\tau)$	$S_{lu}^{lab}(\tau)/S_0^{lab}$ for arrangement of labeled ^{13}CO and ^{15}N nuclei: $\gamma_{1tu}^{lab}(\tau)$, fully constrained model; $\gamma_{1t_1t_2u}^{lab}(\tau)$, unconstrained model	$\gamma_{0tu}^{lab}(\tau) = 1; \gamma_{0t_1t_2u}^{lab}(\tau) = 1; \gamma_{1tu}^{lab}(\tau)$ and $\gamma_{1t_1t_2u}^{lab}(\tau)$ determined by calculation

Table 6 (cont'd).

τ	REDOR dephasing time	2.2, 8.2, 16.2, 24.2, 32.2, 40.2, or 48.2 ms
b	Structural population index	1= min n.a.d. model 2= max n.a.d. model

The data analysis is adapted from literature⁶³ and similar to Chapter III. **Table 6** contains a list of all the indices and parameters. As mentioned in Chapter III, REDOR was used to probe dipolar coupling (d_{CN}) between ^{13}C O and ^{15}N nuclei⁷⁸ where $d_{\text{CN}} = 3080 \text{ Hz-}\text{\AA}^3/r_{\text{CN}}^3$. The average d_{CN} were quantified by comparing the ^{13}C O signal intensity of the “S₀” spectra, S_0 , to the generally reduced intensity of the “S₁” spectra, S_1 , as a function of dephasing time, τ . Signal attenuation was quantified by the normalized dephasing $(\Delta S/S_0) = (S_0 - S_1)/S_0$ where S_0 and S_1 were quantified by 3.0 ppm integrations about the ^{13}C O peak center of the S_0 and S_1 spectra, respectively. These 3.0 ppm regions about the peak center generally encompassed a majority of the ^{13}C O signal. Integration of both larger and smaller regions (1.0 ppm and 5.0 ppm) generally yielded similar results (± 0.01 variation $\Delta S/S_0$). In modeling the experimental data, $r_{\text{CN}} \leq 7 \text{ \AA}$ were responsible for most of the dephasing, and $r_{\text{CN}} > 7 \text{ \AA}$ were not considered in the analysis. The SIMPSON program was used to simulate $\gamma(\tau)$ for ^{13}C O- ^{15}N spin geometries. These spin geometries were based upon d_{CN} and orientations between d_{CN} vectors derived from the crystal structure coordinates of the β barrel outer membrane protein G (OMPG) (PDB file 2IWW). Input parameters included d_{CN} s as well as Euler angles for each d_{CN} vector and for the ^{13}C O chemical shift anisotropy (CSA) principal axis system (PAS) in a fixed crystal frame^{84,87}.

The (α, β, γ) Euler angles of d_{CN} PAS were $(0, 0, 0)$ and $(0, \vartheta, 0)$ for 2 and 3 spin simulations, respectively, where ϑ was the angle between two d_{CN} vectors. The ^{13}CO CSA principal values of 247, 176, and 99 ppm were also inputs to the SIMPSON calculations⁸⁹. The $\gamma_{\text{IC}}^{\text{na}}(\tau)$ and $\gamma_{\text{IN}}^{\text{na}}(\tau)$ were calculated for na ^{13}CO -labeled ^{15}N and labeled ^{13}CO -na ^{15}N for $r_{\text{CN}} \leq 7 \text{ \AA}$. Each membrane-associated sample was labeled by an index u where $u = \zeta + \eta - 1$ (the ^{13}CO labeled residue number minus the ^{15}N labeled residue number minus 1). For each sample, u , the $t = u-1$, u , and $u+1$ registries resulted in labeled ^{13}CO -labeled ^{15}N interstrand $r_{\text{CN}} \leq 7 \text{ \AA}$. These registries resulted in the most significant signal attenuation in S_1 relative to S_0 and $\gamma_{\text{ltu}}^{\text{lab}} \neq 1$. For all other registries (in total denoted X), labeled interstrand $r_{\text{CN}} > 7 \text{ \AA}$, and we approximated $\gamma_{0\text{tu}}^{\text{lab}} = \gamma_{\text{ltu}}^{\text{lab}} = 1$. Intersheet distances were neglected because the typical intersheet $r_{\text{CN}} > 7 \text{ \AA}$ ⁹⁸⁻¹⁰⁰.

Simulated signal was defined as:

$$S_{\text{lu}}^{\text{sim}}(\tau) = \left\{ \left[0.99 \times J \times 0.0037 \times \gamma_{\text{IN}}^{\text{na}}(\tau) \right] + \left[0.98 \times K \times 0.011 \times \gamma_{\text{IC}}^{\text{na}}(\tau) \right] \right\} + \left\{ 0.33 - (K \times 0.011) \right\} + \left\{ \left[0.99 - (J \times 0.0037) \right] \times \left[f_X + \sum_{t=u-1}^{u+1} f_t \gamma_{\text{ltu}}^{\text{lab}}(\tau) \right] \right\} \quad (49)$$

The first braced term in **Eq. (49)** corresponds to labeled and na ^{13}CO that experience *nad*, the second braced term corresponds to na ^{13}CO s that do not experience *nad*, and the third braced term corresponds to labeled ^{13}CO s that do not experience *nad* but may experience dephasing from labeled ^{15}N s where $f_X = 1 - f_{t=u-1} - f_{t=u} - f_{t=u+1}$. The K and J were estimated to be 7.5 and were derived from the *nad* model used in Chapter III (0.5 max *nad* model and 0.5 min *nad*).

$S_{0u}^{sim}(\tau)$ were calculated using $\gamma_{0tu}^{lab}(\tau) = \gamma_{0N}^{na}(\tau) = \gamma_{0C}^{na}(\tau) = 1$ while $S_{1u}^{sim}(\tau)$ were calculated using $\gamma_{1tu}^{lab}(\tau)$ from the fully constrained model and fractional populations, f_t 's. The fractional populations were varied from 0.00 to 1.00 for $7 < t < 25$ where $f_t = 0.00$ for $t < 8$ and $t > 24$. Similar to Chapter III, f_t values were evaluated by using a χ^2 metric, but the larger range of t values considered in this Chapter required minimization of the searching space for each f_t to reduce the computational time associated with a global fitting of 17 variables (**Eq (51)** below). For each t , the maximum f_t , f_t^{max} , was obtained using the $(\Delta S / S_0)^{exp}_{u=t}$ data, and the minimum $\chi^2_{u=t}$ was determined under the constraints $f_{t=u-1} = f_{t=u+1} = 0$ with variation of $f_t = u$ from 0.00 to 1.00 by 0.01 increments by **Eq. (50)**.

$$\chi^2_u = \sum_{m=1}^7 \left\{ \frac{\left[\frac{\Delta S(f_{t=u-1}, f_{t=u}, f_{t=u+1})}{S_0} \right]_{um}^{sim} - \left[\frac{\Delta S}{S_0} \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (50)$$

where m is the index for each τ and σ_{um}^{exp} is the rms deviation determined by twelve 3 ppm integrations of noise regions of both the S_0 and S_1 spectra⁸³. At the minimum $\chi^2_{u=t, f_t=u}^{max}$ was set equal to $f_t = u$ for each sample, u . Similarly, each minimum f_t, f_t^{min} , was obtained by searching for the minimum $\chi^2_{u=t}$ under the constraints $f_{t=u-1} = f_{t=u-1}^{max}$ and $f_{t=u+1} = f_{t=u+1}^{max}$ with variation of $f_t = u$ from 0.00 to 1.00 by 0.01 increments. The $f_t = u^{min}$ was the value of $f_t = u$ which yielded minimum $\chi^2_{u=t}$. The best-fit $f_8 - f_{24}$ were obtained by fitting all $(\Delta S$

$/S_0)^{exp}$ data for $u = 8$ to 24 by **Eq (51)**. For this fitting, each f_t was varied from f_t^{min} to f_t^{max} by 0.01 increments.

$$\chi^2(f_{8-24}) = \sum_{u=8}^{24} \sum_{m=1}^7 \left\{ \frac{\left[\frac{\Delta S(f_{t=u-1}, f_{t=u}, f_{t=u+1})}{S_0} \right]_{um}^{sim} - \left[\frac{\Delta S}{S_0} \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (51)$$

The best-fit f_t 's were those corresponding to the global minimum $\chi^2(f_{8-24})$. This method was used for determining the best-fit f_t 's for both mHFP and mV2E-HFP. Each f_t^{min} , f_t^{max} and best-fit f_t are found in **Table 7** (see Results). These fittings are generally referred to as 3 registry fittings while the equations were modified for the 5 registry fittings, **Appendix VIII**, and the results are found in **Table 8**.

To evaluate the reproducibility of these experiments, a second HFP-F8CG13N was synthesized, and incorporated into a new vesicle sample and new $(\Delta S/S_0)^{exp}$ data were collected. The $(\Delta S/S_0)^{exp}$ data typically varied by < 0.01 between similar data points, τ , of the two HFP-F8CG13N samples, **Table 9**. The new HFP-F8CG13N $(\Delta S/S_0)^{exp}$ data were substituted for the old $(\Delta S/S_0)^{exp}$ data and a second fitting was performed with resultant $f_{19} = 0.02, f_{20} = 0.17, f_{21} = 0.00$ for a fully constrained model at the $\chi^2_{min} = 119$. The second fitting yielded a χ^2_{min} that was ~ 9 higher than the first fitting, but the $\sum_{t=19}^{21} f_t$ differed by only 0.02 between the two fittings. The reproducibility of the $(\Delta S/S_0)^{exp}$ data and best-fit f_t 's validate this data analysis whereas variation

in the χ^2_{\min} associated with the second fitting was due in part to the smaller $\sigma_{20,m}^{\text{exp}}$ s associated with the second data set, **Table 9**. An unconstrained iterative 3 registry fitting was also considered which yielded similar results, see **Appendix IX**.

4.3 Results.

1. Fully Constrained Model mHFP Registry Distribution

Table 7. Three registry fittings for mHFP and mV2E-HFP.

mHFP					mV2E-HFP			
f_t	f_t^{\min}	f_t^{\max}	Best-fit f_t	χ^2	f_t^{\min}	f_t^{\max}	Best-fit f_t	χ^2
f_8	0.00	0.00	0.00	1.4	0.00	0.00	0	4.0
f_9	0.00	0.01	0.00	3.0	0.00	0.00	0	4.0
f_{10}	0.00	0.02	0.01	3.9	0.00	0.00	0	2.7
f_{11}	0.01	0.07	0.04	2.0	0.00	0.00	0	8.1
f_{12}	0.03	0.15	0.07	6.4	0.00	0.03	0.01	6.2
f_{13}	0.08	0.22	0.16	7.8	0.00	0.05	0.05	5.0
f_{14}	0.00	0.19	0.06	9.5	0.00	0.09	0.01	6.4
f_{15}	0.06	0.23	0.15	3.1	0.00	0.17	0.12	4.2
f_{16}	0.04	0.25	0.12	3.4	0.04	0.26	0.13	3.8
f_{17}	0.10	0.28	0.18	5.3	0.05	0.33	0.18	4.8
f_{18}	0.05	0.21	0.13	6.8	0.10	0.36	0.26	3.8
f_{19}	0.00	0.14	0.02	9.4	0.00	0.33	0.03	3.1
f_{20}	0.11	0.18	0.15	8.2	0.28	0.49	0.44	5.0
f_{21}	0.00	0.04	0.00	9.8	0.05	0.24	0.06	10.9
f_{22}	0.00	0.04	0.01	10.6	0.00	0.08	0.01	8.6
f_{23}	0.04	0.05	0.04	6.0	0.03	0.08	0.07	3.3
f_{24}	0.00	0.00	0.00	13.8	0.02	0.04	0.02	3.2

* The $\chi^2_{u=t}$ were calculated using **Eq 45** where the best-fit f_t values were used for $f_t = u-1, f_t = u$, and $f_t = u+1$.

For mHFP, the global fitting resulted in $\sum_{t=8}^{24} f_t = 1.14$ with $\chi^2_{\min} = 110$. The $\sum_{t=8}^{24} f_t^{\min} = 0.52$ and $\chi^2 = 1090$. The $\sum_{t=8}^{24} f_t^{\max} = 2.08$ $\chi^2 = 2110$.

For mV2E-HFP, the global best-fit resulted in $\sum_{t=8}^{24} f_t = 1.39$ with $\chi^2_{\min} = 87$. The $\sum_{t=8}^{24} f_t^{\min} = 0.57$ with $\chi^2 = 2071$. The $\sum_{t=8}^{24} f_t^{\max} = 2.55$ with $\chi^2 = 4444$.

Table 8. Five registry fittings for mHFP and mV2E-HFP.

mHFP					mV2E-HFP			
f_t	f_t^{\min}	f_t^{\max}	Best-fit f_t	χ^2	f_t^{\min}	f_t^{\max}	Best-fit f_t	χ^2
f_8	0.00	0.00	0.00	1.4	0.00	0.00	0.00	4.0
f_9	0.00	0.01	0.00	3.0	0.00	0.00	0.00	4.0
f_{10}	0.00	0.02	0.01	3.9	0.00	0.00	0.00	2.7
f_{11}	0.00	0.07	0.04	2.0	0.00	0.00	0.00	8.2
f_{12}	0.02	0.15	0.07	6.4	0.00	0.03	0.00	6.5
f_{13}	0.07	0.22	0.16	7.8	0.00	0.05	0.05	5.0
f_{14}	0.00	0.19	0.06	9.5	0.00	0.09	0.01	6.7
f_{15}	0.03	0.23	0.15	3.1	0.00	0.17	0.09	2.9
f_{16}	0.02	0.25	0.12	3.4	0.04	0.26	0.14	4.7
f_{17}	0.07	0.28	0.18	5.3	0.03	0.33	0.14	5.1
f_{18}	0.03	0.21	0.13	6.8	0.05	0.36	0.25	4.7
f_{19}	0.00	0.14	0.02	9.4	0.00	0.33	0.01	4.3
f_{20}	0.09	0.18	0.15	8.2	0.24	0.49	0.44	3.9
f_{21}	0.00	0.04	0.00	9.8	0.00	0.24	0.05	10.0
f_{22}	0.00	0.04	0.01	10.6	0.00	0.08	0.00	4.4
f_{23}	0.03	0.05	0.04	6.0	0.01	0.08	0.07	2.7
f_{24}	0.00	0.00	0.00	13.8	0.00	0.04	0.01	2.9

* The χ^2 were calculated using **Eq 60 (Appendix VIII)** where the best-fit f_t values were used for $f_{t=u-2}$, $f_{t=u-1}$, $f_{t=u}$, $f_{t=u+1}$, and for $f_{t=u+2}$.

For mHFP, the global fitting resulted in $\sum_{t=8}^{24} f_t = 1.05$ with $\chi^2_{\min} = 111$. The $\sum_{t=8}^{24} f_t^{\min} = 0.36$ and $\chi^2 = 1561$. The $\sum_{t=8}^{24} f_t^{\max} = 2.08$ $\chi^2 = 2738$.

For mV2E-HFP, the global best-fit resulted in $\sum_{t=8}^{24} f_t = 1.26$ with $\chi^2_{\min} = 83$. The $\sum_{t=8}^{24} f_t^{\min} = 0.37$ with $\chi^2 = 2929$. The $\sum_{t=8}^{24} f_t^{\max} = 2.55$ with $\chi^2 = 5740$.

Table 9. HFP-F8CG13N data sets used in global fittings

τ (ms)	1 st Fitting		2 nd Fitting	
	$(\Delta S/S_0)^{exp}$	σ^{exp}	$(\Delta S/S_0)^{exp}$	σ^{exp}
48.2	0.175	0.015	0.206	0.011
40.2	0.177	0.024	0.170	0.012
32.2	0.161	0.012	0.162	0.011
24.2	0.116	0.015	0.108	0.010
16.2	0.068	0.010	0.066	0.010
8.2	0.017	0.012	0.0326	0.010
2.2	0.022	0.012	0.018	0.010

For a sample u , the $t = u-1$, u and $u+1$ registries resulted in interstrand labeled ^{13}CO -labeled ^{15}N $r_{\text{CN}} < 7$ Å while the corresponding intrastrand r_{CN} were always much greater than 7 Å. A small fraction of the labeled (natural abundance) ^{13}CO s were also close to natural abundance (labeled) ^{15}Ns , and this contribution to S_1 signal attenuation is detailed below. If all pairs of HFP molecules had a specific antiparallel registry t , then each labeled ^{13}CO in a $u = t$ sample would have a $r_{\text{CN}} \sim 4.0$ Å and a $(\Delta S/S_0)^{exp} \sim 1$ for $\tau = 48$ ms. Additionally, the $u = t - 1$ and $u = t + 1$ samples would have significant dephasing whereas other samples would have minimal dephasing. If only a fraction of the molecules have $t = u-1$, u and $u+1$ registries, the $(\Delta S/S_0)^{exp}$ at $\tau = 48$ ms would be reduced for this sample. A $(\Delta S/S_0)^{exp} \sim 0.0-0.1$ at $\tau = 48$ ms implies that $f_t \sim 0.00$, and comparisons between $(\Delta S/S_0)^{exp}$ can be used as a semi-quantitative measure of the fractional population of t registry, f_t . Detection of significant $(\Delta S/S_0)^{exp}$ for many different samples in the $u = 8 - 24$ range, **Figure 27**, demonstrated that there was a broad

distribution of antiparallel β sheet registries rather than a single predominant registry. Only β sheet structures were considered because of the observed β sheet ^{13}C chemical shift distributions.

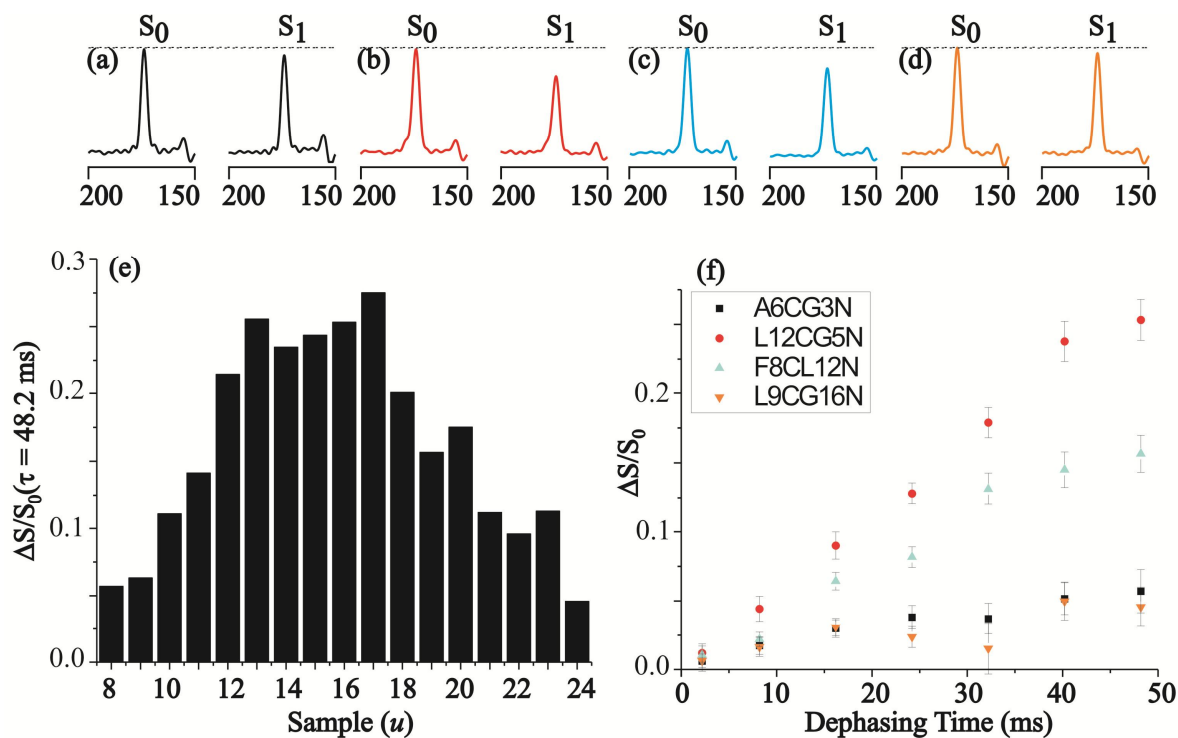


Figure 27. REDOR S_0 and S_1 ^{13}C SSNMR spectra at 48.2 ms dephasing time for (a) mHFP-A6CG3N, (b) HFP-L12CG5N, (c) HFP-F8CL12N, or (d) HFP-L9CG16N. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 23766; (b) 21454; (c) 40331; or (d) 39133 scans. (e) $(\Delta S/S_0)^{\text{exp}}$ ($\tau = 48.2$ ms) for all samples, u . (f) Plots of $(\Delta S/S_0)^{\text{exp}}$ vs dephasing time with the rms error. Isotopic labeling of each mHFP is displayed in the legend that correspond to HFP-A6CG3N (black, square), HFP-L12CG5N (red, circle), HFP-F8CL12N (cyan, triangle), and (d) HFP-L9CG16N (orange, inverted triangle). Variation less than ± 0.02 in $(\Delta S/S_0)^{\text{exp}}$ was also observed between two different preparations of the same sample type, e.g. HFP-F8CG13N (not displayed here).

Quantitative analysis of this distribution considered that the S_0 and S_1 signal intensities each contained contributions from labeled and natural abundance (na) ^{13}C nuclei. Signal attenuation in S_1 resulted from natural abundance dephasing (nad) and registry dependent

dephasing (*rdd*). The *nad* was due to *na* ^{13}CO -labeled ^{15}N and labeled ^{13}CO -*na* ^{15}N proximities. For sample u , the *rdd* of the labeled ^{13}CO -labeled ^{15}N interstrand proximities resulted from the $t = u-1$, $t = u$, and $t = u+1$ registries where the corresponding fractional populations of these registries were denoted $f_{t=u-1}$, $f_{t=u}$, and $f_{t=u+1}$. The SIMPSON software was used to calculate $(S_1/S_0)^{sim} = \gamma^{sim}_l$ for specific β sheet structure spin geometries (see Materials and Methods and **Appendix XVI**) over the range of experimental dephasing times τ . Such calculations were done for: (1) different labeled ^{13}CO -*na* ^{15}N spin geometries with resultant γ^{na}_l C's; and (2) different *na* ^{13}CO -labeled ^{15}N spin geometries with resultant γ^{na}_{lN} 's. Consideration of the fractional populations of *na* ^{13}CO and *na* ^{15}N then allowed calculation of $(\Delta S/S_0)^{nad}$, i.e. the expected dephasing if there were no *rdd*. If the calculation of $(\Delta S/S_0)^{nad}$ were accurate, the $(\Delta S/S_0)^{exp} = (\Delta S/S_0)^{nad}$ for u samples with $f_{t=u-1} = f_{t=u} = f_{t=u+1} = 0.00$. The smallest $(\Delta S/S_0)^{exp}$ for mHFP were derived from the $u = 8$ and the $u = 28$ samples and were considered to only be due to *nad*, i.e. and $f_7 = f_8 = f_9 = 0.00$ and $f_{27} = f_{28} = f_{29} = 0.00$. Comparison of the $(\Delta S/S_0)^{exp}$ to $(\Delta S/S_0)^{nad}$ yielded $\chi^2 = 1.1$ for $u = 8$ and $\chi^2 = 1.8$ for $u = 28$ with 7 data points for each sample. Similarly, the mV2E-HFP $(\Delta S/S_0)^{exp}$ from $u = 8$ sample were compared to $(\Delta S/S_0)^{nad}$, $f_7 = f_8 = f_9 = 0.00$, which yielded $\chi^2 = 4.0$ for 7 data points for each sample. These combined results and the Chapter III results supported the accuracy and adaptability of the $(\Delta S/S_0)^{nad}$ calculation for different membrane-associated samples.

The γ_{ltu}^{lab} were calculated for the registries that aligned labeled ^{13}CO and labeled ^{15}N nuclei within 7 Å, $t = u-1$, $t = u$, and $t = u+1$ registries. Because the γ_{ltu}^{lab} only depended upon generalized spin geometries that were derived from crystal structures (see Materials and Methods), equivalent γ_{ltu}^{lab} values were used for all samples, u , and different γ_{ltu}^{lab} values were used for each generalized spin geometry for the $t = u-1$, $t = u$, and $t = u+1$ registries. Qualitative analysis of the $(\Delta S/S_0)^{exp}$ at $\tau = 48$ ms, **Figure 27e**, suggests that mHFP forms a distribution of antiparallel registries, and a “fully-constrained” model was used for quantitative fitting of the registry populations for which there was only a single antiparallel registry within a local region of the β sheet. Fully constrained oligomers likely pack with minimal void space that could be energetically favorable if the magnitude of the negative enthalpic contribution to the free energy from local contacts is greater than the positive entropic contribution to the free energy from forming a single rather than multiple registries¹⁰¹. It is noted that for mature amyloid fibrils formed in aqueous solution, structural data have been interpreted with a model of a single β sheet registry, i.e. the fully constrained model¹⁰²⁻¹⁰⁵. It is also noted that the number of molecules within a fibril are likely much greater than in mHFP oligomers since most Ala-1 are in close contact with the lipid headgroup of the membrane. This would not occur in larger amyloid fibrils where β sheet stacking would reduce the number of Ala-1 carbonyl in close contact (~ 5 Å) with the phosphorus of the lipid headgroup¹⁰⁶.

For a sample u , the calculated $\Delta(S/S_0)^{sim}$ depended on $f_{t=u-1}$, $f_{t=u}$ and $f_{t=u+1}$ but not on other f_t 's, whose fractional population sum is described by $f_X = 1 - f_{t=u-1} - f_{t=u} - f_{t=u+1}$. As

described in the Materials and Methods, sets of distinct $(\Delta S/S_0)^{sim}$ for any sample u were calculated using: (1) γ^{na}_{IC} 's, and γ^{na}_{IN} 's from *nad* spin geometries along with the probabilities of these geometries; (2) γ^{lab}_{ltu} 's from *rdd* spin geometries; and (3) specific values of $f_{t=u-1}$, $f_t = f_u$ and $f_{t=u+1}$. There were 119 $(\Delta S/S_0)^{exp}$ data from all samples, and these data were globally fit to the f_t 's with t in the 8 – 24 range, see **Materials and Methods**. The best-fit f_t 's, **Table 7**, were correlated with a statistically reasonable $\chi^2 = 110$ which was comparable to the number of degrees of freedom of the fit (102). The sum of best-fit f_t 's, $\sum_{t=8}^{24} f_t = 1.14$, was close to 1, and it is reasonable to approximate the excess to be proportionally distributed among the f_t 's so that $f_t(\text{true}) = f_t(\text{best-fit})/1.14$. In addition to the aforementioned “fully constrained” model with a single local registry, an alternate “unconstrained” model was also considered for which there could be multiple registries within a local oligomeric region⁶³. Similar best-fit χ^2 and f_t 's were obtained with this model, see **Appendix IX**.

The distribution of registries observed in mHFP appears to be close to the thermodynamic equilibrium state. Evidence to support this view is (1) the reproducibility of the $(\Delta S/S_0)^{exp}$ data for two different HFP-F8CG13N samples; (2) samples with similar chemical shifts and line shapes were prepared in aqueous solution (this dissertation) and by organic co-solubilization followed by removal of organic solvent and hydration³⁹. The membrane/HFP interactions affect the thermodynamic equilibrium state of mHFP as evidenced by mixed secondary structures for HFP lyophilized without membranes and membrane-inserted β sheet

structure for mHFP²⁶. This Chapter shows that mHFP has a distribution of β sheet registries, and previous work has demonstrated that many factors contribute toward the free energy of any membrane inserted structure¹⁰⁷. The membrane insertion free energy for each of the twenty naturally occurring amino acids, ΔG_a^{ins} , where a is an index referring to each of the twenty naturally occurring amino acids, has been estimated from measurements of the effects of point mutations on the relative quantities of a protein segment in either transmembrane α helical structure or in a location outside of the membrane⁹³. The effect of a mutation depended on residue position within the segment, and the ΔG_a^{ins} were based on the effects of mutations within the central region of the segment which was located within the membrane interior in the transmembrane population. The secondary, tertiary, and quaternary structures of a sequence result in specific proximities between charged, polar, and hydrophobic residues, and these proximities contribute toward the free energy of the structure and may affect the ΔG_a^{ins} values in other systems such as mHFP.

In the NMR sample preparation, HFP was initially in aqueous solution with random coil structure and then bound to membranes with a consequent transition to a membrane-inserted β sheet structure. Within the fully constrained model, there was consideration of multiple membrane insertion states for each antiparallel registry, t , where each state was differentiated by the number of amino acids, n , considered to be membrane-inserted and therefore used to calculate the total membrane insertion energy of a HFP strand. Since β sheets typically extend over at least 4 sequential residues¹⁰⁸, the minimum value allowed for n was 4 and the index i was used to denote the residue number of the first membrane inserted amino acid. The maximum

value allowed for n corresponded to the number of residues within a registry, t . The membrane insertion energy of each state, $\Delta G_{t,n,i}$, was the sum of the ΔG_a^{ins} for all membrane inserted amino acids and was calculated by:

$$\Delta G_{t,n,i} = \sum_{h=i}^{i+n-1} \Delta G_h^{ins} \quad (52)$$

where h is an index indentifying the residue number of a membrane-inserted residue (i.e. $\Delta G_h^{ins} = \Delta G_a^{ins}$ of amino acid type “ a ” at residue “ h ” in the HFP sequence). For each registry, the minimum $\Delta G_{t,n,i}$ had residue “ i ” in a strand hydrogen bonded to residue $i+n-1$ of an adjacent strand, **Figure 28**. In this model, the first membrane inserted residue, residue i , was the same residue number for adjacent strands. The $\Delta G_{t,n,i}$ depend on n and i for a specific t and the minimum $\Delta G_{t,n,i}$ for each registry was denoted ΔG_t^{\min} and corresponded to a specific i_t and n_t . The ΔG_t^{\min} for $t = 8-24$ are displayed in **Figure 29** and **Appendix XIV**. Registries with more negative ΔG_t^{\min} generally had higher best-fit f_t ’s as determined from the NMR analysis. However, ΔG_t^{\min} is a model that uses approximations and it does not quantitatively predict f_t ’s accurately (**Appendix XII**). There are other free energy contributions associated with a peptide/protein’s membrane insertion energy where membrane binding, lipid composition, and peptide/protein secondary and tertiary structures may have non-negligible energetic contributions that contribute to the free energy of a system (for review of this topic, see reference¹⁰⁷).

u	Lables	t	n	i	
20	F8CG13N	20	8	7	<div> <div>AVGIGALFLGFLGAAGSTMGARS</div> <div>SRAGMTSGAAGLFLGFLLAGIGVA</div> </div>
18	L12CL7N	18	6	7	<div> <div>AVGIGALFLGFLGAAGSTMGARS</div> <div>SRAGMTSGAAGLFLGFLLAGIGVA</div> </div>
18	L12CL7N	18	14	3	<div> <div>AVGIGALFLGFLGAAGSTMGARS</div> <div>SRAGMTSGAAGLFLGFLLAGIGVA</div> </div>
13	L9CG5N	13	11	2	<div> <div>AVGIGALFLGFLGAAGSTMGARS</div> <div>SRAGMTSGAAGLFLGFLLAGIGVA</div> </div>

Figure 28. Sample indices, u , with the corresponding labeling schemes are displayed along with the registries, t , that result in labeled r_{CN} of ~ 4.1 Å and ~ 5.5 Å that respectively correspond to hydrogen-bonded and non-hydrogen bonded $^{13}\text{CO} - \text{H}^{15}\text{N}$. Membrane inserted regions are highlighted in yellow, and the corresponding n and i values are listed.

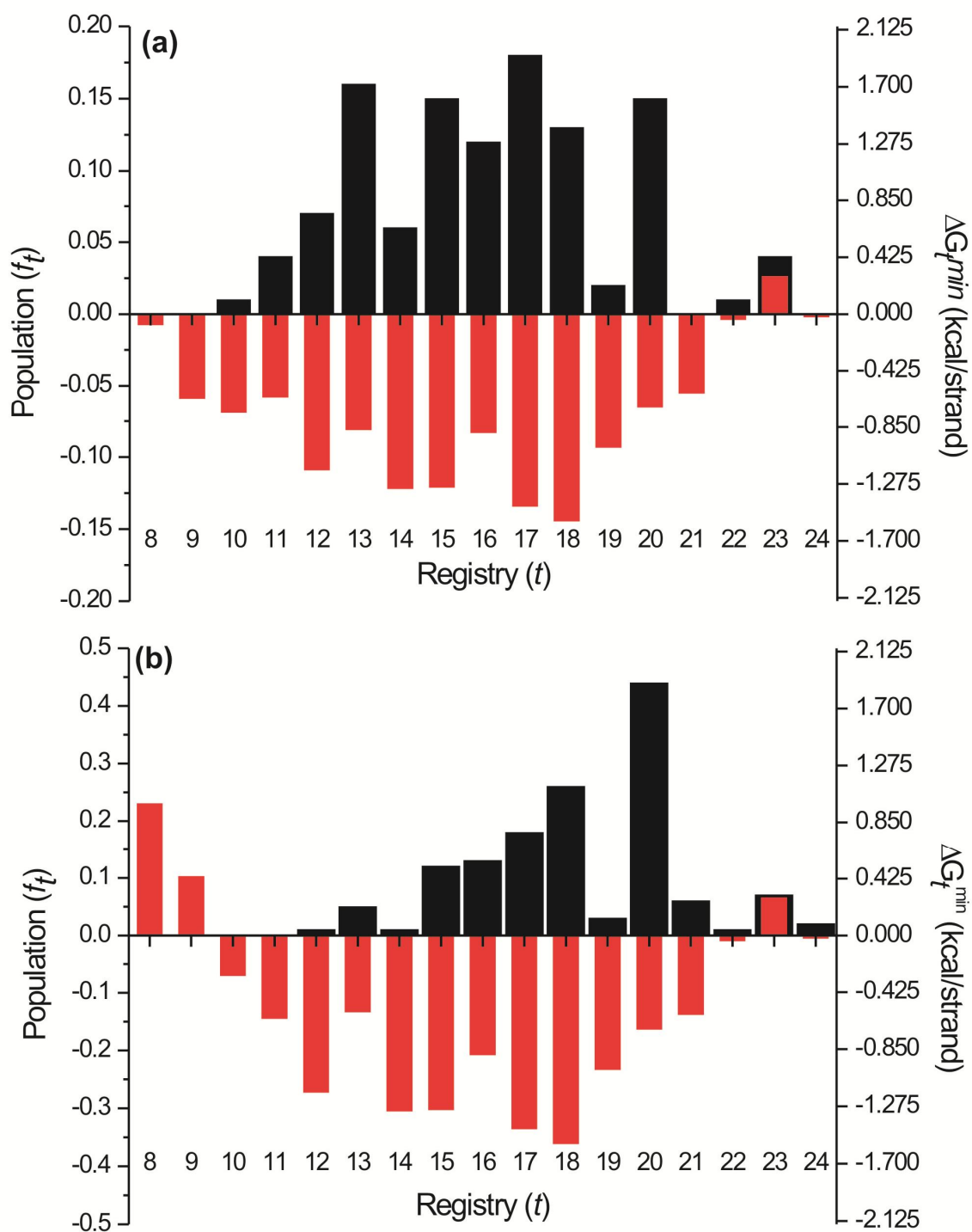


Figure 29. Double-y plot where f_t populations (black) and ΔG_t^{\min} (red) are plotted for each registry for mHFP (a) and mV2E-HFP (b).

2. Fully Constrained Model mV2E-HFP Registry Distribution

A broad distribution of mHFP oligomeric registries were observed, and the importance of a FP oligomeric structure in membrane fusion was suggested by dominant inhibition of fusion and virus infection when a small fraction of V2E mutated gp41 was co-expressed with a major fraction of wild-type gp41²¹. One interpretation of this result is that V2E mutated gp41 can be incorporated into an oligomeric structure of wild-type gp41 and incorporation of V2E mutated gp41 results in the formation of fusion inactive FP oligomers (**Appendix III**). To date, this hypothesized fusion inactive FP oligomeric structure has not been identified. Similarly, V2E-HFP has reduced membrane fusion rates relative to HFP³¹, and structural comparisons between mHFP and mV2E-HFP may help to explain the dominant inhibition of membrane fusion by the gp41 V2E mutant. To investigate potential differences between mHFP and mV2E-HFP, samples were prepared to quantify the registry populations in mV2E-HFP with labeling schemes identical to the mHFP samples. Relative to HFP-WT, larger $(\Delta S/S_0)^{\text{exp}}$ at $\tau = 48.2$ ms were observed for $u > 17$ while smaller $(\Delta S/S_0)^{\text{exp}}$ at $\tau = 48.2$ ms were observed for $u < 16$, **Figure 30a**. These results qualitatively suggest that longer registries ($t > 17$) are more populated in mV2E-HFP while shorter registries, $t < 16$, are more populated in mHFP. These qualitative comparisons are in good agreement with the quantitative comparisons between mHFP and mV2E-HFP populations,

Figure 30b and **Table 7**. For the mV2E-HFP registry population fittings, the $\sum_{t=8}^{t=24} f_t = 1.39$ and

the $\chi^2_{\min} = 87$ for 102 degrees of freedom. The mV2E-HFP χ^2_{\min} was lower than the mHFP χ^2_{\min} , and one contributing factor is the narrower registry distribution for mV2E-HFP since the

χ^2_{min} were generally lower for $\sum_{t=u-1}^{t=u+1} f_t = 0.00$ than for data sets with populated f_t 's, **Table 7**.

Another possible contributing factor may be that our calculations only considered the spectral noise contribution to the error and did not account for other errors associated with our data analysis. The largest unaccounted for source of error is likely the error associated with the f_t based $(\Delta S/S_0)^{sim}$ data that were fit to the $(\Delta S/S_0)^{exp}$ in the χ^2 analyses. As demonstrated by the duplicate analysis of the global χ^2_{min} with two different F8CG13N mHFP sample preparations, **Table 9**, smaller rms error lead to similar best fit f_t , but larger χ^2_{min} (+9). This suggests that our experiments were reproducible, but the experimental error may be under approximated when quantified by rms error. Additionally, out of 119 data points, mHFP had smaller spectral noise as indicated by smaller rms error in 64 points while mV2E-HFP had smaller spectral noise in 55 points. If the overall experimental error is generally greater than the spectral noise error, then we would expect global $\chi^2_{min}(\text{mHFP}) > \text{global } \chi^2_{min}(\text{mV2E-HFP})$ since the spectral noise error is smaller for mHFP.

There are multiple potential contributing factors that made $\sum_{t=8}^{t=24} f_t > 1.00$: (1) The r_{CN}

distances may be shorter in mHFP and mV2E-HFP constructs than the modeled r_{CN} that were derived from the 2IWW.pdb β barrel crystal structure. This would result in larger actual d_{CN} than the modeled d_{CN} . Therefore, larger $f_{t=u-1}$, $f_{t=u}$, and $f_{t=u+1}$ registries would be required to model the actual dephasing. Other experiments have demonstrated that larger side chain

groups can affect the hydrogen bond strength of 3_{10} -helices¹⁰⁹. Since the HFP sequence contains many Gly residues (i.e. the amino acid with the smallest sidechain), the hydrogen bonds in mHFP β sheets may be stronger resulting in shorter r_{CN} than those found in the 2IWW β sheets. However, to my knowledge, there is currently no experimental data for β sheets to support this hypothesis. (2) The dephasing due to the $t = u-2$ and $t = u+2$ registries was not accounted for in the 3 registry population fittings for each sample. These registries were accounted for in the 5 parameter (registry) fittings where $\sum_{t=8}^{24} f_t$ was reduced by 0.09 and 0.13

for the mHFP and mV2E-HFP, respectively, **Table 7** and **Table 8**. The modified equations for these fittings are found in **Appendix VIII**; (3) The β sheets may not be fully constrained β sheets. Other models, such as the unconstrained model (**Appendix IX**), result in smaller

$\sum_{t=8}^{24} f_t$ than the fully constrained model for the best-fit f_t 's.

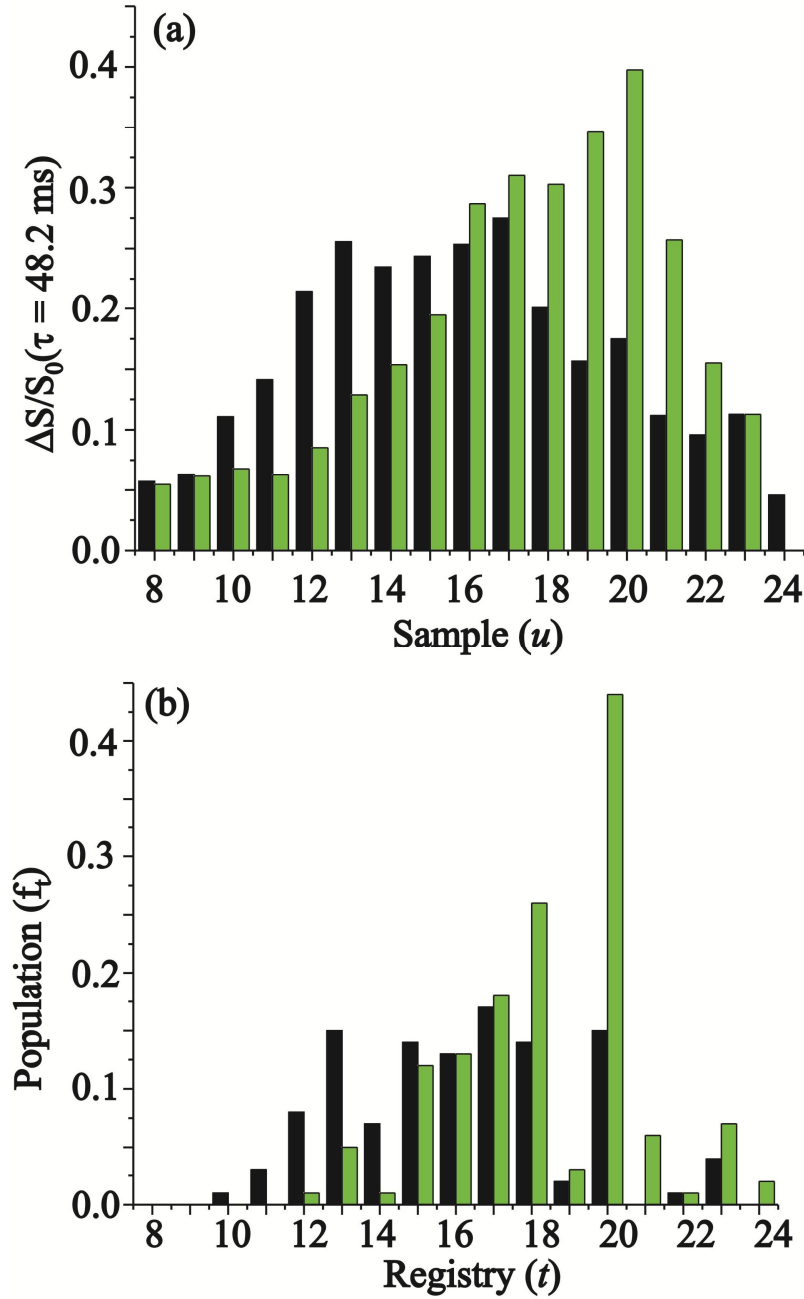


Figure 30. (a) The $(\Delta S/S_0)^{\text{exp}} (\tau = 48.2 \text{ ms})$ for mHFP (black) and mV2E-HFP (green). (b) The f_t for mHFP (black) and mV2E-HFP (green) for the fully constrained model using the 3 registry fitting method. Both (a) and (b) demonstrate that mV2E-HFP has a smaller population of shorter registries ($t < 16$) and that mV2E-HFP has a larger population of longer registries ($t > 17$).

This Chapter has shown that mHFP forms a broad distribution of registries that are correlated to membrane insertion energy of each registry, ΔG_t^{\min} , in mHFP. Previous work has shown that a HFP construct with Phe-11 mutated to Gly has significantly reduced membrane fusion activity while the Phe-11 to Val mutation has reduced membrane fusion activity compared to the wild-type HFP construct, but higher activity than the F11G mutation¹¹⁰. These results are consistent with the idea that forming structures with more negative ΔG_t^{\min} are important for forming fusion active structures since the F11G mutation increases the calculated ΔG_t^{\min} for most registries with $t > 10$ (**Appendix X**). However, mV2E-HFP has a surface membrane location²⁶, and it is therefore counterintuitive to expect the negative magnitude of ΔG_t^{\min} to be correlated to the magnitude of the populated registries in mV2E-HFP. The calculated ΔG_t^{\min} for mV2E-HFP are similar to mHFP, and f_t 's are very different for mV2E-HFP and mHFP, **Figure 30** and **Table 7**. Therefore, other energy contributions associated with V2E mutation likely affect the distribution of registries and fusion activity (see Discussion). However, it is of note that populated registries in mV2E-HFP generally had negative ΔG_t^{\min} values.

Since mV2E-HFP has a shallower membrane insertion depth than mHFP and a larger population of $t > 17$ registries, these data are consistent with all or a subset of “longer registries”, $t > 17$, having shallower membrane insertion depth and reduced membrane fusion activity relative to all or a subset of the “shorter registries” $t < 16$. The $t \sim 17$ was chosen as a distinguishing registry in this qualitative model because f_{17} for HFP is approximately equal to the f_{17} for V2E-HFP, **Figure 30**. Additionally, for $t > 13$ in V2E-HFP registries, the charged

Glu-2 residue is placed further away from the most hydrophobic region (LFLGFL) of the adjacent strands as t becomes larger. This likely has a favorable contribution to the free energy of these longer registries over the $t \leq 13$ registries that contribute to the minimal population of $t \leq 13$ registries in V2E-HFP. HFP membrane insertion is likely related to favorable hydrophobic interactions with the membrane relative to water, and the hydrophobic residues in HFP are primarily found in the first 12 N-terminal residues. As previously noted, shorter registries, such as $t = 13$ registry, cluster hydrophobic residues in HFP oligomers which may be important for achieving deeper membrane insertion and more rapid fusion. The $t = 13$ registry has $i_t = 2$ and $n_t = 11$ at the ΔG_t^{\min} where i_t and n_t are respectively the number of the first membrane-inserted residue and the total number of inserted residues in registry t that correspond to ΔG_t^{\min} . The $t = 20$ registry has $i_t = 7$ and $n_t = 8$. Registries with more residues incorporated into the membrane inserted region may allow for the hydrophobic segments to protrude deeper into the membrane. Additionally, residues that flank the membrane inserted region could affect the membrane insertion depth as well. For mHFP, the $t = 13$ registry has a membrane inserted region that spans 11 residues (-0.9 kcal/strand) with one flanking residue on each side of the membrane inserted region (+0.9 kcal/strand). For the $t = 20$ registry, the membrane inserted region spans 8 residues (-0.7 kcal/strand) while the residues flanking this region span six residues on each side (+3.6 kcal/strand). While flanking residues with disordered secondary structure had minimal contributions to the actual ΔG_t^{\min} of helical structure in the Hessa study (**Appendix XV**)⁹³, the large flanking regions of the $t = 20$ registry may contribute toward the actual ΔG_t^{\min} since the backbone dihedral angles within sheets are more restricted than disordered flanking regions of

the α helices in the Hessa experiments. The extension of the β sheet from the membrane inserted region to the flanking region for the $t = 20$ registry is supported by ^{13}C O chemical shift data for both mHFP and mV2E-HFP where the Ala-1 and Ile-4 residues have chemical shifts consistent with β sheet structure²⁶. In general, this suggests that the β sheet structure extends from the membrane inserted region to the flanking regions in both mHFP and mV2E-HFP. This is different from the Hessa experiments where Pro and Gly residues were inserted adjacent to the membrane inserted region to disrupt secondary structure (**Appendix XV**). The flanking regions' strong preference for the bilayer-water interface over the hydrophobic membrane interior may anchor the $t = 20$ registry to the bilayer-water interface and result in an overall shallower membrane insertion depth than the $t = 13$ registry. The $u = 20$ sample, F8CG13N labeling, was also prepared with the highly fusogenic HFPtr. Relative to mHFP, smaller $(\Delta S/S_0)^{\text{exp}}$ were observed for mHFPtr which was opposite to the mV2E-HFP result, **Figure 30**. These data suggest relative to mHFP and mV2E-HFP, longer registries are less populated in mHFPtr and provide corollary evidence for the hypothesis that the $t \sim 20$ registries are less fusogenic than shorter registries. These combined data suggest that the FP may provide a target for fusion inhibitory drug design by stabilization of longer ($u \geq 20$) FP registries.

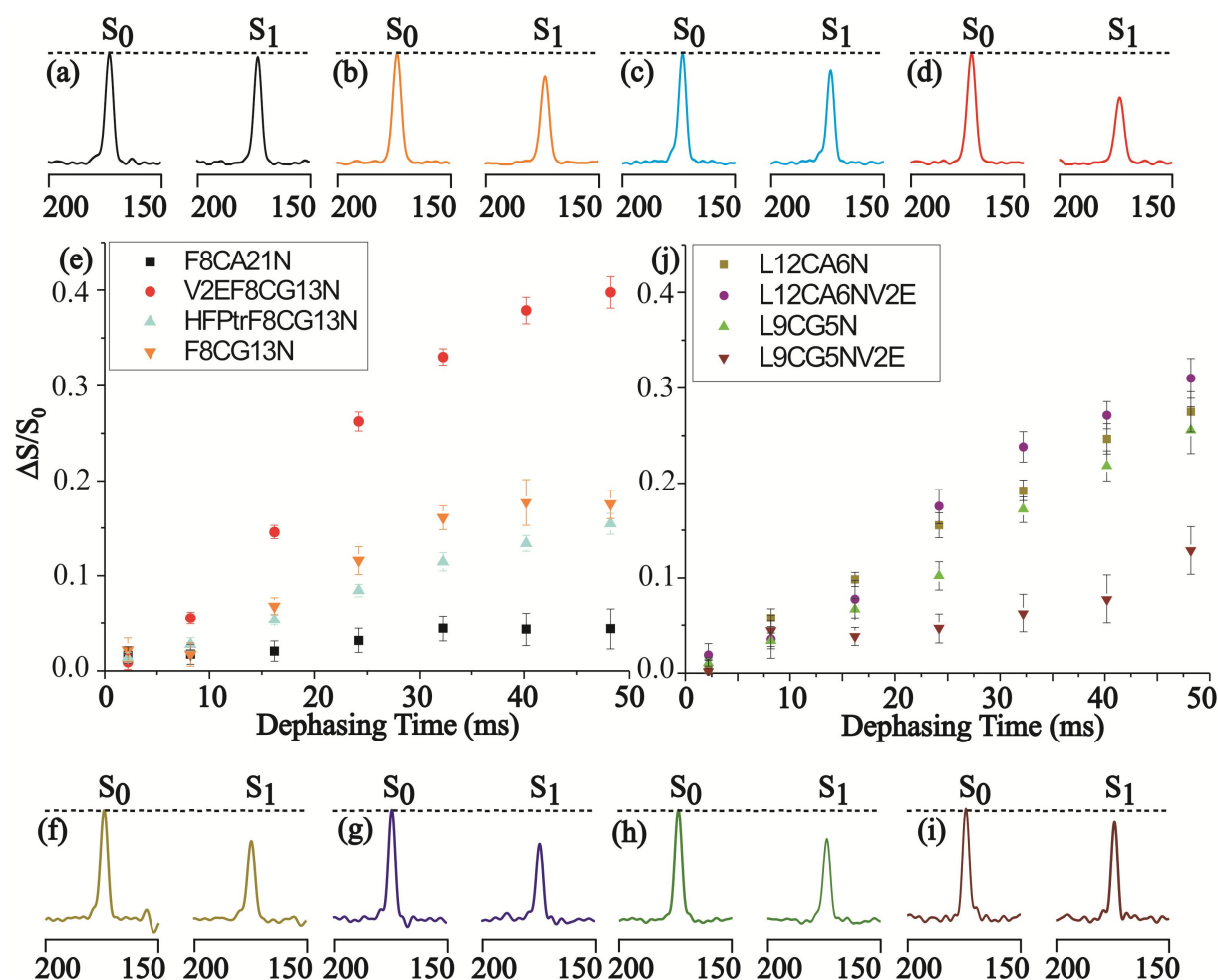


Figure 31. REDOR S_0 and S_1 ^{13}C SSNMR spectra at 48.2 ms dephasing time for (a) HFP-F8CA21N, (b) HFP-F8CG13N, (c) HFPtr-F8CG13N, (d) V2E-F8CG13N, (f) HFP-L12CA6N, (g) V2E-L12CA6N, (h) HFP-L9CG5N, or (i) V2E-L9CG5N. Each spectrum was processed with 200 Hz line broadening and baseline correction and was the sum of: (a) 46816; (b) 36665; (c) 19372; (d) 34271; (f) 44931; (g) 46231; (h) 40272; or (i) 46809 scans. (e), (j) Plots of $(\Delta S/S_0)^{exp}$ vs dephasing time with the rms error. Isotopic labeling of each mHFP is displayed in the legend that correspond to HFP-F8CA21N, (black, square), HFP-F8CG13N (orange, inverted triangle), HFPtr-F8CG13N (cyan, triangle), (d) V2E-F8CG13N (red, circle), (f) HFP-L12CA6N (dark yellow, square), (g) V2E-L12CA6N (purple, circle), (h) HFP-L9CG5N (green, triangle), or (i) V2E-L9CG5N (wine, inverted triangle). Variation less than ± 0.02 in $(\Delta S/S_0)^{exp}$ was also observed between two different preparations of the same sample type, e.g. HFP-F8CG13N.

4.4 Discussion

1. Modeled Membrane Insertion Depth mHFP

Earlier SSNMR data showed that the deepest membrane insertion depth for mHFP occurred for the Ala-6 and Leu-9 residues²⁶. The results of these prior experiments were compared to newly calculated membrane insertion depths of each residue's ^{13}CO for Ala-1 through Gly-16. In general, these equations have dependences upon the indices: (1) i , the residue number for the first membrane inserted residue (N-terminal most residue of the membrane inserted region); (2) n , the number of membrane inserted residues; and (3) h , the residue number of the ^{13}CO for which the membrane insertion depth was calculated. The inserted region of a β sheet was modeled as a semi-circle, and the membrane insertion depth of each residue's ^{13}CO , $D_{h,i,n}$, was calculated, **Eqs (53)-(58)**. Additionally, **Figure 32** provides a visual aid to follow the $D_{h,i,n}$ calculations.

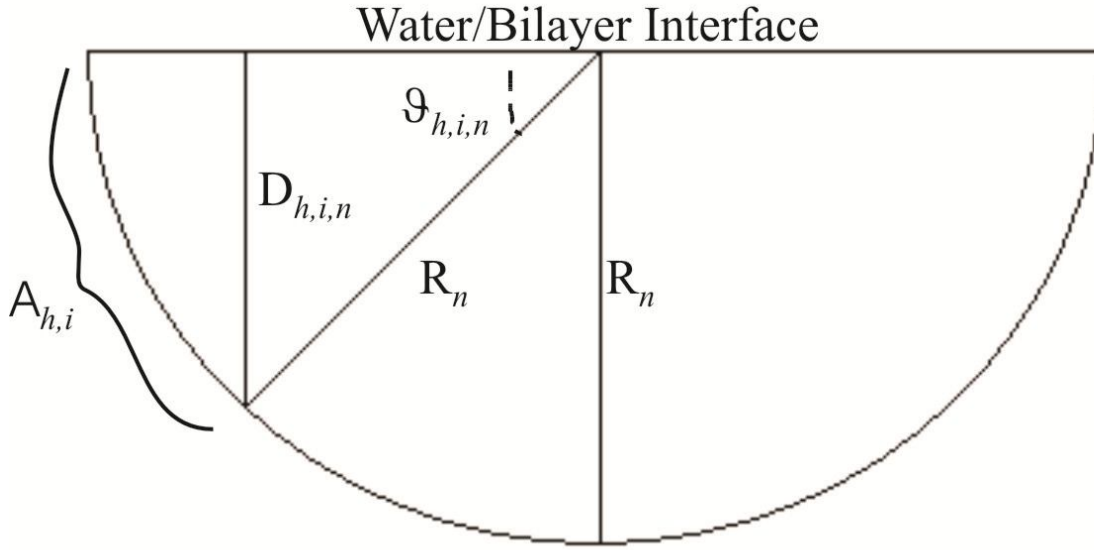


Figure 32. Membrane insertion depth model as described in the text. Previous work has demonstrated that the Ala-1 carbonyl carbon is ~ 5 Å from the lipid phosphorus, and this phosphate region is referred to as the water/bilayer interface.

$$D_{h,i,n} = R_n \sin(\phi_{h,i,n}) \quad (53)$$

where R_n is the radius of the semi-circle for registry t and $\phi_{h,i,n}$ is the angle between the plane of the water/bilayer interface and R_n when R_n is extended to the carbonyl of residue h .

$$R_n = \frac{3.5n}{\pi} \quad (54)$$

where each amino acid in β sheet conformation spans ~ 3.5 Å. The angle $\phi_{h,i,n}$ was calculated by:

$$\phi_{h,i,n} = \frac{A_{h,i}\pi}{3.5n} \quad (55)$$

where $A_{h,i}$ was the arc length along the semi-circle from residue i to h where residue i is N-terminal to the ^{13}CO of residue h . The arc length, $A_{h,i}$ was calculated by **Eq (56)**

$$A_{h,i} = 3.5(h - i + 1) \quad (56)$$

where the N-terminus of the semi-circle is the ^{13}CO of residue $i - 1$, and the C-terminus of the semi-circle is the ^{13}CO of residue $i + n - 1$. For $h < i$, $h > i + n - 1$, or $n_t = 0$ (i.e. the respective conditions where h was N-terminal of the membrane inserted region, h was C-terminal of the membrane inserted region or there was no membrane inserted region) the ^{13}CO of residue h is not located within the membrane inserted region and $D_{h,i,n} = R_n = \varphi_{h,i,n} = A_{h,i} = 0$.

Each t registry could be an ensemble of states which differ in their i and n . In any ensemble, the lowest energy state will be the most populated. In calculating the membrane insertion depth of residue h in registry t , it was approximated that the only populated state was the one with $i = i_t$ and $n = n_t$ corresponding to ΔG_t^{\min} . Under these conditions, **Eq (53)** can be rewritten as:

$$D_{h,i_t,n_t} = \frac{3.5n_t}{\pi} \sin\left(\frac{(h - i_t + 1)\pi}{n_t}\right) \quad (57)$$

Thus, the average membrane insertion depth of each residue's ^{13}CO in mHFP for the fully constrained model, **Figure 33**, was calculated by:

$$D_{h,i_t,n_t}^{avg} = \frac{\sum_{t=8}^{24} f_t D_{h,i_t,n_t}}{\sum_{t=8}^{24} f_t} \quad (58)$$

where 1.14 and 1.39 correspond to the $\sum_{t=8}^{24} f_t$ for mHFP and mV2E-HFP in the fully constrained model, respectively. This calculation is consistent with the membrane insertion depth data in that Ala-6 and Leu-9 were inserted deeper than Ile-4, Leu-12, and Ala-14. For most $t < 16$ registries,

the calculated Ala-6 $D_{h,i_t,n_t} > 5 \text{ \AA}$ while $D_{h,i_t,n_t} \sim 0 \text{ \AA}$ for $t \geq 16$. Relative to mHFP, mV2E-HFP have a larger population of $t > 16$ registries with consequent smaller D_{h,i_t,n_t}^{avg} for Ala-6 in mV2E-HFP, **Figure 34**. This is consistent with earlier SSNMR measurements of deeper membrane insertion of the Ala-6 of mHFP relative to the Ala-6 of mV2E-HFP.

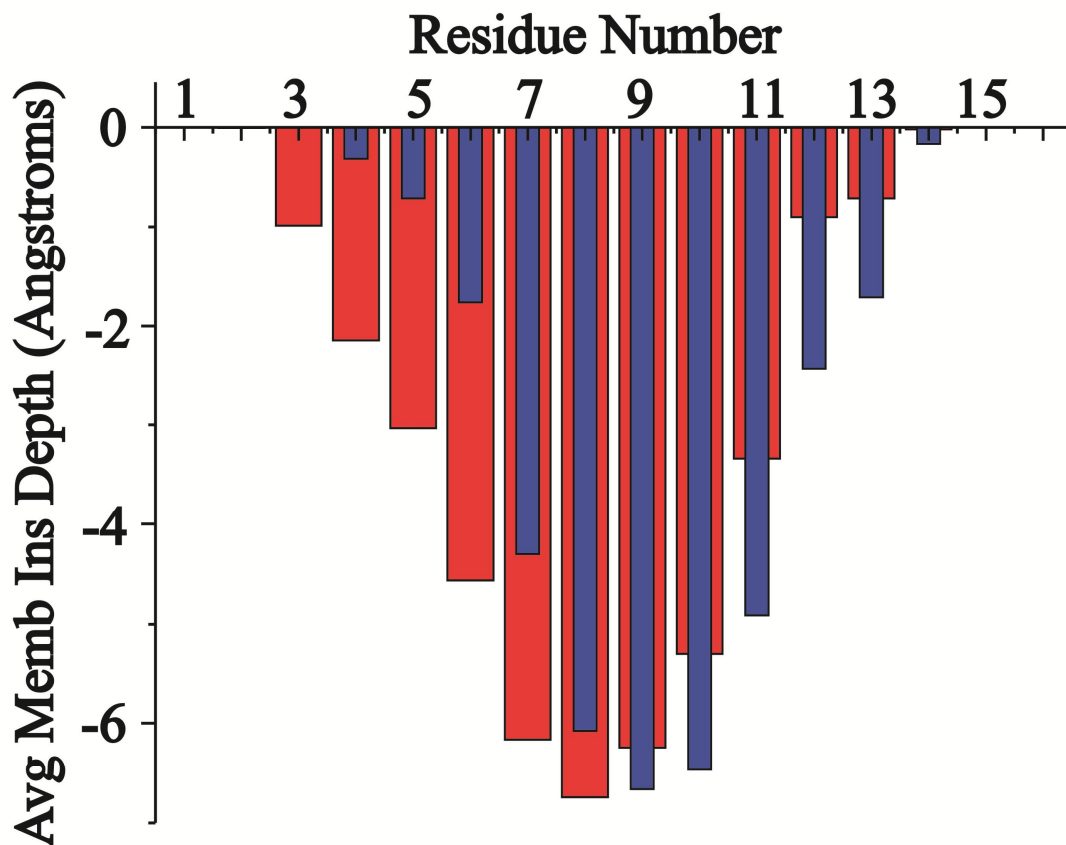


Figure 33. The calculated average membrane insertion depth of each residue's ^{13}CO , D_{h,i_t,n_t}^{avg} , is plotted for HFP (red) and V2E-HFP (blue).

Table 10. mHFP and mV2E-HFP $^{13}\text{CO } D_{h,i_t,n_t}^{avg}$

	$^{13}\text{CO } D_{h,i_t,n_t}^{avg}$	
Residue Number	mHFP	mV2E-HFP
1	0.00	0.0
2	-0.0091	0.0
3	-0.99	0.0
4	-2.1	-0.32
5	-3.0	-0.72
6	-4.6	-1.8
7	-6.2	-4.3
8	-6.7	-6.1
9	-6.2	-6.7
10	-5.3	-6.5
11	-3.3	-4.9
12	-0.90	-2.4
13	-0.72	-1.7
14	-0.030	-0.17
15	-4.3 E-17	-1.1 E-16
16	0	0.0

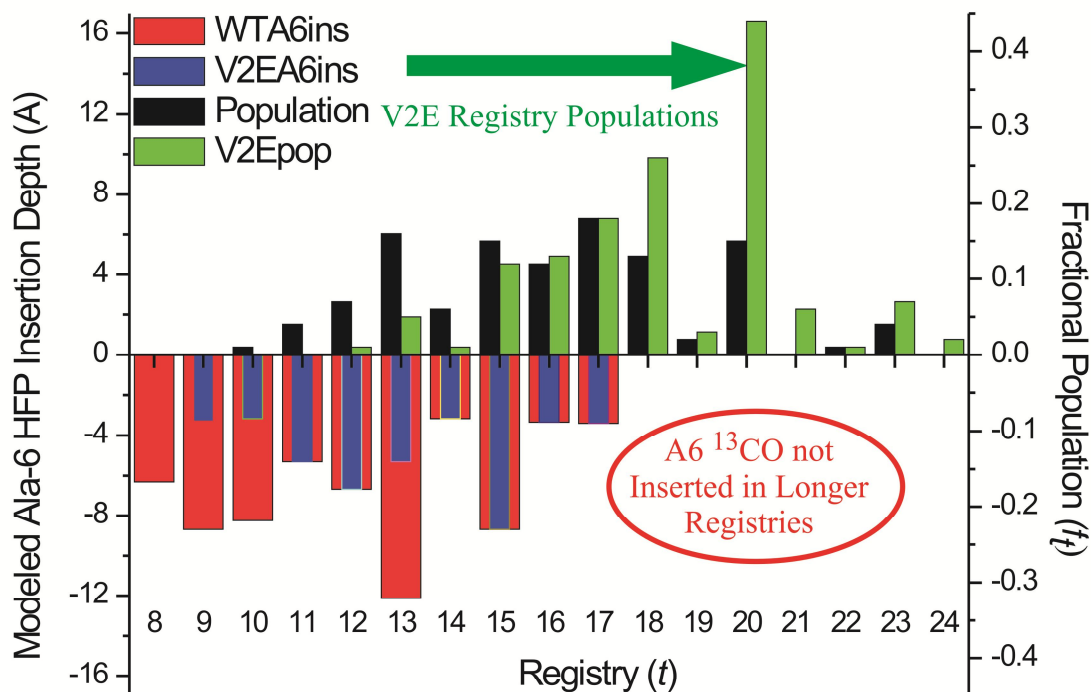


Figure 34. Calculated membrane insertion depth of the Ala-6 ^{13}CO for HFP (red) and V2E-HFP (blue) is plotted for each registry. Additionally, f_t for mHFP (black) and mV2E-HFP (green) are plotted for each registry. The mV2E-HFP registry populations are shifted toward longer registries ($t > 17$) relative to mHFP and the calculated membrane insertion depth of the Ala-6 ^{13}CO is ≥ 0 for these registries. This is consistent with previous work where the Ala-6 ^{13}CO is inserted deeper in mHFP relative to mV2E-HFP²⁶.

2. Relevance of Broad Distribution

The quantitative determination of the f_t 's is the most significant result of Chapter IV.

There is qualitative correlation between the f_t 's and the respective negative magnitudes of the

ΔG_t^{\min} 's in mHFP, but there is not quantitative agreement between the f_t 's and those from

thermodynamic calculations, f_t^{calc} , using the ΔG_t^{\min} 's (**Appendix XII**). For α helices, one error

in the estimation of ΔG_t^{\min} was neglect of the (negative) contribution from local clustering of

hydrophobic residues (Support by **Appendix XIII, Figure 52**). For mHFP, such clustering is most prominent in the first 12 residues in the sequence and may contribute to relatively large f_{13} since the $t = 13$ overlaps residues 2-12 (i.e. all of the hydrophobic residues with inclusion of the least possible number of non-hydrophobic residues). In addition, each registry's energy will include a (negative) enthalpic contribution if void space is minimized²⁹. The contribution from such tight packing in β sheets has been evidenced by frequent observation of hydrogen bonding between Phe and Gly residues which have large and small sidechains, respectively¹¹¹. For mHFP, the Phe''Gly hydrogen bonding in the $t = 20$ and 23 registries may contribute to the results $f_{20} > f_{19}, f_{21}$ and $f_{23} > f_{22}, f_{24}$. In mV2E-HFP, there is poor qualitative correlation between the negative magnitude of ΔG_t^{\min} and the best-fit f_t from the NMR analysis, **Figure 29**. For example, the $\Delta G_{20}^{\min} \sim (1/2)\Delta G_{18}^{\min}$, but $f_{20} > f_{18}$ by 0.18. Electrostatic interactions and sidechain packing appear to contribute to the free energy of a particular registry in mV2E-HFP. The $t = 20$ registry may be favored in mV2E-HFP due to the favorable Phe-Gly sidechain packing. Additionally, published crystal structures with charged residues have shown that charged residues are not typically found in β sheet oligomers adjacent to hydrophobic residues¹⁰⁵. This suggests that it is not energetically favorable to have glutamic acid adjacent to hydrophobic residues which may be due to unfavorable electrostatic interactions between charged and hydrophobic sidechains. This is statistically represented where glutamic acid is more frequently found in β sheets when adjacent to charged residues relative to hydrophobic residues¹⁰⁵. In mV2E-HFP, longer β sheet registries place Glu-2 near more polar residues, but

there is currently no residue specific structural data for the Glu-2 residue in mV2E-HFP. However, the most favorable electrostatic interactions may result from proximity to the Ser-17, Thr-18 or Arg-22 residues (i.e. $t \sim 17-23$). This may present favorable interstrand electrostatic interactions between sidechains, but having multiple charged residues near the edges of the β sheets could provide more favorable electrostatic interactions for membrane binding and an overall lower free energy than for shorter registries (i.e. $t \sim <17$). Additionally, mV2E-HFP has shallower membrane insertion relative to mHFP²⁶. Relative to mHFP, there may therefore be less correlation between membrane insertion energies and f_t 's calculated for mV2E-HFP.

Plausible hypotheses to explain the low populations of $t < 17$ registries and higher populations of $t > 17$ registries in mV2E-HFP relative to mHFP include: (1) Longer β sheet registries place Glu-2 near more polar residues, and the most favorable electrostatic interactions may result from Glu-2 adjacent to the presumably positively charged Ser-17, Thr-18 or Arg-22 residues (at pH ~ 7). (2) Our experiments sample the membrane bound structures, and shallow membrane insertion may be required for mHFP and mV2E-HFP to stay bound to the membrane (see **Appendix XIII**, L9R data, for counter example). The $\Delta G_{\text{Glu-2}} = +2.68$ kcal/mol while the ΔG_t^{min} of populated registries were typically > -1.7 kcal/strand and $< \sim -0.4$ kcal/strand, **Figure 29**. The $\Delta G_{\text{Glu-2}} = +2.68$ kcal/mol reflects Glu's high preference to stay in a hydrophilic environment over the hydrophobic environment of the interior of a membrane. For mV2E-HFP registries where Glu-2 is located near "membrane inserted" region, it is plausible that Glu's energetic gain for being located in a hydrophilic environment may outweigh the energetic gain for having the hydrophobic region membrane inserted. As the Glu-2 residue is placed further away from the "membrane inserted" region, it seems possible that the Glu-2 could be placed in a hydrophilic

aqueous environment while the “membrane inserted” region could be in a more hydrophobic environment such as the membrane interior or a β sheet aggregate where the hydrophobic residues of adjacent strands are proximate to each other. In the highly populated registries of mHFP-V2E, $t > 17$, Glu-2 is located far away from the “membrane inserted” region which supports the above hypothesis. This theory is also supported by previous ^{13}CO - ^{31}P mV2E-HFP REDOR experiments where residues that were ^{13}CO labeled at the Ala-6, Leu-9, and Leu-12 (more hydrophobic region of sequence) dephase less than Ala-1, Ile-4, and Ala-14 (more hydrophilic regions of the sequence)¹⁰⁶. Additional experiments to further validate this hypothesis have not been designed to date.

Relative to the mHFP $u = 20$ sample, much larger $(\Delta S/S_0)^{exp}$ were observed for the $u = 20$ mV2E-HFP sample and the opposite trend was detected for the $u = 13$ samples. These data demonstrate significant differences in the registry distributions of the two peptides which may correlate to the large differences in their rates of vesicle fusion. Previous experiments have shown that faster vesicle fusion rates are correlated with deeper membrane insertion²⁶, and my data and calculations suggest that registries with higher populations ($f_t > 0.05$) have negative free energies for membrane insertion. In general, previous work has supported that membrane insertion of peptides/proteins are governed by thermodynamics^{93,107,112,113}. The culmination of the collected data and previous work are consistent with the mHFP and mV2E-HFP registries being governed by thermodynamics. Because thermodynamics appear to govern the membrane inserted structures for proteins/peptides, the following hypotheses are reasonable: (1) The gp41 FP that is membrane bound, has deep membrane insertion depth, and is fusion active may have a

distribution of registries similar to mHFP or weighted toward the “shorter” registries (possibly $t < 17$); and (2) The gp41 FP that is membrane bound, has shallow membrane insertion depth, and is fusion inactive may have a distribution of registries that is similar to mV2E-HFP and weighted toward “longer” registries (possibly $t > 17$). Alternatively, the FP registry distributions in gp41 may result from interactions with residues C-terminal of FP, but this hypothesis is not supported by the mHFP and mV2E-HFP data. However, this hypothesis has not been proven wrong to date and therefore cannot be entirely ruled out. Previous work has shown that the fusion peptide synthesized with the N-terminal helical region of gp41 (i.e. a longer construct commonly referred to as N70) has a higher vesicle fusion rate relative to HFP, but it has not been shown whether: (1) The faster lipid mixing rate of the N70 construct results from a structural change with the fusion peptide region; or (2) The faster lipid mixing rate of the N70 construct results from an interaction between the N-terminal helix (possibly dehydration of the membrane surface that results from the N-terminal helix binding to the membrane surface).

Detection of registry distributions may also be significant for other peptides and proteins. For example, recent data support higher neurotoxicity of small oligomers of amyloid peptides and lower toxicity of large and mature fibrils with well-defined β sheet structure and a single registry (usually in-register parallel)^{102-105,114}. There are little data about the registry distribution of the oligomers and the present approach could be applied to determine this distribution. Comparison with the distribution in the fibril will improve understanding of the amyloid structure-function relationship and aid development of new inhibitors of amyloid oligomer formation and amyloid disease.

Chapter V. Dissertation Summary and Future Work

5.1 Summary

The preliminary focus of this dissertation was to contribute toward the literature of conflicting reports regarding the population of in-register parallel β sheets for membrane-associated HFP constructs^{29,47}. To achieve this, NMR samples were prepared with more sparsely isotopically labeled peptides which provided more specific registry detection and less interpretational ambiguity than for previous studies^{29,47}. Additionally, a more quantitative model to account for natural abundance dephasing contributions was developed to improve upon previous models^{36,54} which improved the accuracy of quantitative data analysis for REDOR experiments. Very little in-register parallel β sheets were detected (< 15% of the β sheet structure)⁶³, and this result argues against previous hypotheses that in-registry parallel β sheets are required for fusion activity. Later experiments identified the presence of a broad distribution of antiparallel registries in mHFP. The distribution of registries appears to be functionally relevant since the registry distribution of mV2E-HFP is narrower and shifted toward longer registries that delocalize hydrophobic residues. The more fusion active mHFP has a distribution of registries that includes shorter registries that cluster hydrophobic residues. This clustering of hydrophobic residues in mHFP may explain mHFP's overall deeper membrane insertion relative to mV2E-HFP²⁶.

For mHFP, there was qualitative correlation between the magnitude of each registry's f_i and the negative magnitude of each registry's ΔG_i^{\min} which suggests that calculating ΔG_i^{\min} energies for registries may be useful in predicting β sheet registries for any sequence of

membrane-bound β sheet structure. Alternatively, the distribution of registries within the less functional mV2E-HFP displayed poor qualitative correlation between the magnitude of each registry's f_t and the negative magnitude of each registry's ΔG_t^{\min} . While this result may initially be discouraging, it is of note that mV2E-HFP has shallow membrane insertion depth, and therefore, the energy minimized structure should not result from an energetic contribution resulting from the change in free energy in going from aqueous solution to a membrane inserted β sheet. It is also important to keep in mind that the NMR samples contain registries that stay bound to the vesicle aggregates and that the registries observed in mV2E-HFP typically had negative ΔG_t^{\min} . Additionally, the L9R mutant gp41 does not transdominantly inhibit wild type gp41²¹ and L9R-HFP registries had predominantly positive ΔG_t^{\min} values, **Figure 35**, which may inhibit formation of a membrane bound oligomeric structure with the wild type fusion peptide. The β sheet structure was not predominantly formed in mL9R-HFP as observed with mHFP and mV2E-HFP, **Appendix XIII**. Typical values for ΔG_t^{\min} in populated registries in mHFP ranged from -0.5 to -1.6 kcal/mol. Other literature has shown that sidechain-sidechain hydrogen bonds located at the water-bilayer interfacial region have been shown to contribute \sim -0.4-0.8 kcal/mol¹¹⁵. This suggests that hydrogen bonds between polar or charged amino acids contribute to the free energy of a structure in a similar magnitude as that of the ΔG_t^{\min} . These energy contributions that result from charged sidechains may explain why there is poor qualitative agreement between the f_t 's in mV2E-HFP and the negative magnitude of the ΔG_t^{\min} 's. Additionally, sidechain packing interactions may also contribute to the energy

minimized structure as previous suggested in the glycine mutant studies²⁸ and as supported by the higher than expected f_{20} and f_{23} in mHFP (Chapter IV). Thus, predicting β sheet registries based upon ΔG_t^{\min} is likely most accurate for hydrophobic sequences lacking or separated from charged residues (i.e. similar to HFP), but other energy contributions should also be considered. Thus, predicting registry populations from each registry's ΔG_t^{\min} may prove useful for sequences such as the measles fusion peptide and simian immunodeficiency virus (SIV) fusion peptide whereas sequences containing charged residues within or near the hydrophobic sequence (i.e. similar to V2E-HFP and L9R-HFP), will likely have significant energetic contributions from the charged sidechains and poor qualitative agreement between the negative magnitude of ΔG_t^{\min} and f_t 's. Alternatively, it seems that charged residues may be incorporated into a sequence with minimally effecting the membrane insertion energies when they are separated by a break in the secondary structure as observed in the Hessa experiments (**Appendix XV**)⁹³ and as observed with the K₆ tag in mHFP^{36,106}. Future experiments in other protein-membrane samples should consider quantifying distributions of registries/structures rather than only attempting to qualitatively identify a single registry/structure since the mHFP and mV2E-HFP data have demonstrated that broad distributions of structures can have approximately equal free energies in a membrane environment. In doing this, functional and nonfunctional structures in complex systems may be identified.

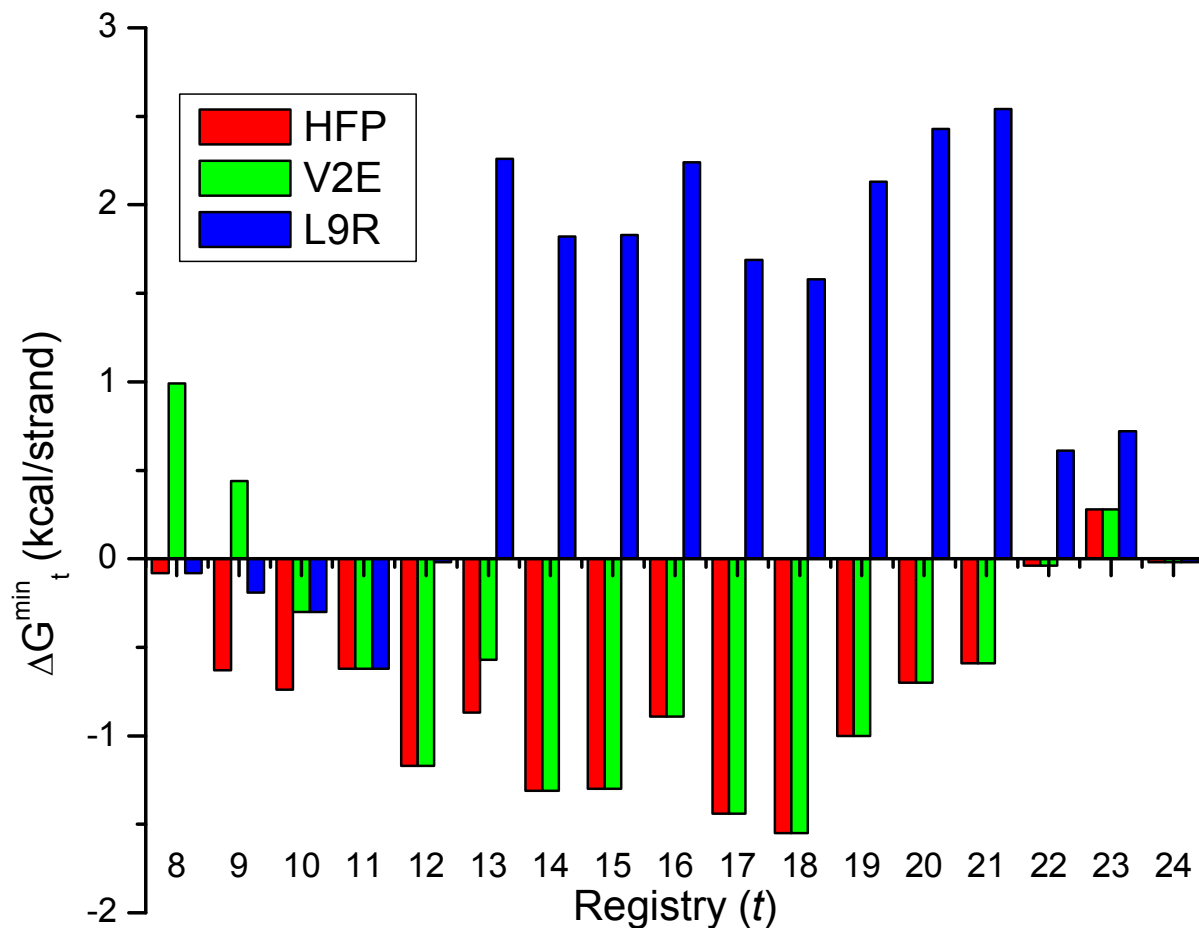


Figure 35. The membrane insertion energies were derived from the Hessa biological hydrophobicity scale for the HFP, V2E-HFP and L9R-HFP by the methods described in Chapter IV. The L9R-HFP has predominantly positive ΔG_t^{\min} whereas both HFP and V2E-HFP have many registries with negative ΔG_t^{\min} which suggests that the distribution of registries should be different between constructs. Additionally, it is not obvious that mL9R-HFP should form membrane inserted β sheets since $t < 12$ registries were minimally populated in mHFP, and $t > 12$ registries have positive ΔG_t^{\min} in mL9R-HFP. Unlike mHFP and mV2E-HFP, mL9R does not predominantly form membrane inserted sheets, **Appendix XIII**.

Table 11. ΔG_t^{\min} values for HFP constructs.

Registry (<i>t</i>)	ΔG_t^{\min} (kcal/mol)		
	HFP	V2E-HFP	L9R-HFP
8	-0.08	0.99	-0.08
9	-0.63	0.44	-0.19
10	0.545	-0.3	-0.3
11	0.065	-0.62	-0.62
12	-0.76	-1.17	-0.02
13	-0.87	-0.57	2.26
14	-0.56	-1.31	1.82
15	-1.3	-1.3	1.83
16	-0.7	-0.89	2.24
17	-1.44	-1.44	1.69
18	-1.55	-1.55	1.58
19	-0.81	-1	2.13
20	-0.7	-0.7	2.43
21	-0.59	-0.59	2.54
22	0.15	-0.04	0.61
23	0.99	0.28	0.72
24	1.51	-0.02	-0.02

* ΔG_t^{\min} values were calculated from the methods described in Chapter IV.

5.2 Membrane Location

Previous membrane location studies from this group have made lipid ^{31}P – HFP ^{13}CO distance measurements as well as lipid ^{19}F – HFP ^{13}CO distance measurements to determine the membrane location of mHFP, mV2E-HFP and mHFPtr^{26,106}. The ^{31}P - ^{13}CO distance measurements provide a great quantitative tool for measuring ^{13}CO distance from the water-

bilayer interface of a vesicle where ~86% of the Ala-1 ^{13}C O had a ^{31}P - ^{13}C distance of ~5 Å. Additionally, ^{19}F - ^{13}C O distance measurements were performed to probe membrane insertion depth. The ^{19}F - ^{13}C O setup experiments used a helical peptide with the sequence EQLLKALEFLLKELLEKL where the Phe-9 was substituted with p-fluorophenylalanine from Sigma-Aldrich and the Leu-10 was ^{13}C O labeled. This setup compound was effective in setting the π pulses, but the fluorinated lipids are not naturally occurring and required H to F substitution at a single site on the lipid acyl chain. Incorporation of fluorinated lipids into lipid vesicles present complications where (1) fluorinated chains tend to cluster together within vesicles, (2) fluorinated chains can form fibrous bands instead of vesicles, (3) fluorinated chains can affect membrane permeability and surface tension^{116,117}, and (4) fluorinated chains can disrupt the bilayer phase¹¹⁸. These effects were considered in the ^{19}F - ^{13}C O REDOR experiments and the mol fraction of ^{19}F -DPPC lipids used was varied. A 0.09 mol fraction of ^{19}F -DPPC lipids yielded the largest $(\Delta S/S_0)^{exp}$, and this mol fraction was considered to maximize potential peptide ^{13}C O-lipid ^{19}F contacts with minimal disruption to the bilayer²⁶. Much like the Chapter III and Chapter IV χ^2 analyses, the $(\Delta S/S_0)^{exp}$ derived from the ^{13}C O-lipid ^{19}F experiments were compared to $(\Delta S/S_0)^{sim}$ by χ^2 analysis. However, in these experiments, the $(\Delta S/S_0)^{sim}$ depended upon two parameters: (1) the population of ^{13}C O nuclei in contact with one or more ^{19}F , and (2) the average dipolar coupling (or distance) between these ^{13}C O and the lipid ^{19}F . For all labeled ^{13}C O in mHFP, mV2E-HFP, and mHFPtr samples, typical best-fit distances between labeled ^{13}C O and lipid ^{19}F were ~7-8 Å, and typical best-fit populations ranged from ~0.00 to 0.40. One interpretation of these results is that detection of ~0.4 fractional population may correspond to obtaining maximum $(\Delta S/S_0)^{exp}$ values for samples with a 0.09 mole fraction of fluorinated lipid. The small molar fraction of fluorinated lipid may make maximal observable $(\Delta S/S_0)^{exp}$ ~0.4, instead of ~1.0 as could be observed with 1.0 mole fraction of fluorinated lipid. Another interpretation is that for any ^{19}F experiment using a 0.09 mol fraction of fluorinated lipid, 0.60

of the labeled ^{13}C 's have a different membrane location than the ^{13}C 's close to the ^{19}F s. The previous interpretation assumes that the unidentified ~ 0.6 fraction have the same membrane location as the identified ~ 0.4 fraction. These two interpretations lie on the extreme ends of how the ^{19}F data could be interpreted. A positive control experiment could be run to determine the precise meaning of these fractional populations by probing the membrane location using the KALP peptide where KALP peptides form transmembrane α helices in lipid vesicles composed of 1,2-Dilauroyl-*sn*-glycero-3-phosphocholine (di-C12:0-PC), 1,2-ditridecanoyl-*sn*-glycero-3-phospholcholine (di-C13:0-PC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (di-C14:0-PC), or 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (di-C18:1-PC)¹¹⁹.

While membrane location studies using ^{19}F - ^{13}C REDOR experiments provided some insight regarding the membrane location of HFP constructs, future membrane location studies could avoid the problems associated with fluorinated lipids by performing ^2H - ^{13}C REDOR^{120,121} experiments. Deuterated cholesterol is available through Sigma Aldrich in the form of cholesterol-2,2,3,4,4,6-d₆ and cholesterol-25,26,26,26,27,27,27-d₇, **Figure 36**. Cholesterol orientation within membranes has been studied by ^{13}C spin-lattice relaxation time experiments where ^{13}C spin-lattice relaxation times depended upon the proximity to the paramagnetic agents. These studies used vesicles prepared with ^{13}C substituted cholesterol and compared changes in ^{13}C spin-lattice relaxation times in the presence and absence of a paramagnetic agent to determine the average membrane location of the ^{13}C atoms at each position¹²². Similar experiments could be run using our NMR sample preparation methods to confirm the average cholesterol orientation in our NMR samples both before and after the addition of the HFP constructs. These studies are important since previous studies have shown that cholesterol position and orientation within a membrane is dependent upon membrane composition. Studies have suggested that cholesterol may lie parallel to the lipid acyl chain with the ringed structure end toward the membrane surface, and it has also been suggested that cholesterol can lie perpendicular to the lipid acyl chain in the center of the membrane^{123,124}.

These differences in membrane location result from different membrane compositions where membranes with higher ratios of poly-unsaturated lipid acyl chains, chains with multiple double bonds, resulted in the cholesterol being located in the center of the membrane. In general, double bonds in unsaturated lipid acyl chains result in “kinked” acyl chains, and membranes containing higher ratios of poly-unsaturated lipids are believed to pack together less favorable than the saturated lipid chains. These less ordered membranes were shown to be correlated to cholesterol being located in the center of the bilayer. While poly-unsaturated lipid chains were not used in the mHFP studies, mHFP constructs have been shown to insert into the membrane bilayer which disorders the membrane bilayer and could alter the cholesterol membrane location. Therefore, defining the cholesterol location before and after the addition of HFP is necessary before using cholesterol to probe mHFP’s membrane location since cholesterol may not be oriented parallel to the lipid acyl chain after the addition of HFP.

Aside from the membrane location data that could be obtained from ^2D - ^{13}C REDOR experiments, the presence of cholesterol in membranes favors β sheet secondary structure for HFP, and the cause of the β sheet conformational preference is not understood. Determining the location of HFP relative to cholesterol molecules could demonstrate if or where HFP is in contact with cholesterol which could contribute toward understanding why membranes with cholesterol favor β sheet structure over α helices. Additionally, the mV2E-HFP experiments demonstrated distinct differences in the ^{31}P - ^{13}CO distance for ^{13}CO labeled Leu-12 residues in β sheet conformation where best-fit ^{31}P - ^{13}CO distances and populations were $5.7 \pm 0.02 \text{ \AA}$, 0.39 ± 0.02 and $8.4 \text{ \AA} \pm 0.02$, 0.97 ± 0.02 for membranes with and without cholesterol, respectively. The difference in Leu-12 ^{13}CO proximity to ^{31}P could be due to β sheet stacking in the presence of cholesterol or aggregation of β sheets within a local cholesterol domain that may segregate HFP constructs away from the lipids. Either of these scenarios could additionally result in ^{19}F - ^{13}CO REDOR experiments that never reach 100% dephasing in the aforementioned samples prepared with fluorinated lipids. Alternatively, the increased ^{19}F - ^{13}CO and ^{31}P - ^{13}CO distances in membranes containing cholesterol may be due to the decreased lipid density within vesicles with cholesterol relative to vesicles without cholesterol.

Determining the proximity between mHFP constructs and cholesterol has scholarly value and may offer new ideas for inhibitory drug design that target the fusion peptide. Recently, inhibitory drug design has incorporated attachment of peptides to cholesterol¹²⁵. This method is attractive toward inhibiting HIV entry since the host cells that HIV infects contain membranes composed of ~30% cholesterol³⁷. Because cholesterol is native to the HIV host cells and the fusion peptide structure is altered by the addition of cholesterol to membranes, future schemes to inhibit HIV fusion may include attaching peptides or even small molecules to cholesterol that inhibit fusion active mHFP structures. In the context of this dissertation, that could involve attaching a small molecule or even peptide, possibly a variation of the V2E sequence, to cholesterol that shifts the distribution of β sheet registries toward longer registries. Attaching molecules to cholesterol has had preliminary success where the inhibitory effects of the C34 peptide towards HIV/host-cell fusion are enhanced when attached to cholesterol¹²⁵. If the fusion peptide is found to be proximate to cholesterol, attaching an inhibitor drug that targets the fusion peptide may prove to be an exceptionally efficient method for inhibitory drug design since the HIV host cells contain high concentrations of cholesterol. This may lengthen the time that the inhibitory drug is incorporated into the host cell which could improve its inhibitory effect as observed with the C34 peptide¹²⁵.

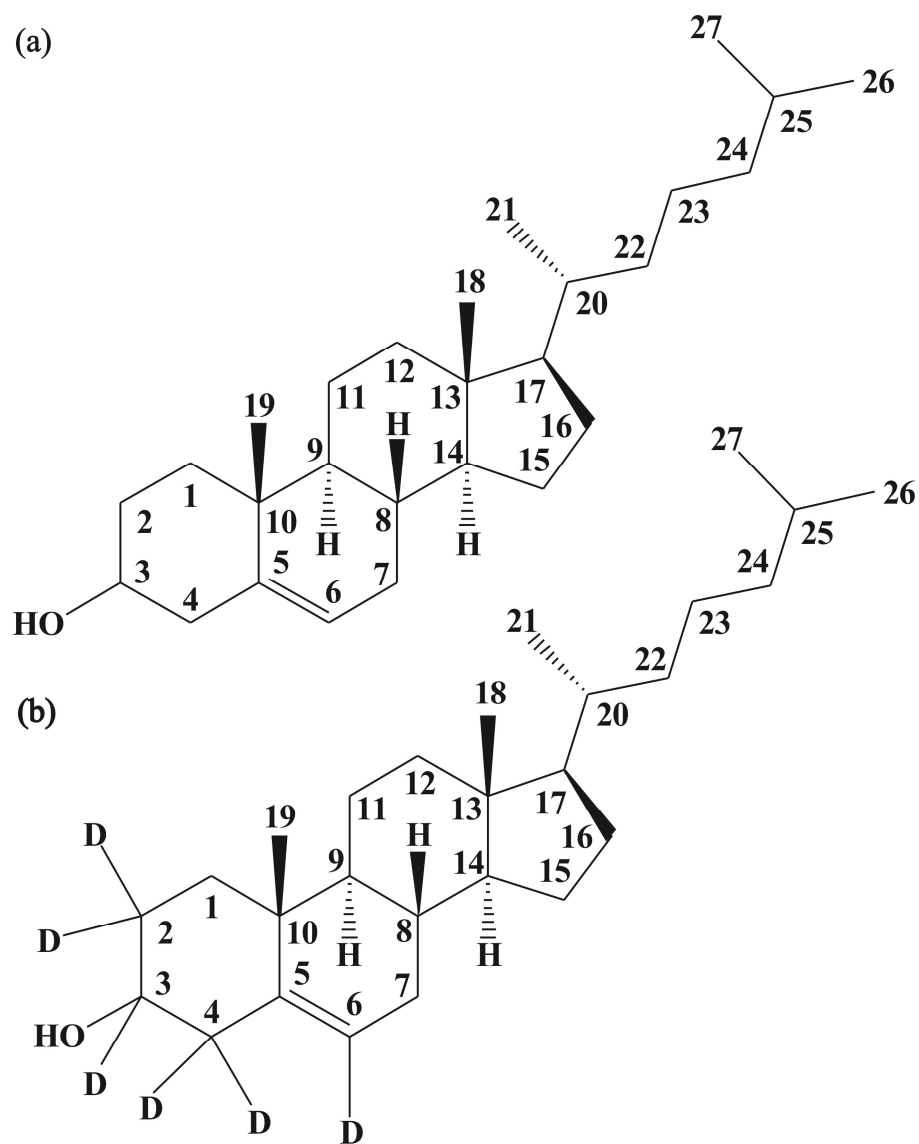
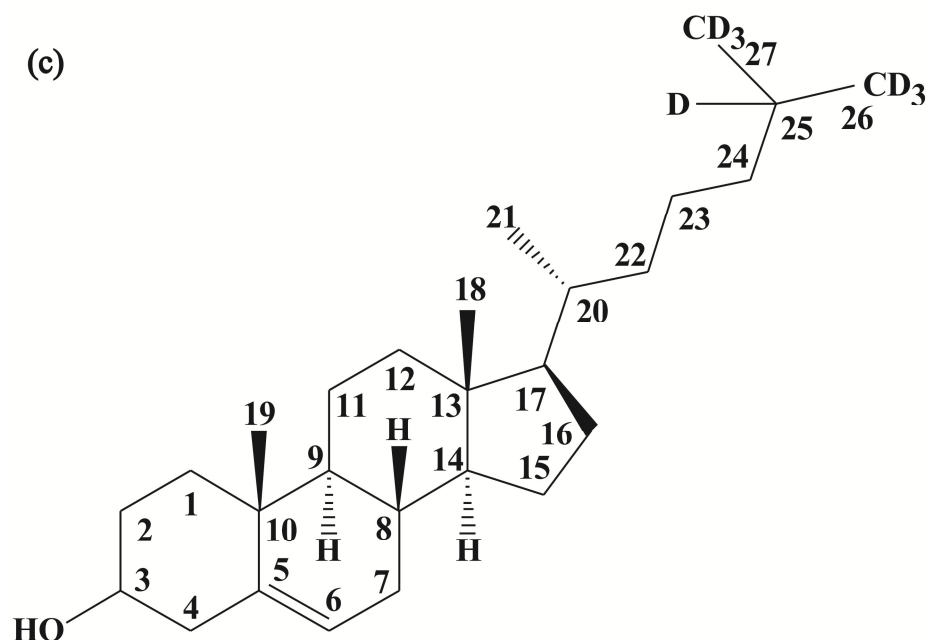


Figure 36. Cholesterol molecules with carbon atoms numbered (a) Cholesterol (b) Cholesterol-2,2,3,4,4,6-d₆ and (c) Cholesterol-25,26,26,26,27,27,27-d₇.

Figure 36 (cont'd)



5.3 Resin Bound Structure

Solid phase synthesis using Fmoc chemistry sequentially adds amino acids from the C-terminus to the N-terminus of the synthesized sequence. In general, the coupling step (i.e. attaching a residue) becomes less efficient for residues that are added later in the synthesis (N-terminal residues) than for residues that are added earlier in the synthesis (C-terminal residues)⁵⁷. For HFP, the N-terminal residues are hydrophobic and become more challenging to add than C-terminal residues²⁶. One potential explanation for this is that the added amino acids form aggregate structures, such as β sheets, while attached to the resin which may reduce the accessibility of the attachment site making the coupling reactions slower. To test the structural basis for this hypothesis, cross polarization spectra of a resin bound V2E-L9CI4N peptide was acquired, **Figure 37** and **Figure 38**. The Leu-9 carbonyl had a 174.3 ppm chemical shift which is similar to the 174 ppm chemical shift observed for β sheet HFP and different from the 179 ppm

chemical shift that is observed for helical HFP^{36,77}. Additionally, the Leu-9 has well defined secondary structure as evident by the 4.0 ppm full-width at half maximum height linewidth where as lyophilized HFP had line widths of ~7-8 ppm, Chapter III. These spectra serve as preliminary data to support the aforementioned hypothesis. By sampling more positions on HFP, one would hope to learn (1) where the β sheet structure begins; and (2) whether the β sheet forms as a result of peptide length or hydrophobicity. If the reasons for the formation of β sheet structure were well understood, new approaches for synthesis may be designed to enhance the efficiency of HFP synthesis and possibly the synthesis of other peptides as well. Additionally, sequences greater than ~30 residues are rarely synthesized due to poor yields of these sequences. If the efficiency were improved, synthesis of longer sequences may be more feasible.

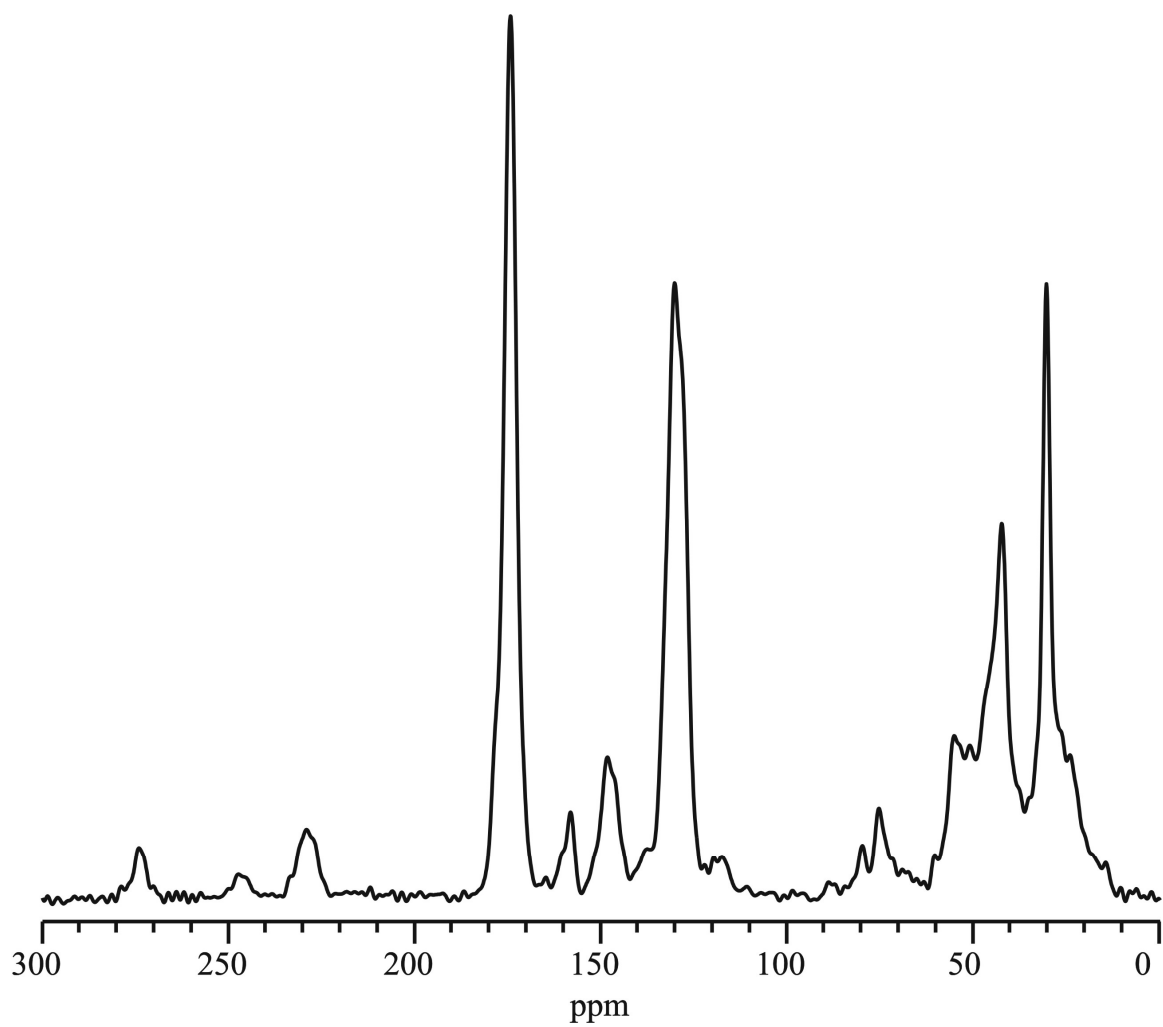


Figure 37. V2E-L9CI4N resin bound (prior to cleavage) with a MAS speed of 10 kHz.

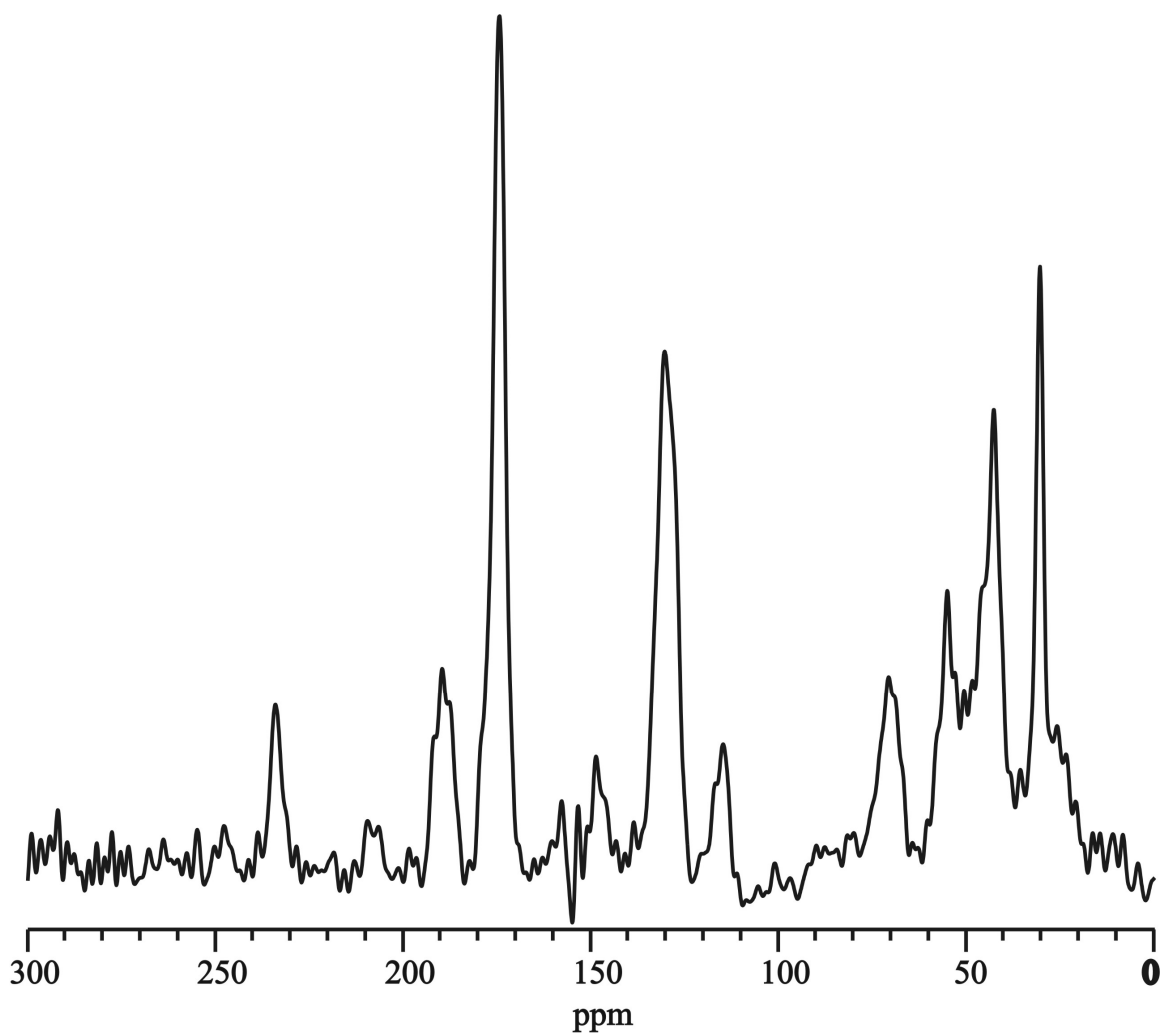


Figure 38. V2E-L9CI4N resin bound (prior to cleavage) with a MAS speed of 6 kHz.

APPENDICES

Appendix I. Files Checklist

Chapter 2

Figure 12 Magic Angle Spinning @ ..\home\sunyan4c\data\Scott\Setup\KBr091809

Magic Angle Spinning Procedure

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\
KBr_Magic_angle.pdf

Figure 13 ^1H Pulse @..\home\mb4b\data\Scott\REDOR\Setup\90pulse_array060610

Figure 14 CP @..\home\mb4b\data\Scott\REDOR\Setup\cp_array_012711

Figure 15 Cp_zfilter @..\home\mb4b\data\Scott\REDOR\Setup\cp_zfilter_012711

Figure 16 ^{15}N π pulse @..\home\mb4b\data\Scott\REDOR\Setup\15Narray_012711

Figure 17 @..\home\mb4b\data\Scott\REDOR\Setup\I4\I4_15NH1Rabi_arrays\20kHz_15N\
I4_15N1HRabi_array162

@..\home\mb4b\data\Scott\REDOR\Setup\I4\I4_15NH1Rabi_arrays\20kHz_15N\
I4_15N1HRabi_array242

@..\home\mb4b\data\Scott\REDOR\Setup\I4\I4_15NH1Rabi_arrays\20kHz_15N\
I4_15N1HRabi_array322

@..\home\mb4b\data\Scott\REDOR\Setup\I4\I4_15NH1Rabi_arrays\20kHz_15N\
I4_15N1HRabi_array402

@..\home\mb4b\data\Scott\REDOR\Setup\I4\I4_15NH1Rabi_arrays\20kHz_15N\
I4_15N1HRabi_array482

Figure 18 Adamantane @..\home\mb4b\data\Scott\REDOR\Setup\adam_102010

Chapter 3

SIMPSON SIMULATED Results @..\poohbah.chemistry.msu.edu\welikyshare\$\
SchmickThesis\Chp3Thesis\Par_anti_na.xls

Figures Coral Draw File @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\
Chp3Thesis\Figures_mod13.cdr

Figure 2a. @..\home\mb4b\data\Scott\REDOR2\neg_control\ -50_482_REDOR_050908

Figure 2b. @..\home\mb4b\data\Scott\REDOR2\2ndL12wG13A14\482_REDOR_041008

Figure 2c. @..\home\mb4b\data\Scott\REDOR2\L12wG5A6\482_REDOR_112407

Figure 2d. @..\home\mb4b\data\Scott\REDOR2\F8wL9G10\482REDOR_021708

Figure 3a,b. @..\home\mb4b\data\Scott\REDOR2\neg_control\
-50_482_REDOR_050908

-50_402_REDOR_050908

-50_322_REDOR_050908

-50_242_REDOR_050908

-50_162_REDOR_050908

-50_82_REDOR_050908

-50_22_REDOR_050908

Figure 3b. @..\home\mb4b\data\Scott\REDOR2\2ndL12wG13A14\
482_REDOR_041008

402_REDOR_041008

322_REDOR_041008

242_REDOR_041008

162_REDOR_041008

82_REDOR_041008

22_REDOR_041008
 @..\home\mb4b\data\Scott\REDOR2\L12wG5A6\
 482_REDOR_112407
 402_REDOR
 322_REDOR_112407
 242_REDOR_112407
 162_REDOR
 82_REDOR_112407
 22_REDOR
 @..\home\mb4b\data\Scott\REDOR2\F8wL9G10\
 482REDOR_021708
 402REDOR_021708
 322REDOR_021708
 242REDOR_021708
 162REDOR_021708
 82REDOR_021708
 22REDOR_021708

Chapter IV and all other. mHFP, mV2E-HFP, mHFPdm, and mHFPtr NMR Files.

Table 15. File directories for the 48.2, 40.2, 32.2, 24.2, 16.2, 8.2, and 2.2 ms data are shown.

@..\home\mb4b\data\Scott\REDOR\happi\3channel\Dimer\F8C-G13N\
 @..\home\mb4b\data\Scott\REDOR\dim\L12C-A5N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-A6N\dried\
 @..\home\mb4b\data\Scott\REDOR\dim\L12C-A6N\
 @..\home\mb4b\data\Scott\REDOR\dim\L9C-G5N\

Table 18. File directories for the 48.2, 40.2, 32.2, 24.2, 16.2, 8.2, and 2.2 ms data are shown.

@..\home\mb4b\data\Scott\REDOR\HFP\A6C-G3N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-A14N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-A15N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-A21N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-G13N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-G16N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-G3N\
 @..\home\mb4b\data\Scott\REDOR\HFP\F8C-L12N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-A6N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-G3N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-G5N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-I4N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L12C-L7N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L7C-G3N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L9C-G16N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L9C-G3N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L9C-G5N\
 @..\home\mb4b\data\Scott\REDOR\HFP\L9C-I4N\

Table 19. File directories for the 48.2, 40.2, 32.2, 24.2, 16.2, 8.2, and 2.2 ms data are shown.

@..\home\mb4b\data\Scott\REDOR\V2E\A6CG3N\
 @..\home\mb4b\data\Scott\REDOR\V2E\F8CA14N\

@..\home\mb4b\data\Scott\REDOR\V2E\F8CA15N\
 @..\home\mb4b\data\Scott\happi\3channel\V2E\F8CG13N\
 @..\home\mb4b\data\Scott\REDOR\V2E\F8CG16N\
 @..\home\mb4b\data\Scott\REDOR\V2E\F8CG3N\
 @..\home\mb4b\data\Scott\REDOR\V2E\F8C-L12N\062210\
 @..\home\mb4b\data\Scott\REDOR\V2E\L12CA6N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L12CG3N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L12CG5N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L12CI4N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L12CL7N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L7CG3N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L9CG16N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L9CG3N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L9CG5N\
 @..\home\mb4b\data\Scott\REDOR\V2E\L9CI4N\

Chapter IV. Excel Files for Iterative Fittings

HFP 3 Registry Fitting Fully Constrained @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\3Reg\F.C. WT compiled.xls

HFP 3 Registry Fitting Unconstrained

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\3Reg\2nd_U.C. WT.xls

V2E-HFP 3 Registry Fitting Fully Constrained

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\3Reg\F.C. V2E compiled.xls

HFP 5 Registry Fitting Fully Constrained @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\5Reg\2nd_F.C. WT compiled.xls

V2E-HFP 5 Registry Fitting Fully Constrained

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\5Reg\2nd_F.C. WT compiled.xls

Chapter IV and Appendices. HFP, V2E-HFP, L9R-HFP Excel Files for ΔG^{ins} , Fractional Populations Calculated from a Boltzmann Distribution using ΔG^{ins} for Energy, and the Modeled Membrane Insertion Depth.

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\HFP Chi^2_2nd.xls

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\V2E Chi^2_2nd.xls

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\L9R Chi^2_2nd.xls

Chapter V.

Figure 37. @..\home\mb4b\data\Scott\Thesis\resin_10kHz_V2E_L9CI4N_012011p

Figure 38. @..\home\mb4b\data\Scott\Thesis\resin_6kHz_V2E_L9CI4N_012011p

Appendix II. Freed Mutant Excel Spreadsheet Calculations.

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\1992_mutant.xls

Appendix IV. Clean HPLC Column Protocol.

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\cleancolumn.pdf
 L9R CP @..\home\mb4b\data\Scott\REDOR\L9R\F8CG13N\cp_ramp_F8CG13N_081210

Resin

@..\home\mb4b\data\Scott\cp\resin\cp10kHz_V2E_resinL9CI4N_012011

@..\home\mb4b\data\Scott\cp\resin\cp6kHz_V2E_resinL9CI4N_012011

SIMMOL and SIMPSON files

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\
Coordinate\Antiparallel_5_spin\Leu132_change_cord MOL file
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\Coordinate\
Antiparallel_5_spin\Leu132_change_cord SPINSYS file
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Antiparallel_5spin_4left\redor-ALA12_050608.in
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Antiparallel_5spin_4left\redor-ALA12-050608-122.45-138.46-49.79-31.824-82.087-131.98--
17.21--71.217-17.1718-44.177-16.768-19.126.fid
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Antiparallel_3spin_bleft\redor-ALA12_050608.in
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Antiparallel_3spin_bleft\redor-ALA12-050608-49.79-31.824--17.21--71.217-16.768-19.126.fid
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\
3_strand_CNC_REDOR\redor-98.0-050608.in
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\
3_strand_CNC_REDOR\redor-98.0-050608-0-98.0-0-0-12.35-12.97.fid
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Natural_abundance\b-sheet\redor-2spin-4.0-050608.in
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\Simpson\REDOR\REDOR_files\
Natural_abundance\b-sheet\redor-2spin-4.00-050608-0-0-48.2.fid

Fortran Files

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
HFPdata2\x2_fixed
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
HFPdata2\HFP.f
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\HFPdata2\
output
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
HFP2ndF8CG13N\x2_fixed
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
HFP2ndF8CG13N\HFP.f
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
HFP2ndF8CG13N\output
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\data\AppenThesis\Fortran\
5regHFP\f9=0.00\x2
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
5regHFP\f9=0.00\HFP_5var.f
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\5regHFP\
f9=0.00\output
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\5regHFP\
f9=0.01\x2
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\5regHFP\
f9=0.01\HFP_5var.f
@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\5regHFP\

f9=0.01\output
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\V2E_data\
 x2_V2E
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\V2E_data\
 V2E.f
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\V2E_data\
 output
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.00\x2_V2E
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.00\V2E_5var.f
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.00\output
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.01\x2_V2E
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.01\V2E_5var.f
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.01\output
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.02\x2_V2E
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.02\V2E_5var.f
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.02\output
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.03\x2_V2E
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.03\V2E_5var.f
 @..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\AppenThesis\Fortran\
 V2E5reg_data\f12=0.03\output

Appendix II. Current HIV Inhibitor Drugs

<u>Nucleoside/Nucleotide Reverse Transcriptase Inhibitors</u>	
Combivir (zidovudine + lamivudine)	
Emtriva (emtricitabine)	
Epivir (lamivudine)	
Epzicom (abacavir + lamivudine)	
Retrovir (zidovudine)	
Trizivir (abacavir + zidovudine + lamivudine)	
Truvada (tenofovir + emtricitabine)	
Videx EC (didanosine)	
Viread (tenofovir)	
Zerit (stavudine)	
Ziagen (abacavir)	
<u>Non-Nucleoside Reverse Transcriptase Inhibitors</u>	<u>Protease Inhibitors</u>
Atripia (efavirenz + tenofovir + emtricitabine)	Aptivus (tipranavir)
Intelence (etravirine)	Crixivan (indinavir)
Rescriptor (delavirdine)	Invirase (saquinavir)
Sustiva (efavirenz)	Kaletra (lopinavir + ritonavir)
	Lexiva (fosamprenavir)
	Norvir (ritonavir)
	Prezista (darunavir)
	Reyataz (atazanavir)
	Viracept (nelfinavir)
<u>Fusion and Entry Inhibitors</u>	
Fuzeon (enfuvirtide, injected)	
Selzentry (maraviroc)	
	<u>Integrase Inhibitors</u>
	Isentress (raltegravir)

Figure 39. Chart of commercially available anti HIV drugs. This chart was last updated 12\14\2010 and was taken from www.aidsmeds.com.

Appendix III. Simple Number of Strands for Fusion Model.

As mentioned in Chapters I, III, IV, and V, there is literature that supports that β sheet oligomers are a reasonable structure for the fusion active HFP structure. Additionally, syncytia fusion assays have demonstrated that an oligomeric structure is essential for membrane fusion where dilute amounts of V2E mutant gp41 were expressed with WT gp41 and fusion activity of WT gp41 was dominantly inhibited, **Table 12**²¹. A simplistic model was created to relate the number of strands within an oligomer required to initiate membrane fusion, s , as a function of experimentally observed fusion activity, $A(s)$. Incorporation of a single V2E gp41 strand into an oligomer was assumed to abrogate the fusion activity of an oligomer. Assuming that oligomerization of V2E gp41 with WT gp41 is random, the fusion activity as a function of s can be described by **Eq (59)**.

$$A(s) = F^s \quad (59)$$

where F is the fraction of WT gp41 from **Table 12**. The calculated activity was compared to the experimental activity by the χ^2 analysis metric.

$$\chi^2 = \sum_{j=1}^4 \frac{\{A(s)_j - (A)_j^{\text{exp}}\}^2}{(\sigma_j^{\text{exp}})^2} \quad (60)$$

where j was an index for each of the experimental activity data points from **Table 12**. The $\chi^2_{\text{min}} = 3.1$ for $s = 7$, Table 13, which suggested that 2-3 gp41 trimers were required to form fusion active β sheet oligomers. Additionally, χ^2 analysis was performed where the activity was modeled a fraction of 6 and a fraction of 9 strand β sheet oligomers (2 and 3 trimers) where the

$\chi^2_{\min} = 2.6$ which corresponded to 50:50 ratio of 6 strand to 9 strand oligomers in this model

(See @..\\poohbah.chemistry.msu.edu\\welikyshare\$\\SchmickThesis\\AppenThesis\\

1992_mutant.xls)

Table 12. Freed fusion activity²¹

Fraction of WT HFP (F)	0.50	0.67	0.83	0.91
Fusion Activity (σ^{exp})	0.04 (.03)	0.06 (.01)	0.18 (0.08)	0.39 (0.20)

Table 13. Strands Model 1.

Number of Strands (s)	χ^2
1	3989
2	1573
3	594
4	206
5	61
6	13
7	3
8	6
9	13
10	19
11	25
12	30

An alternative model was also considered where a minimum of two V2E gp41 strands were needed to be incorporated to abrogate the fusion activity of an oligomer.

$$A(s) = F^s + F^{(s-1)} \quad (61)$$

The $\chi^2_{\text{min}} = 14$ for $s = 10$, **Table 14**, which suggested that 3 or more gp41 trimers were required to form fusion active β sheet oligomers. While these models are not sufficient for determining the number of oligomers required for membrane fusion, both overwhelming suggest that more than one gp41 trimer is required for membrane fusion since $s = 3$ yields χ^2 values of 594 and 5,000 for the respective models. The excel spreadsheet with complete analysis was uploaded to the ftp.

Table 14. Strands Model 2.

Number of Strands (s)	χ^2
1	29000
2	12000
3	5000
4	2100
5	830
6	320
7	120
8	41
9	17
10	14
11	16
12	21
13	25
14	29
15	32

Appendix IV. RP-HPLC Purification, Optimization and Troubleshooting

From the currently available columns in our laboratory, the preferred column used for RP-HPLC was a Custom Bioseparation Symmetry 300 C4 steel column with 19 x 300mm dimensions, 5 μm particle size, and 300 Å pore size. The 19 mm diameter provides larger loading volumes than columns with 10 mm diameters, and the 5 μm particle size provided better peak resolution than columns with larger particle sizes, such as 10-15 μm .

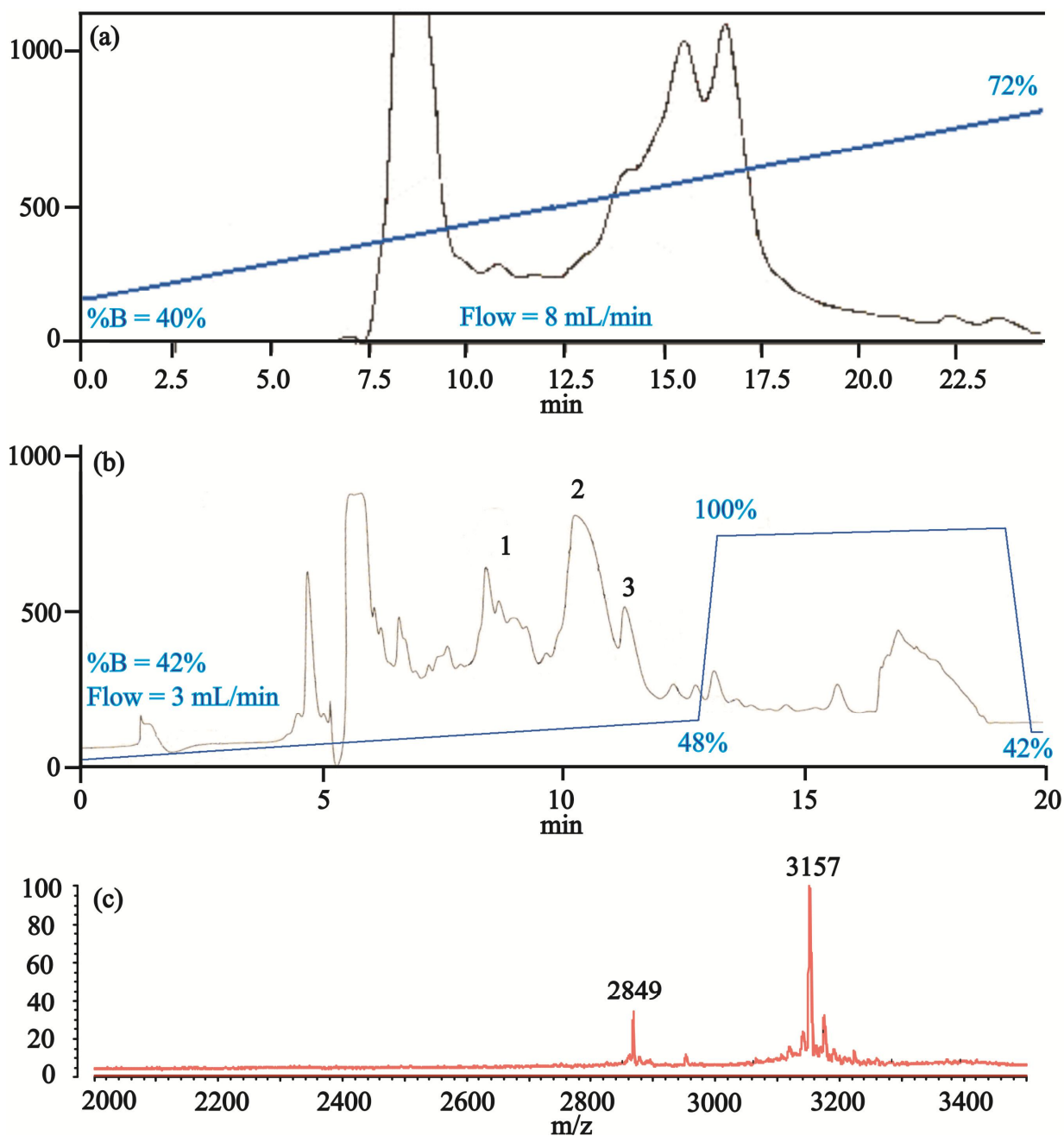


Figure 40. (a) HFP purification with a “large” C18 column (10-15 μm pore size). (b) HFP purification with a “small” C4 column (10 x 250 mm and 5 μm pore size). Better peak resolution was obtained with the C4 column. (c) Typical MALDI-TOF mass spectroscopy of peak 2 HFP from a purification similar to (b) where the expected mass was 3151 +2 g/mol where the +2 refers to the mass gain from the ^{13}C and ^{15}N isotopes.

The RP-HPLC chromatograms below, **Figure 41-Figure 46**, demonstrate how to develop a purification protocol to purify crude V2E-HFP. In these gradients, solvent A, 100% water with 0.01% TFA, was mixed with solvent B, 90% acetonitrile with 0.01% TFA. The flow rate of

solvent A + solvent B was set to 8 mL/min. Products or impurities that “elute” or pass through the column are detected at ~ 8 minutes under these conditions, Figure 42, and the peak at ~ 8 minutes is generally referred to as the elution peak. Therefore, we can approximate the percent elution, P_e , of a peak as a function of the solvent B gradient slope, G_s , the initial concentration of solvent B, C_i , and the elution time, t_e .

$$P_e = G_s (t_e - 8) + C_i \quad (62)$$

The figures below illustrate changes that can be made to the gradient to better resolve peaks.

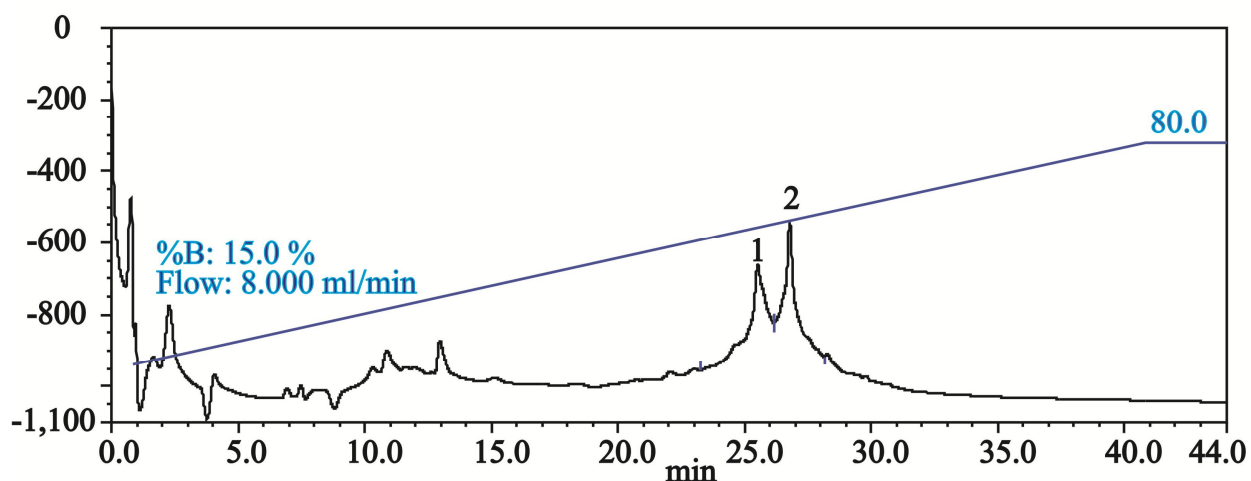


Figure 41. Preliminary gradient of 15% to 80% solvent B over 40 minutes. In developing purification protocols, small amounts of crude peptide were used to make product peaks narrow. By mass spec, Peak 2 is the confirmed product peak. From **Eq (62)**, the variables have the following values: $C_i = 15\%$; $G_s = 1.625 \text{ \%/min}$; $t_e \sim 26.5 \text{ min}$; and $P_e = 45\%$.

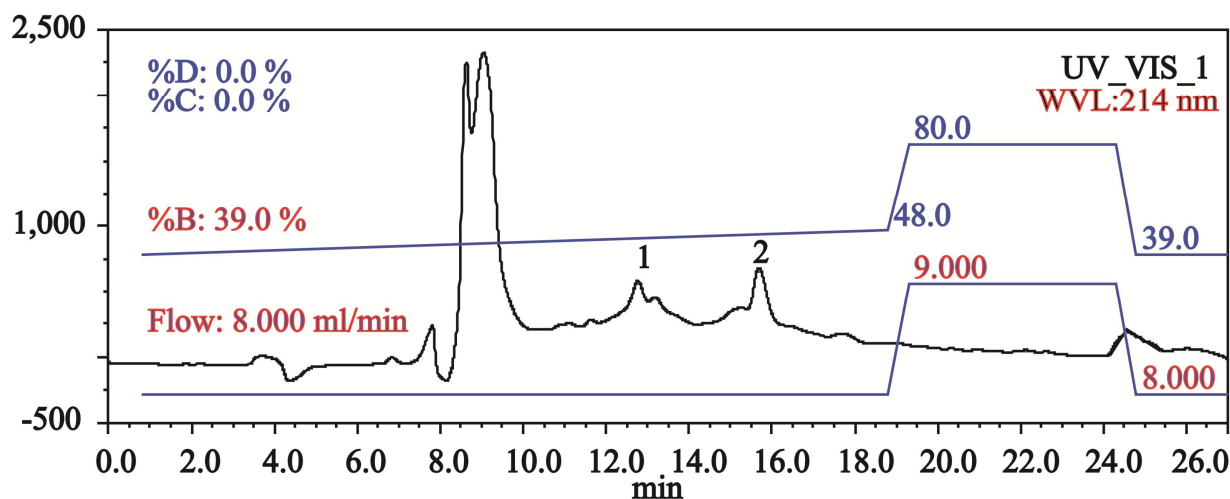


Figure 42. A linear 39-48% gradient was run over 18.5 minutes for purifying the product peak. From **Eq (62)**, the variables have the following values: $C_i = 39\%$; $G_s = 0.5 \text{ \%}/\text{min}$; $t_e \sim 15.5 \text{ min}$; and $P_e = 46.5\%$. The initial starting concentration was chosen to make the elution time around 15 minutes which was calculated by **Eq (62)** and a more gradual gradient was used to better separate peaks 1 and 2 from **Figure 41**. Also, at the end of the program, the gradient was ramped up to 80% solvent B over 0.5 minutes and the flow rate was increased to 9 mL/min to clean the column after each run. After 5 minutes, the gradient concentration of solvent B was returned to 39% over 0.5 minutes and the column equilibrated at this concentration for 3 minutes to prepare for the next run.

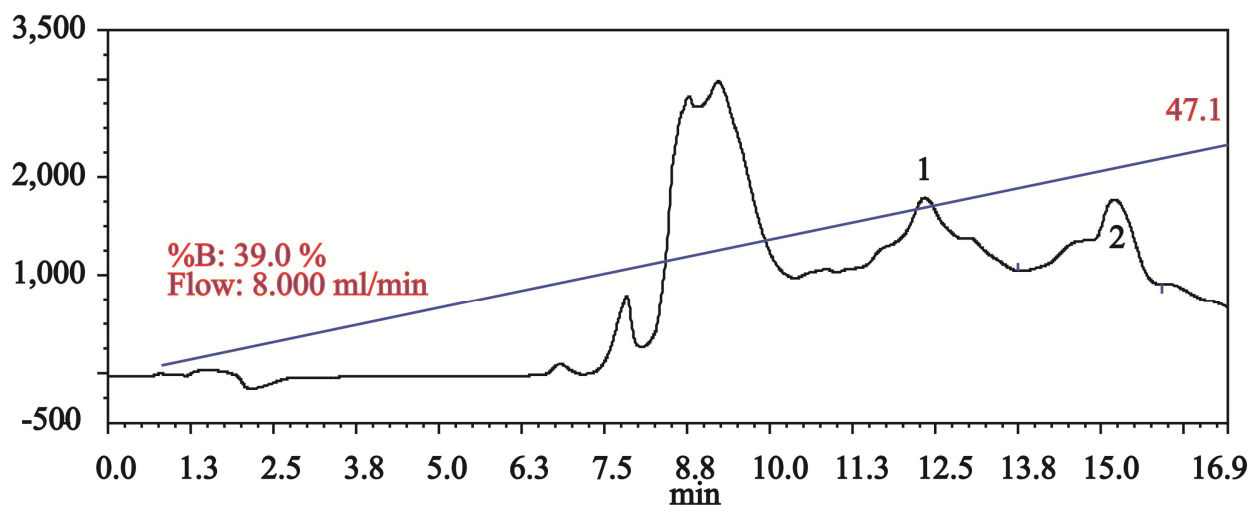


Figure 43. The program from **Figure 42** was used, but a higher loading volume of the crude peptide was used which resulted in poor separation of our product peak.

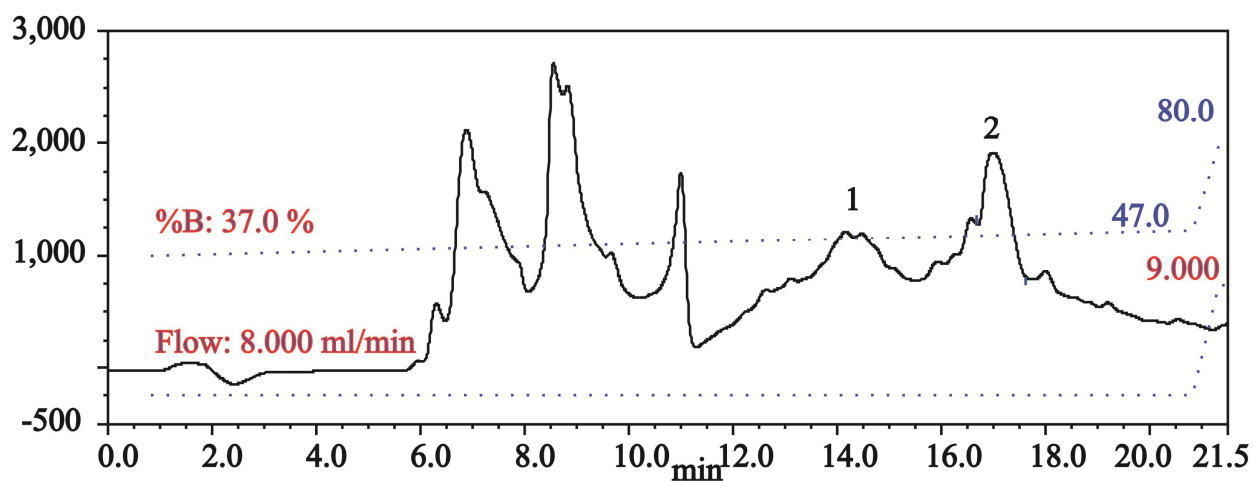


Figure 44. A linear 37-47% gradient was run over 20 minutes. The peaks were separated better with minimal peak broadening.

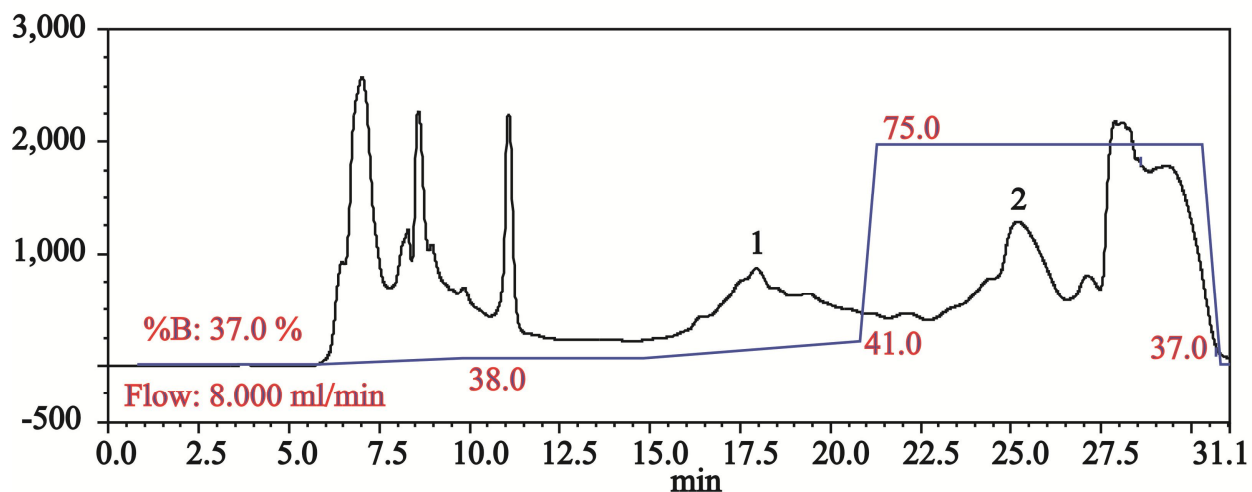


Figure 45. Nonlinear gradients can be used to separate peak 1 from peak 2. The gradient broadened peak 1 using a more gradual slope initially while the gradient was steeper from 15 to 21 minutes to retain the sharpness of peak 2. To optimize the time of the program, it's best to have your product elute during the period where the ramp is up to 75% solvent B since nothing is achieved during this time in **Figure 41-Figure 44**. Recall, the elution time of a peak is 8 minutes. Therefore, in this figure, peak 2 was collected during 75% use of solvent B, but peak 2 actually began coming off the column at time $t_e - 8$ or ~16-18 minutes.

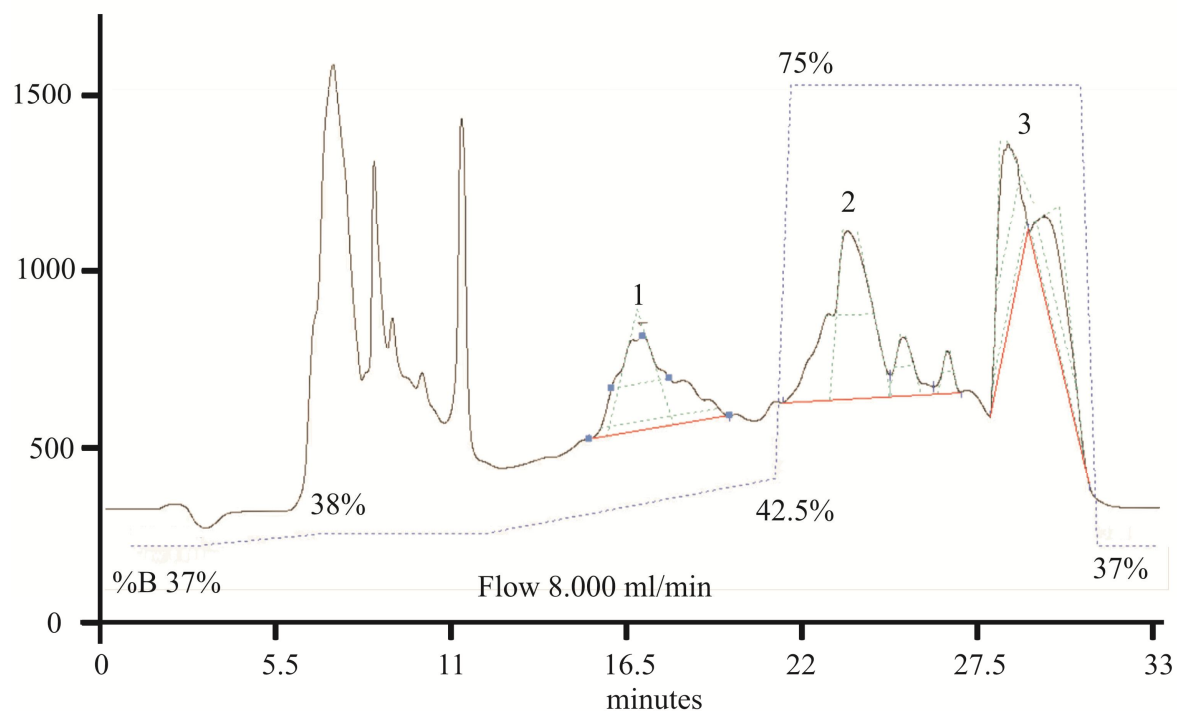


Figure 46. The ramp was modified to separate peak 2 from peak 3. This program was created because peak 1 also contains peptide with our products molecular weight which was collected for potential future use. Peak 3 should also be collected if the product peak is low relative to other syntheses. Peak 3 can contain HFP with N-terminal or sidechain protecting groups.

1. Specific Problems

1.1 Well separated peak contains “impurities” by MALDI-TOF analysis

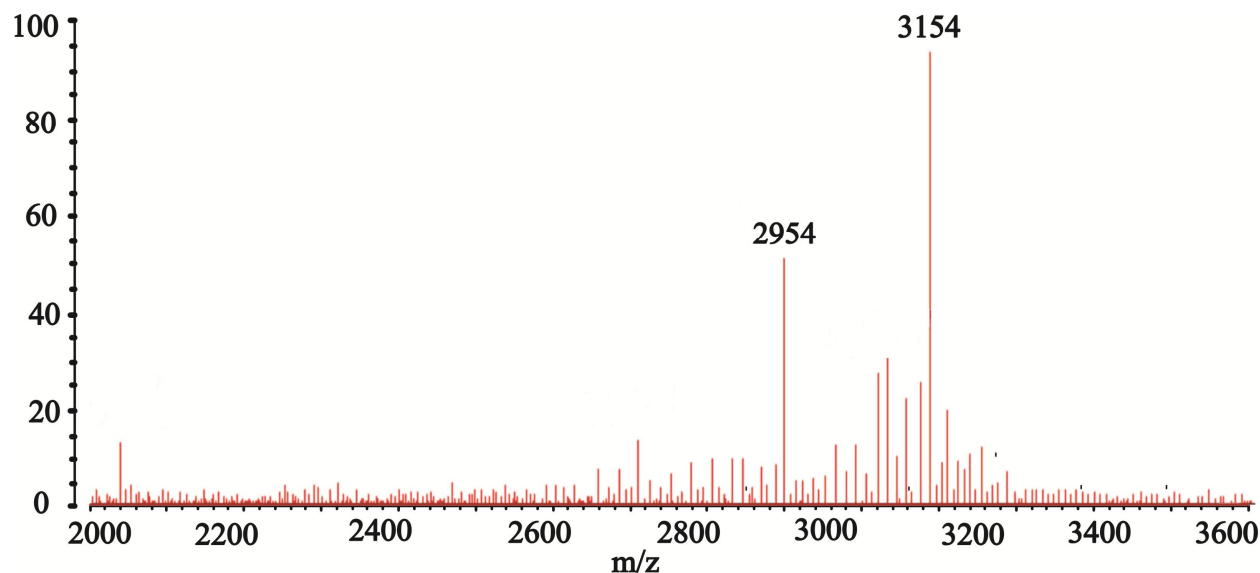


Figure 47. MALDI-TOF mass spectroscopy of purified HFP-L9G10. In MALDI-TOF experiments, increasing the laser power can increase the signal to noise, but it can also lead to peptide fragmentation where fragmentation can occur C-terminal of amino acids with basic sidechain groups^{65,66}. Alternatively, gas phase degradation of the peptide may be unlikely. The peptide degradation may result from hydrolytic cleavage in the matrix or possibly during isolation⁶⁴. Fragmentation of the HFP-L9G19 peptide appeared to occur C-terminal of the Arg-22, Lys-29, and Lys-30 where the respective fragments detected were likely AVGIGALFLGFLGAAGSTMGAR (2038 +2 g/mol), AVGIGALFLGFLGAAGSTMGARSWKKKKK (2952 +2 g/mol), and AVGIGALFLGFLGAAGSTMGARSWKKKKKK (3080 +2 g/mol), and the HFP product had an expected mass of 3151 + 2 g/mol.

1.2 Forgot to deprotect Fmoc group.

Dissolve crude peptide in 10 mL of 20% piperidine in DMF solution in a 50 mL conical vial and gently vortex for 20-25 minutes. Cap the conical vial and sonicate every 3 minutes to increase the solubility of the crude peptide. The crude peptide will not fully dissolve in the deprotection solution. Also, try to dissolve peptide that may be stuck to the conical vial.

1.3 Column pressure is increasing over time.

You may be loading precipitated peptides or dust onto the column in which case you can sonicate and centrifuge your peptide solution. Additionally, you can reverse the column and run a cleaning protocol. Note: when switching solvents, the pressure of the column will change due to swelling of the resin. The recommended flow rates are suggested to keep the column pressure low, but can be increased as long as you pay attention to the pressure. This procedure is effective in reducing the column pressure as it has reduced the “small” C4 column from ~2100-2500 psi to ~1300-1400 psi at a flow rate of 3 mL/min of 40% solvent B. A pdf file with a cleaning protocol can be found in:

@..\poohbah.chemistry.msu.edu\welikyshare\$\SchmickThesis\OtherThesis\cleancolumn.pdf

Appendix V. HFPdm Data and Lyophilized HFP

Chapter IV detected differences in registry distributions between the mV2E-HFP, mHFP, and mHFPtr constructs by using isotopic labeling schemes and the REDOR pulse sequence. Similar labeling schemes were incorporated into mHFPdm, and no clear structural differences were distinguished between mHFPdm and mHFP using this REDOR method. The structure of mHFP and mHFPdm may be identical or the differences are too subtle to detect using the current experimental design. Additionally, the HFP-L12CA6N sample was lyophilized for 24 hours after collection of the 7 data points and the sample was not rehydrated. The HFP-L12CA6N and lyophilized HFP-L12CA6N data were indistinguishable which suggests that mHFP and lyophilized mHFP have the same or similar registries.

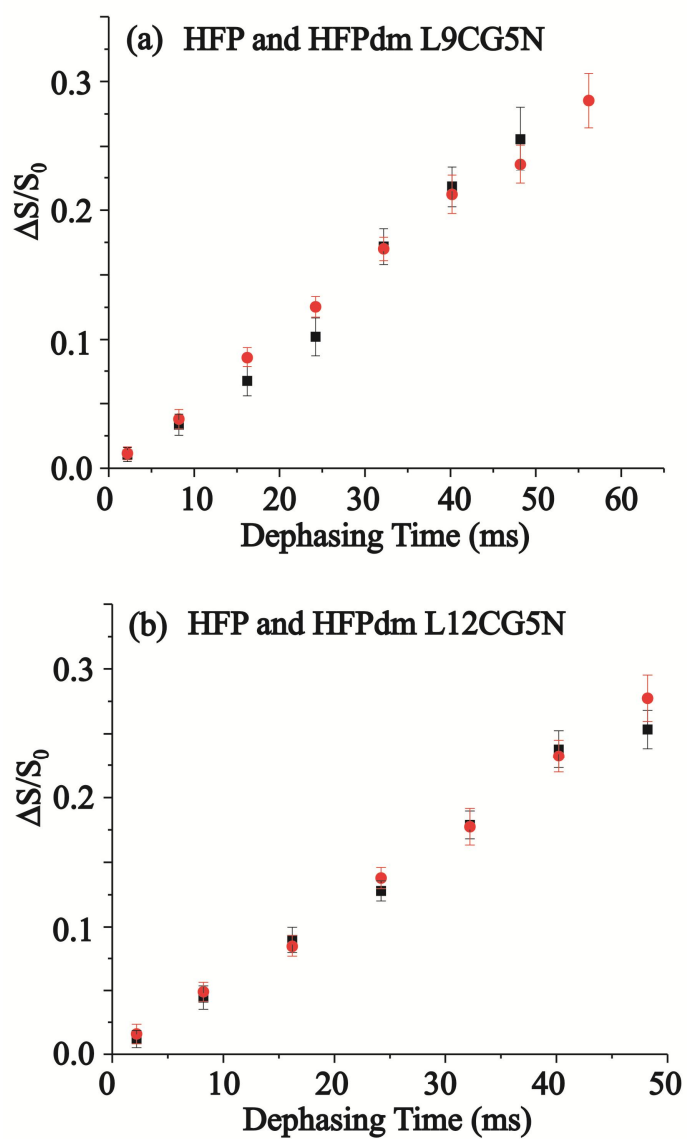


Figure 48. REDOR data for HFP (black boxes) and HFPdm (red circles) are displayed with error bars that are associated with rms deviation and labeling corresponds to (a) L9CG5N and (b) L12CG5N.

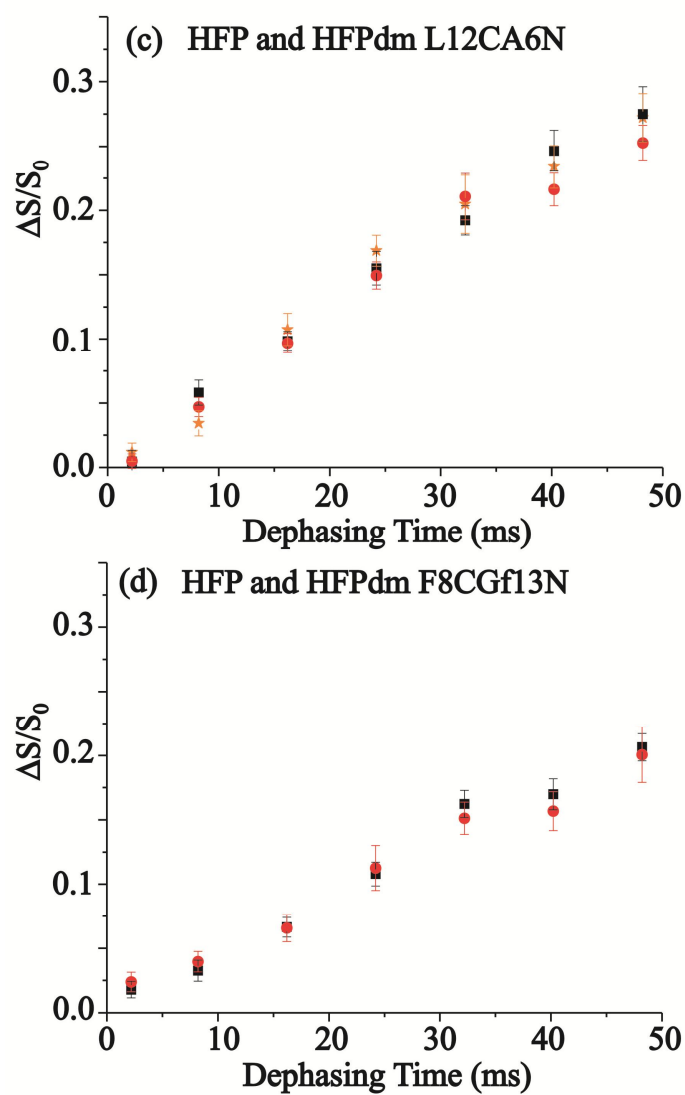


Figure 48 (cont'd). REDOR data for HFP (black boxes), lyophilized HFP (orange stars), and HFPdm (red circles) are displayed with error bars that are associated with rms deviation and labeling corresponds to (c) L12CA6N and (d) L12CG5N.

Table 15. HFPdm and L12CA6Nmn Lyophilized $(\Delta S/S_0)^{exp}$ and σ^{exp}

$(\Delta S/S_0)^{exp} + \sigma^{exp}$ in parentheses					
Dephasing time (ms)	L9C- G5Ndm	L12C- G5Ndm	L12C- A6Ndm	L12C- A6Nmn Lyophilized	F8C- G13Ndm
2.2	0.012 (.005)	0.016 (.008)	0.005 (.006)	0.012 (.007)	0.024 (.008)
8.2	0.039 (.007)	0.049 (.007)	0.047 (.007)	0.034 (.010)	0.040 (.008)
16.2	0.086 (.008)	0.085 (.008)	0.097 (.008)	0.107 (.013)	0.066 (.011)
24.2	0.125 (.008)	0.138 (.008)	0.149 (.011)	0.161 (.012)	0.113 (.018)
32.2	0.170 (.009)	0.178 (.014)	0.211 (.018)	0.205 (.017)	0.151 (.013)
40.2	0.212 (.015)	0.232 (.013)	0.216 (.013)	0.234 (.017)	0.157 (.015)
48.2	0.236 (.015)	0.277 (.018)	0.253 (.013)	0.272 (.019)	0.201 (.020)

Appendix VI. SIMMOL, SIMPSON, and Fortran Files.

1. Sample SIMMOL Files

1.1 SIMMOL input file for Leu-132 ¹³CO from the 2IWW.pdb file.

```
regsub ".mol" $argv0 {\1.oogl} oogl
regsub ".mol" $argv0 {\1.spinsys} spinsys

set m [mload "2IWW.pdb"]
mloadtensors $m -default
#mloadjcouplings $m -default
msetspinsysfile $m $spinsys -numbered

mselect $m 1 atom 536
mselect $m 2 atom 357
mselect $m 3 atom 352
mselect $m 4 atom 344
mselect $m 5 atom 727
mselect $m 6 atom 716
mselect $m 7 atom 705

#mset $m -solid -ellipsoid shielding -color cpk -nice
mdipole $m 1 2 0AA 20AA
mdipole $m 1 3 0AA 20AA
mdipole $m 1 4 0AA 20AA
mdipole $m 1 5 0AA 20AA
mdipole $m 1 6 0AA 20AA
mdipole $m 1 7 0AA 20AA

#mclosespinsysfile $m
munload $m
puts "Generated: $spinsys"
```

1.2. SIMMOL output file for the Leu-132 residue ¹³CO from the 2IWW.pdb file

```
spinsys {
#      1      2      3      4      5      6      7
#  536N   357N   352N   344C   727N   716N   705N
#
  channels 15N 13C
  nuclei   15N 15N 15N 13C 15N 15N 15N
  dipole 1 4 21.0712 0 101.81 72.787
  dipole 2 4 35.747 0 98.531 33.511
```

```

dipole 3 4 13.194 0 72.5 9.7837
dipole 4 5 16.339 0 76.984 9.9556
dipole 4 6 19.871 0 103.81 36.023
dipole 4 7 12.326 0 129.81 61.397
}

```

2. Sample SIMPSON Files

2.1. Input file 5 spin (NNCNN) SIMPSON File for Ala-12 ¹³CO from 2WIL.pdb

```

# REDOR simulation for pulse sequence redorxy8xy_pm for a model three spin
system
# ver. 1.0, last revised on 02/05/07 w/o all real pulses
# no cp phase cycling
# no output of s0 and s1

spinsys {
channels 13C 15N
nuclei   13C 15N 15N 15N 15N
dipole   1 2 $par(dp12) 0 $par(a) $par(e)
dipole   1 3 $par(dp13) 0 $par(b) $par(f)
dipole   1 4 $par(dp14) 0 $par(c) $par(g)
dipole   1 5 $par(dp15) 0 $par(d) $par(h)
shift 1 174.3p -75.0p 0.97333 0 109.37 -171.38
}

par {
proton_frequency 400.7797840e6
spin_rate        10000

sw                50000

np                7

crystal_file      rep320

dipole_check      false
gamma_angles      18
  start_operator   I1x

detect_operator    I1p
verbose            1101

variable Crf       61728

variable Nrf       61728
variable dp12_min  44.177
variable dp12_max  44.177
variable dp12_incr  1
variable dp13_min  12.43
variable dp13_max  12.43
variable dp13_incr  1
variable dp14_min  19.126
variable dp14_max  19.126
variable dp14_incr  1

```

```

variable dp15_min 13.485
variable dp15_max 13.485
variable dp15_incr 1
variable a_min 138.46
variable a_max 138.46
variable a_incr 1
variable b_min 118.5
variable b_max 118.5
variable b_incr 1
variable c_min 31.824
variable c_max 31.824
variable c_incr 1
variable d_min 40.492
variable d_max 40.492
variable d_incr 1
variable e_min 131.98
variable e_max 131.98
variable e_incr 1
variable f_min 165.09
variable f_max 165.09
variable f_incr 1
variable g_min -71.217
variable g_max -71.217
variable g_incr 1
variable h_min -132.18
variable h_max -132.18
variable h_incr 1
variable bestkaisqr 1e6
variable c_off 15398
variable n_off 0

```

```

}

```

```

proc pulseseq { } {
global par

maxdt 1
set Ct180 [expr 0.5e6/$par(Crf)]
set Nt180 [expr 0.5e6/$par(Nrf)]
set tr [expr 0.5e6/$par(spin_rate)]
set tr1 [expr $tr-0.5*$Ct180]
set tr2 [expr $tr-0.5*$Ct180-0.5*$Nt180]

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
    pulse $Ct180 $par(Crf) x 0 x
store 1
#s1

reset
offset $par(c_off) $par(n_off)

```

```

delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 2
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr1
store 3
#s1

reset
  foreach i {1 2 1 2 2 1 2 1} {
prop $i
  }

store 8
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
  pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
store 4
#s0
  reset offset $par(c_off) $par(n_off)
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 5
#s0

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr1
store 6

```

```

    #s0
    reset
    foreach i {4 5 4 5 5 4 5 4} {
    prop $i
    }

    store 9
    #s0

    foreach j {2 8 16 24 32 40 48} {
    reset
    prop $par(swth1) $j
    prop $par(swth2)
    acq
    }

}

proc main { } {
global par
for {set par(a) $par(a_min)} {$par(a) <= $par(a_max)} {set par(a) [expr
$par(a)+$par(a_incr)]} {
for {set par(b) $par(b_min)} {$par(b) <= $par(b_max)} {set par(b) [expr
$par(b)+$par(b_incr)]} {
for {set par@c $par(c_min)} {$par@c <= $par(c_max)} {set par@c [expr
$par@c+$par(c_incr)]} {
for {set par(d) $par(d_min)} {$par(d) <= $par(d_max)} {set par(d) [expr
$par(d)+$par(d_incr)]} {
for {set par(e) $par(e_min)} {$par(e) <= $par(e_max)} {set par(e) [expr
$par(e)+$par(e_incr)]} {
for {set par(f) $par(f_min)} {$par(f) <= $par(f_max)} {set par(f) [expr
$par(f)+$par(f_incr)]} {
for {set par(g) $par(g_min)} {$par(g) <= $par(g_max)} {set par(g) [expr
$par(g)+$par(g_incr)]} {
for {set par(h) $par(h_min)} {$par(h) <= $par(h_max)} {set par(h) [expr
$par(h)+$par(h_incr)]} {
for {set par(dp12) $par(dp12_min)} {$par(dp12) <= $par(dp12_max)} {set
par(dp12) [expr $par(dp12)+$par(dp12_incr)]} {
for {set par(dp13) $par(dp13_min)} {$par(dp13) <= $par(dp13_max)} {set
par(dp13) [expr $par(dp13)+$par(dp13_incr)]} {
for {set par(dp14) $par(dp14_min)} {$par(dp14) <= $par(dp14_max)} {set
par(dp14) [expr $par(dp14)+$par(dp14_incr)]} {
for {set par(dp15) $par(dp15_min)} {$par(dp15) <= $par(dp15_max)} {set
par(dp15) [expr $par(dp15)+$par(dp15_incr)]} {
    foreach p {{0 9 6} {1 8 3}} {
set par(swth1) [lindex $p 1]
set par(swth2) [lindex $p 2]
set f [fsimpson]
        set g[lindex $p 0] [fdup $f]
    }

#fsave $g0 $par(name)-$par(dp12)-s0.fid
#fsave $g1 $par(name)-$par(dp12)-s1.fid
        set gsub [fdup $g0]
fsub  $gsub $g1

```



```

spinsys {
channels 13C 15N
nuclei   13C 15N 15N
dipole   1 2 $par(dp12) 0 $par(a) $par(c)
dipole   1 3 $par(dp13) 0 $par(b) $par(d)
shift 1 174.3p -75.0p 0.97333 0 109.37 -171.38
}
par {
proton_frequency 400.7797840e6
spin_rate        10000

sw                50000

np                7

crystal_file      rep320

dipole_check      false
gamma_angles      18
  start_operator   I1x

detect_operator    I1p
verbose            1101

variable Crf       61728

variable Nrf        61728
variable dp12_min  16.768
variable dp12_max  16.768
variable dp12_incr    1
variable dp13_min  19.126
variable dp13_max  19.126
variable dp13_incr    1
variable a_min      49.79
variable a_max      49.79
variable a_incr      1
variable b_min      31.824
variable b_max      31.824
variable b_incr      1
variable c_min      -17.21
variable c_max      -17.21
variable c_incr      1
variable d_min      -71.217
variable d_max      -71.217
variable d_incr      1
variable bestkaisqr    1e6
variable c_off 15398
variable n_off 0
}

proc pulseseq { } {
global par

maxdt 1

```

```

set Ct180 [expr 0.5e6/$par(Crf)]
set Nt180 [expr 0.5e6/$par(Nrf)]
set tr [expr 0.5e6/$par(spin_rate)]
set tr1 [expr $tr-0.5*$Ct180]
set tr2 [expr $tr-0.5*$Ct180-0.5*$Nt180]

```

```

reset
offset $par(c_off) $par(n_off)

```

```

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
    pulse $Ct180 $par(Crf) x 0 x
store 1
#s1

```

```

reset
offset $par(c_off) $par(n_off)

```

```

delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 2
#s1

```

```

reset
offset $par(c_off) $par(n_off)

```

```

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr1
store 3
#s1

```

```

reset
    foreach i {1 2 1 2 2 1 2 1} {
prop $i
    }

```

```

store 8
#s1

```

```

reset
offset $par(c_off) $par(n_off)

```

```

delay $tr2
    pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
store 4
#s0

```

```

    reset    offset $par(c_off) $par(n_off)
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 5
#s0

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr1
store 6
    #s0
reset
foreach i {4 5 4 5 5 4 5 4} {
prop $i
}

store 9
#s0

foreach j {2 8 16 24 32 40 48} {
reset
prop $par(swth1) $j
prop $par(swth2)
acq
}

}

proc main { } {
global par
for {set par(a) $par(a_min)} {$par(a) <= $par(a_max)} {set par(a) [expr
$par(a)+$par(a_incr)]} {
for {set par(b) $par(b_min)} {$par(b) <= $par(b_max)} {set par(b) [expr
$par(b)+$par(b_incr)]} {
for {set par@c $par(c_min)} {$par@c <= $par(c_max)} {set par@c [expr
$par@c+$par(c_incr)]} {
for {set par(d) $par(d_min)} {$par(d) <= $par(d_max)} {set par(d) [expr
$par(d)+$par(d_incr)]} {
for {set par(dp13) $par(dp13_min)} {$par(dp13) <= $par(dp13_max)} {set
par(dp13) [expr $par(dp13)+$par(dp13_incr)]} {
for {set par(dp12) $par(dp12_min)} {$par(dp12) <= $par(dp12_max)} {set
par(dp12) [expr $par(dp12)+$par(dp12_incr)]} {
    foreach p {{0 9 6} {1 8 3}} {
set par(swth1) [lindex $p 1]
set par(swth2) [lindex $p 2]
set f [fsimpson]

```

```

        set g[lindex $p 0] [fdup $f]
    }

#fsave $g0 $par(name)-$par(dp12)-s0.fid
#fsave $g1 $par(name)-$par(dp12)-s1.fid
    set gsub [fdup $g0]
fsub  $gsub $g1
    for {set j 1} {$j <= $par(np)} {incr j} {
        set a0 [findex $g0 $j -re]
    set a1 [findex $gsub $j -re]
fsetindex $g1 $j [expr $a1/$a0] 0
}

    #fphase $g1 -scale $par(scl_avg)
fsave $g1 $par(name)-$par(a)-$par(b)-$par(c)-$par(d)-$par(dp12)-
$par(dp13).fid
}
}
}
}
}
}
#set outpt [open $par(name).out a+ 0600]
#puts $outpt "$dp12_opt $par(bestkaisqr)"
#close $outpt
#funload
}

```

2.4 Output file 3 spin (CNN) SIMPSON File for Ala-12 ¹³CO from 2WIL.pdb

```

SIMP
NP=7
SW=50000
TYPE=FID
DATA
0.0021804517 0
0.0289860854 0
0.108652919 0
0.228572947 0
0.373536138 0
0.526272979 0
0.670419088 0
END

```

2.5 Input file 3 spin (CNC) SIMPSON File from 2WIL.pdb

```

# REDOR simulation for pulse sequence redorxy8xy_pm for a model three spin
system
# ver. 1.0, last revised on 02/05/07 w/o all real pulses
# no cp phase cycling
# no output of s0 and s1

spinsys {

```

```

channels 13C 15N
nuclei 13C 15N 13C
dipole 1 2 $par(dp12) 0 $par(a) $par(c)
dipole 2 3 $par(dp23) 0 $par(b) $par(d)
shift 1 174.3p -75.0p 0.97333 0 109.37 -171.38
    }
par {
proton_frequency 400.7797840e6
spin_rate 10000

sw 50000

np 7

crystal_file rep320

dipole_check false
gamma_angles 18
    start_operator I1x

detect_operator I1p
verbose 1101

variable Crf 61728

variable Nrf 61728
variable dp12_min 12.35
variable dp12_max 12.35
variable dp12_incr 1
variable dp23_min 12.97
variable dp23_max 12.97
variable dp23_incr 1
variable a_min 0
variable a_max 0
variable a_incr 1
variable b_min 98.0
variable b_max 98.0
variable b_incr 1
variable c_min 0
variable c_max 0
variable c_incr 1
variable d_min 0
variable d_max 0
variable d_incr 1
variable bestkaisqr 1e6
variable c_off 15398
variable n_off 0
}

proc pulseseq { } {
global par

maxdt 1
set Ct180 [expr 0.5e6/$par(Crf)]
set Nt180 [expr 0.5e6/$par(Nrf)]

```

```

set tr [expr 0.5e6/$par(spin_rate)]
set tr1 [expr $tr-0.5*$Ct180]
set tr2 [expr $tr-0.5*$Ct180-0.5*$Nt180]

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
store 1
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 2
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr1
store 3
#s1

reset
foreach i {1 2 1 2 2 1 2 1} {
prop $i
}

store 8
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
store 4
#s0
reset offset $par(c_off) $par(n_off)
delay $tr2

```

```

pulse $Nt180 0 y 0 y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 5
#s0

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr1
store 6
#s0
reset
foreach i {4 5 4 5 5 4 5 4} {
prop $i
}

store 9
#s0

foreach j {2 8 16 24 32 40 48} {
reset
prop $par(swth1) $j
prop $par(swth2)
acq
}

}

proc main { } {
global par
for {set par(a) $par(a_min)} {$par(a) <= $par(a_max)} {set par(a) [expr
$par(a)+$par(a_incr)]} {
for {set par(b) $par(b_min)} {$par(b) <= $par(b_max)} {set par(b) [expr
$par(b)+$par(b_incr)]} {
for {set par© $par(c_min)} {$par© <= $par(c_max)} {set par© [expr
$par©+$par(c_incr)]} {
for {set par(d) $par(d_min)} {$par(d) <= $par(d_max)} {set par(d) [expr
$par(d)+$par(d_incr)]} {
for {set par(dp23) $par(dp23_min)} {$par(dp23) <= $par(dp23_max)} {set
par(dp23) [expr $par(dp23)+$par(dp23_incr)]} {
for {set par(dp12) $par(dp12_min)} {$par(dp12) <= $par(dp12_max)} {set
par(dp12) [expr $par(dp12)+$par(dp12_incr)]} {
foreach p {{0 9 6} {1 8 3}} {
set par(swth1) [lindex $p 1]
set par(swth2) [lindex $p 2]
set f [fsimpson]
set g[lindex $p 0] [fdup $f]
}
}
}

```



```

#fsave $g0 $par(name)-$par(dp12)-s0.fid
#fsave $g1 $par(name)-$par(dp12)-s1.fid
      set gsub [fdup $g0]
fsub  $gsub $g1
      for {set j 1} {$j <= $par(np)} {incr j} {
          set a0 [findex $g0 $j -re]
set a1 [findex $gsub $j -re]
fsetindex $g1 $j [expr $a1/$a0] 0

}

      #fphase $g1 -scale $par(scl_avg)
fsave $g1 $par(name)-$par(a)-$par(b)-$par(c)-$par(d)-$par(dp12)-
$par(dp23).fid
}
}
}
}
}
}
#set outpt [open $par(name).out a+ 0600]
#puts $outpt "$dp12_opt $par(bestkaisqr)"
#close $outpt
#funload

}

```

2.6 Output file 3 spin (CNC) SIMPSON File from 2WIL.pdb

```

SIMP
NP=7
SW=50000
TYPE=FID
DATA
0.000517198113 0
0.00693752805 0
0.0267554081 0
0.0589735416 0
0.102783454 0
0.157079165 0
0.220491486 0
END

```

2.7 Input file 2 spin (CN) SIMPSON File from 2WIL.pdb

```

# REDOR simulation for pulse sequence redorxy8xy_pm for a model three spin
system
# ver. 1.0, last revised on 02/05/07 w/o all real pulses
# no cp phase cycling
# no output of s0 and s1

spinsys {
channels 13C 15N

```

```

nuclei    13C 15N
dipole    1 2 $par(dp12) 0 0 0
shift 1 174.3p -75.0p 0.97333 0 109.37 -171.38
}
par {
proton_frequency 400.7797840e6
spin_rate        10000

sw                50000

np                13

crystal_file      rep320

dipole_check      false
gamma_angles      18
  start_operator   I1x

detect_operator    I1p
verbose            1101

variable Crf       61728

variable Nrf       61728
variable dp12_min  48.125
variable dp12_max  48.125
variable dp12_incr      1
variable a_min        0
variable a_max        0
variable a_incr       1
variable e_min        0
variable e_max        0
variable e_incr       1
variable bestkaisqr    1e6
variable c_off 15398
variable n_off 0
}

proc pulseseq { } {
global par

maxdt 1
set Ct180 [expr 0.5e6/$par(Crf)]
set Nt180 [expr 0.5e6/$par(Nrf)]
set tr [expr 0.5e6/$par(spin_rate)]
set tr1 [expr $tr-0.5*$Ct180]
set tr2 [expr $tr-0.5*$Ct180-0.5*$Nt180]

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2

```

```

    pulse $Ct180 $par(Crf) x 0 x
store 1
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 2
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x $par(Nrf) x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y $par(Nrf) y
delay $tr1
store 3
#s1

reset
    foreach i {1 2 1 2 2 1 2 1} {
prop $i
    }

store 8
#s1

reset
offset $par(c_off) $par(n_off)

delay $tr2
    pulse $Nt180 0 x 0 x
delay $tr2
pulse $Ct180 $par(Crf) x 0 x
store 4
#s0
    reset offset $par(c_off) $par(n_off)
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr2
pulse $Ct180 $par(Crf) y 0 y
store 5
#s0

reset
offset $par(c_off) $par(n_off)

delay $tr2
pulse $Nt180 0 x 0 x

```

```

delay $tr2
pulse $Ct180 $par(Crf) x 0 x
delay $tr2
pulse $Nt180 0 y 0 y
delay $tr1
store 6
    #s0
reset
foreach i {4 5 4 5 5 4 5 4} {
prop $i
    }

store 9
#s0

foreach j {2 8 16 24 32 40 48 56 64 72 80 88 96} {
reset
prop $par(swth1) $j
prop $par(swth2)
acq
    }

}

proc main { } {
global par
for {set par(a) $par(a_min)} {$par(a) <= $par(a_max)} {set par(a) [expr
$par(a)+$par(a_incr)]} {
for {set par(e) $par(e_min)} {$par(e) <= $par(e_max)} {set par(e) [expr
$par(e)+$par(e_incr)]} {
for {set par(dp12) $par(dp12_min)} {$par(dp12) <= $par(dp12_max)} {set
par(dp12) [expr $par(dp12)+$par(dp12_incr)]} {
    foreach p {{0 9 6} {1 8 3}} {
set par(swth1) [lindex $p 1]
set par(swth2) [lindex $p 2]
set f [fsimpson]
        set g[lindex $p 0] [fdup $f]
    }

#fsave $g0 $par(name)-$par(dp12)-s0.fid
#fsave $g1 $par(name)-$par(dp12)-s1.fid
        set gsub [fdup $g0]
fsub $gsub $g1
        for {set j 1} {$j <= $par(np)} {incr j} {
            set a0 [findex $g0 $j -re]
set a1 [findex $gsub $j -re]
fsetindex $g1 $j [expr $a1/$a0] 0
        }

        #fphase $g1 -scale $par(scl_avg)
fsave $g1 $par(name)-$par(a)-$par(e)-$par(dp12).fid

    }
}

```

```

}
#set outpt [open $par(name).out a+ 0600]
#puts $outpt "$dp12_opt $par(bestkaisqr)"
#close $outpt
#funload

}

```

2.8 Output file 2 spin (CN) SIMPSON File from 2WIL.pdb

```

SIMP
NP=7
SW=50000
TYPE=FID
DATA
0.00785343775 0
0.102004628 0
0.355299988 0
0.661681946 0
0.912493344 0
1.03832576 0
1.03547994 0
END

```

3. Sample Fortran Input Script Files For Global χ^2 Fittings

Reducing computational time was essentially for executing the global fittings where up to 144 nodes were simultaneously performing calculations. This work was not possible without using the Michigan State High Performance Computing Center. Additionally, I must thank Dirk Colbry for providing assistance in developing the script files below to reduce computational times. Please see **Appendix I** for the location of each file. Additionally, each file has an associated compiled output file that can be viewed in excel. Please see **Appendix I** for the location of the output file for each fitting. For each fitting, 2 input script files were used: (1) A “qsub file” used to split the main job across many nodes; (2) The main script fortran file. The name of each file below is denoted in quotations in the title of each section.

3.1 HFP 3 Registry Fitting qsub Script, “x2_fixed”

```

#!/bin/bash
#PBS -l nodes=1:ppn=1,walltime=144:00:00,mem=2gb,feature=gbe
#PBS -j oe

```

```

#PBS -t 0-109

#change to the original working directory
cd ${PBS_O_WORKDIR}

# Define number of loops for i16
cols=21
# Define number of loops for i22 (not used)
rows=4

# Number of jobs = cols*rows (i.e. -t 0-109)

#math to figure out the variable values based on the array id
i16=`echo "${PBS_ARRAYID} % ( ${cols} + 1 )" | bc`
i22=`echo "${PBS_ARRAYID} / ( ${cols} + 1 )" | bc`

#display the command we are going to run
echo "./x2 ${i16} ${i22} > ${i16}_${i22}.txt"

#run the command with the input variables
./x2 ${i16} ${i22} > ${i16}_${i22}.txt

# Calculate the runtiem for the job
qstat -f ${PBS_JOBID}

```

3.2 HFP 3 Registry Fitting Main Script, “HFP.f”

```

* This is a comment
* This program was written by Scott Schmick 030111
*
* t values (1 = a = 482, 2 = b = 402, etc.)
  real f7 f8 f9 f10 f11 f12 f13 f14 f15 f16
  real f17 f18 f19 f20 f21 f22 f23 f24 f25
  real sa8 sb8 sc8 sd8 se8 sg8 sh8
  real sa9 sb9 sc9 sd9 se9 sg9 sh9
  real sa10 sb10 sc10 sd10 se10 sg10 sh10
  real sa11 sb11 sc11 sd11 se11 sg11 sh11
  real sa12 sb12 sc12 sd12 se12 sg12 sh12
  real sa13 sb13 sc13 sd13 se13 sg13 sh13
  real sa14 sb14 sc14 sd14 se14 sg14 sh14
  real sa15 sb15 sc15 sd15 se15 sg15 sh15
  real sa16 sb16 sc16 sd16 se16 sg16 sh16
  real sa17 sb17 sc17 sd17 se17 sg17 sh17
  real sa18 sb18 sc18 sd18 se18 sg18 sh18
  real sa19 sb19 sc19 sd19 se19 sg19 sh19
  real sa20 sb20 sc20 sd20 se20 sg20 sh20
  real sa21 sb21 sc21 sd21 se21 sg21 sh21
  real sa22 sb22 sc22 sd22 se22 sg22 sh22
  real sa23 sb23 sc23 sd23 se23 sg23 sh23
  real sa24 sb24 sc24 sd24 se24 sg24 sh24

  real S0
  real xa8 xb8 xc8 xd8 xe8 xg8 xh8
  real xa9 xb9 xc9 xd9 xe9 xg9 xh9
  real xa10 xb10 xc10 xd10 xe10 xg10 xh10

```

*

```

real xa11 xb11 xc11 xd11 xe11 xg11 xh11
real xa12 xb12 xc12 xd12 xe12 xg12 xh12
real xa13 xb13 xc13 xd13 xe13 xg13 xh13
real xa14 xb14 xc14 xd14 xe14 xg14 xh14
real xa15 xb15 xc15 xd15 xe15 xg15 xh15
real xa16 xb16 xc16 xd16 xe16 xg16 xh16
real xa17 xb17 xc17 xd17 xe17 xg17 xh17
real xa18 xb18 xc18 xd18 xe18 xg18 xh18
real xa19 xb19 xc19 xd19 xe19 xg19 xh19
real xa20 xb20 xc20 xd20 xe20 xg20 xh20
real xa21 xb21 xc21 xd21 xe21 xg21 xh21
real xa22 xb22 xc22 xd22 xe22 xg22 xh22
real xa23 xb23 xc23 xd23 xe23 xg23 xh23
real xa24 xb24 xc24 xd24 xe24 xg24 xh24

real goff1 goff2 goff3 goff4 goff5 goff6 goff7
  real gon1 gon2 gon3 gon4 gon5 gon6 gon7
  real gnad1 gnad2 gnad3 gnad4 gnad5 gnad6 gnad7
real ft8t1
real ft8t2
real ft8t3
real ft8t4
real ft8t5
real ft8t6
real ft8t7
real ft8t1e
real ft8t2e
real ft8t3e
real ft8t4e
real ft8t5e
real ft8t6e
real ft8t7e
real ft9t1
real ft9t2
real ft9t3
real ft9t4
real ft9t5
real ft9t6
real ft9t7
real ft9t1e
real ft9t2e
real ft9t3e
real ft9t4e
real ft9t5e
real ft9t6e
real ft9t7e
real ft10t1
real ft10t2
real ft10t3
real ft10t4
real ft10t5
real ft10t6
real ft10t7
real ft10t1e
real ft10t2e
real ft10t3e
real ft10t4e

```

real	ft10t5e
real	ft10t6e
real	ft10t7e
real	ft11t1
real	ft11t2
real	ft11t3
real	ft11t4
real	ft11t5
real	ft11t6
real	ft11t7
real	ft11t1e
real	ft11t2e
real	ft11t3e
real	ft11t4e
real	ft11t5e
real	ft11t6e
real	ft11t7e
real	ft12t1
real	ft12t2
real	ft12t3
real	ft12t4
real	ft12t5
real	ft12t6
real	ft12t7
real	ft12t1e
real	ft12t2e
real	ft12t3e
real	ft12t4e
real	ft12t5e
real	ft12t6e
real	ft12t7e
real	ft13t1
real	ft13t2
real	ft13t3
real	ft13t4
real	ft13t5
real	ft13t6
real	ft13t7
real	ft13t1e
real	ft13t2e
real	ft13t3e
real	ft13t4e
real	ft13t5e
real	ft13t6e
real	ft13t7e
real	ft14t1
real	ft14t2
real	ft14t3
real	ft14t4
real	ft14t5
real	ft14t6
real	ft14t7
real	ft14t1e
real	ft14t2e
real	ft14t3e
real	ft14t4e
real	ft14t5e

real	ft14t6e
real	ft14t7e
real	ft15t1
real	ft15t2
real	ft15t3
real	ft15t4
real	ft15t5
real	ft15t6
real	ft15t7
real	ft15t1e
real	ft15t2e
real	ft15t3e
real	ft15t4e
real	ft15t5e
real	ft15t6e
real	ft15t7e
real	ft16t1
real	ft16t2
real	ft16t3
real	ft16t4
real	ft16t5
real	ft16t6
real	ft16t7
real	ft16t1e
real	ft16t2e
real	ft16t3e
real	ft16t4e
real	ft16t5e
real	ft16t6e
real	ft16t7e
real	ft17t1
real	ft17t2
real	ft17t3
real	ft17t4
real	ft17t5
real	ft17t6
real	ft17t7
real	ft17t1e
real	ft17t2e
real	ft17t3e
real	ft17t4e
real	ft17t5e
real	ft17t6e
real	ft17t7e
real	ft18t1
real	ft18t2
real	ft18t3
real	ft18t4
real	ft18t5
real	ft18t6
real	ft18t7
real	ft18t1e
real	ft18t2e
real	ft18t3e
real	ft18t4e
real	ft18t5e
real	ft18t6e

real	ft18t7e
real	ft19t1
real	ft19t2
real	ft19t3
real	ft19t4
real	ft19t5
real	ft19t6
real	ft19t7
real	ft19t1e
real	ft19t2e
real	ft19t3e
real	ft19t4e
real	ft19t5e
real	ft19t6e
real	ft19t7e
real	ft20t1
real	ft20t2
real	ft20t3
real	ft20t4
real	ft20t5
real	ft20t6
real	ft20t7
real	ft20t1e
real	ft20t2e
real	ft20t3e
real	ft20t4e
real	ft20t5e
real	ft20t6e
real	ft20t7e
real	ft21t1
real	ft21t2
real	ft21t3
real	ft21t4
real	ft21t5
real	ft21t6
real	ft21t7
real	ft21t1e
real	ft21t2e
real	ft21t3e
real	ft21t4e
real	ft21t5e
real	ft21t6e
real	ft21t7e
real	ft22t1
real	ft22t2
real	ft22t3
real	ft22t4
real	ft22t5
real	ft22t6
real	ft22t7
real	ft22t1e
real	ft22t2e
real	ft22t3e
real	ft22t4e
real	ft22t5e
real	ft22t6e
real	ft22t7e

```

real          ft23t1
real          ft23t2
real          ft23t3
real          ft23t4
real          ft23t5
real          ft23t6
real          ft23t7
real          ft23t1e
real          ft23t2e
real          ft23t3e
real          ft23t4e
real          ft23t5e
real          ft23t6e
real          ft23t7e
real          ft24t1
real          ft24t2
real          ft24t3
real          ft24t4
real          ft24t5
real          ft24t6
real          ft24t7
real          ft24t1e
real          ft24t2e
real          ft24t3e
real          ft24t4e
real          ft24t5e
real          ft24t6e
real          ft24t7e
real kaisq
real bestkaisq
real bf8
real bf9
real bf10
real bf11
real bf12
real bf13
real bf14
real bf15
real bf16
real bf17
real bf18
real bf19
real bf20
real bf21
real bf22
real bf23
real bf24

* DIRK

integer i16 i22
character inputarg*128

CALL getarg(1,inputarg)
read(inputarg,*) i16

CALL getarg(2,inputarg)
read(inputarg,*) i22

```

```

goff1 = 0.45011191
goff2 =      0.581357776
goff3 =      0.710329562
goff4 =      0.826254485
goff5 =      0.918588976
goff6 =      0.978470252
goff7 =      0.998385086
gon1  =      0.093897695
gon2  =      0.118554369
gon3  =      0.196398598
gon4  =      0.378629871
gon5  =      0.645290442
gon6  =      0.893821899
gon7  =      0.991709633
gnad1 =      0.281406262
gnad2 =      0.283440096
gnad3 =      0.298896559
gnad4 =      0.309714762
gnad5 =      0.320805841
gnad6 =      0.325732761
gnad7 =      0.342143946

```

```

ft8t1 =      0.05708
ft8t2 =      0.05172
ft8t3 =      0.03724
ft8t4 =      0.03827
ft8t5 =      0.02986
ft8t6 =      0.01733
ft8t7 =      0.00602
ft8t1e =      0.01557
ft8t2e =      0.01152
ft8t3e =      0.01107
ft8t4e =      0.01
ft8t5e =      0.01
ft8t6e =      0.01
ft8t7e =      0.01
ft9t1  =      0.06324
ft9t2  =      0.06789
ft9t3  =      0.04697
ft9t4  =      0.03341
ft9t5  =      0.03188
ft9t6  =      0.00943
ft9t7  =      0.01236
ft9t1e =      0.01814
ft9t2e =      0.01219
ft9t3e =      0.01063
ft9t4e =      0.01
ft9t5e =      0.01
ft9t6e =      0.01
ft9t7e =      0.01
ft10t1 =      0.1108
ft10t2 =      0.06207
ft10t3 =      0.03901
ft10t4 =      0.04643

```

ft10t5	=	0.03261
ft10t6	=	0.02239
ft10t7	=	0.01478
ft10t1e	=	0.02111
ft10t2e	=	0.01698
ft10t3e	=	0.01879
ft10t4e	=	0.01561
ft10t5e	=	0.01157
ft10t6e	=	0.01
ft10t7e	=	0.01
ft11t1	=	0.1414
ft11t2	=	0.09741
ft11t3	=	0.08076
ft11t4	=	0.06556
ft11t5	=	0.04604
ft11t6	=	0.02563
ft11t7	=	0.01426
ft11t1e	=	0.02165
ft11t2e	=	0.01693
ft11t3e	=	0.01136
ft11t4e	=	0.012
ft11t5e	=	0.01
ft11t6e	=	0.01
ft11t7e	=	0.01
ft12t1	=	0.21467
ft12t2	=	0.16973
ft12t3	=	0.1133
ft12t4	=	0.09493
ft12t5	=	0.05973
ft12t6	=	0.01603
ft12t7	=	0.01078
ft12t1e	=	0.023
ft12t2e	=	0.0164
ft12t3e	=	0.0113
ft12t4e	=	0.01236
ft12t5e	=	0.01
ft12t6e	=	0.01
ft12t7e	=	0.01
ft13t1	=	0.25555
ft13t2	=	0.21818
ft13t3	=	0.17189
ft13t4	=	0.10223
ft13t5	=	0.06738
ft13t6	=	0.03356
ft13t7	=	0.01049
ft13t1e	=	0.02456
ft13t2e	=	0.01572
ft13t3e	=	0.01381
ft13t4e	=	0.01466
ft13t5e	=	0.01151
ft13t6e	=	0.01
ft13t7e	=	0.01
ft14t1	=	0.23479
ft14t2	=	0.17063
ft14t3	=	0.13751
ft14t4	=	0.10895
ft14t5	=	0.08805

ft14t6	=	0.03327
ft14t7	=	0.00304
ft14t1e	=	0.02056
ft14t2e	=	0.01094
ft14t3e	=	0.01379
ft14t4e	=	0.01283
ft14t5e	=	0.012
ft14t6e	=	0.01
ft14t7e	=	0.01
ft15t1	=	0.24363
ft15t2	=	0.21509
ft15t3	=	0.17268
ft15t4	=	0.12293
ft15t5	=	0.09325
ft15t6	=	0.04324
ft15t7	=	0.00788
ft15t1e	=	0.01916
ft15t2e	=	0.01912
ft15t3e	=	0.01382
ft15t4e	=	0.01091
ft15t5e	=	0.01227
ft15t6e	=	0.01
ft15t7e	=	0.01
ft16t1	=	0.25327
ft16t2	=	0.23779
ft16t3	=	0.17889
ft16t4	=	0.12801
ft16t5	=	0.09009
ft16t6	=	0.0444
ft16t7	=	0.0119
ft16t1e	=	0.0148
ft16t2e	=	0.01449
ft16t3e	=	0.01081
ft16t4e	=	0.01
ft16t5e	=	0.01
ft16t6e	=	0.01
ft16t7e	=	0.01
ft17t1	=	0.27503
ft17t2	=	0.24656
ft17t3	=	0.19215
ft17t4	=	0.15509
ft17t5	=	0.0985
ft17t6	=	0.05809
ft17t7	=	0.00448
ft17t1e	=	0.02104
ft17t2e	=	0.01588
ft17t3e	=	0.01137
ft17t4e	=	0.01298
ft17t5e	=	0.01
ft17t6e	=	0.01
ft17t7e	=	0.01
ft18t1	=	0.20074
ft18t2	=	0.18812
ft18t3	=	0.17372
ft18t4	=	0.12562
ft18t5	=	0.08545
ft18t6	=	0.05519

ft18t7	=	0.01112
ft18t1e	=	0.0208
ft18t2e	=	0.02035
ft18t3e	=	0.01215
ft18t4e	=	0.01004
ft18t5e	=	0.01077
ft18t6e	=	0.01061
ft18t7e	=	0.01
ft19t1	=	0.15663
ft19t2	=	0.14513
ft19t3	=	0.13131
ft19t4	=	0.08177
ft19t5	=	0.06441
ft19t6	=	0.02196
ft19t7	=	0.01039
ft19t1e	=	0.01323
ft19t2e	=	0.01279
ft19t3e	=	0.01142
ft19t4e	=	0.01
ft19t5e	=	0.01
ft19t6e	=	0.01
ft19t7e	=	0.01
ft20t1	=	0.17509
ft20t2	=	0.17716
ft20t3	=	0.16072
ft20t4	=	0.11588
ft20t5	=	0.06816
ft20t6	=	0.01721
ft20t7	=	0.02198
ft20t1e	=	0.01534
ft20t2e	=	0.02447
ft20t3e	=	0.01242
ft20t4e	=	0.0147
ft20t5e	=	0.01
ft20t6e	=	0.01226
ft20t7e	=	0.01231
ft21t1	=	0.11176
ft21t2	=	0.07229
ft21t3	=	0.07415
ft21t4	=	0.05246
ft21t5	=	0.02774
ft21t6	=	0.00517
ft21t7	=	0.0098
ft21t1e	=	0.01604
ft21t2e	=	0.01604
ft21t3e	=	0.01327
ft21t4e	=	0.01275
ft21t5e	=	0.01133
ft21t6e	=	0.01032
ft21t7e	=	0.01
ft22t1	=	0.09585
ft22t2	=	0.08392
ft22t3	=	0.06998
ft22t4	=	0.02077
ft22t5	=	0.041
ft22t6	=	0.04159
ft22t7	=	0.01122

```

ft22t1e = 0.01307
ft22t2e = 0.01559
ft22t3e = 0.01072
ft22t4e = 0.01187
ft22t5e = 0.01
ft22t6e = 0.01
ft22t7e = 0.01
ft23t1 = 0.11271
ft23t2 = 0.08928
ft23t3 = 0.05727
ft23t4 = 0.05885
ft23t5 = 0.04872
ft23t6 = 0.0138
ft23t7 = 0.02577
ft23t1e = 0.01934
ft23t2e = 0.01941
ft23t3e = 0.01625
ft23t4e = 0.01762
ft23t5e = 0.01068
ft23t6e = 0.01402
ft23t7e = 0.01024
ft24t1 = 0.0458
ft24t2 = 0.04996
ft24t3 = 0.0152
ft24t4 = 0.02364
ft24t5 = 0.03095
ft24t6 = 0.01628
ft24t7 = 0.00624
ft24t1e = 0.01433
ft24t2e = 0.0138
ft24t3e = 0.01858
ft24t4e = 0.01
ft24t5e = 0.01
ft24t6e = 0.01
ft24t7e = 0.01
bestkaisq = 100000000

```

S0 = 1.32

Slab = 0.9625275

f7 = 0.00

f25 = 0.00

f8 = 0.00

do while (f8 <= 0.00)

f9 = 0.00

do while (f9 <= 0.01)

f10 = 0.00

sa8 = (S0-Slab*(gon1*f8+goff1*(f7+f9)+1-f7-f8-f9)-gnad1)/S0

sb8 = (S0-Slab*(gon2*f8+goff2*(f7+f9)+1-f7-f8-f9)-gnad2)/S0

sc8 = (S0-Slab*(gon3*f8+goff3*(f7+f9)+1-f7-f8-f9)-gnad3)/S0

sd8 = (S0-Slab*(gon4*f8+goff4*(f7+f9)+1-f7-f8-f9)-gnad4)/S0

se8 = (S0-Slab*(gon5*f8+goff5*(f7+f9)+1-f7-f8-f9)-gnad5)/S0

sg8 = (S0-Slab*(gon6*f8+goff6*(f7+f9)+1-f7-f8-f9)-gnad6)/S0

sh8 = (S0-Slab*(gon7*f8+goff7*(f7+f9)+1-f7-f8-f9)-gnad7)/S0


```

    xa8 = (ft8t1-sa8)**2/ft8t1e**2
    xb8 = (ft8t2-sb8)**2/ft8t2e**2
    xc8 = (ft8t3-sc8)**2/ft8t3e**2
    xd8 = (ft8t4-sd8)**2/ft8t4e**2
    xe8 = (ft8t5-se8)**2/ft8t5e**2
    xg8 = (ft8t6-sg8)**2/ft8t6e**2
    xh8 = (ft8t7-sh8)**2/ft8t7e**2
do while (f10 <= 0.02)
sa9 = (S0-Slab*(gon1*f9+goff1*(f8+f10)+1-f8-f9-f10)-gnad1)/S0
sb9 = (S0-Slab*(gon2*f9+goff2*(f8+f10)+1-f8-f9-f10)-gnad2)/S0
sc9 = (S0-Slab*(gon3*f9+goff3*(f8+f10)+1-f8-f9-f10)-gnad3)/S0
sd9 = (S0-Slab*(gon4*f9+goff4*(f8+f10)+1-f8-f9-f10)-gnad4)/S0
se9 = (S0-Slab*(gon5*f9+goff5*(f8+f10)+1-f8-f9-f10)-gnad5)/S0
sg9 = (S0-Slab*(gon6*f9+goff6*(f8+f10)+1-f8-f9-f10)-gnad6)/S0
sh9 = (S0-Slab*(gon7*f9+goff7*(f8+f10)+1-f8-f9-f10)-gnad7)/S0
    xa9 = (ft9t1-sa9)**2/ft9t1e**2
    xb9 = (ft9t2-sb9)**2/ft9t2e**2
    xc9 = (ft9t3-sc9)**2/ft9t3e**2
    xd9 = (ft9t4-sd9)**2/ft9t4e**2
    xe9 = (ft9t5-se9)**2/ft9t5e**2
    xg9 = (ft9t6-sg9)**2/ft9t6e**2
    xh9 = (ft9t7-sh9)**2/ft9t7e**2
    f11 = 0.01
do while (f11 <= 0.07)
sa10 = (S0-Slab*(gon1*f10+goff1*(f9+f11)+1-f9-f10-f11)-gnad1)/S0
sb10 = (S0-Slab*(gon2*f10+goff2*(f9+f11)+1-f9-f10-f11)-gnad2)/S0
sc10 = (S0-Slab*(gon3*f10+goff3*(f9+f11)+1-f9-f10-f11)-gnad3)/S0
sd10 = (S0-Slab*(gon4*f10+goff4*(f9+f11)+1-f9-f10-f11)-gnad4)/S0
se10 = (S0-Slab*(gon5*f10+goff5*(f9+f11)+1-f9-f10-f11)-gnad5)/S0
sg10 = (S0-Slab*(gon6*f10+goff6*(f9+f11)+1-f9-f10-f11)-gnad6)/S0
sh10 = (S0-Slab*(gon7*f10+goff7*(f9+f11)+1-f9-f10-f11)-gnad7)/S0
    xa10 = (ft10t1-sa10)**2/ft10t1e**2
    xb10 = (ft10t2-sb10)**2/ft10t2e**2
    xc10 = (ft10t3-sc10)**2/ft10t3e**2
    xd10 = (ft10t4-sd10)**2/ft10t4e**2
    xe10 = (ft10t5-se10)**2/ft10t5e**2
    xg10 = (ft10t6-sg10)**2/ft10t6e**2
    xh10 = (ft10t7-sh10)**2/ft10t7e**2
    f12 = 0.03
do while (f12 <= 0.15)
sa11 = (S0-Slab*(gon1*f11+goff1*(f10+f12)+1-f10-f11-f12)-gnad1)/S0
sb11 = (S0-Slab*(gon2*f11+goff2*(f10+f12)+1-f10-f11-f12)-gnad2)/S0
sc11 = (S0-Slab*(gon3*f11+goff3*(f10+f12)+1-f10-f11-f12)-gnad3)/S0
sd11 = (S0-Slab*(gon4*f11+goff4*(f10+f12)+1-f10-f11-f12)-gnad4)/S0
se11 = (S0-Slab*(gon5*f11+goff5*(f10+f12)+1-f10-f11-f12)-gnad5)/S0
sg11 = (S0-Slab*(gon6*f11+goff6*(f10+f12)+1-f10-f11-f12)-gnad6)/S0
sh11 = (S0-Slab*(gon7*f11+goff7*(f10+f12)+1-f10-f11-f12)-gnad7)/S0
    xa11 = (ft11t1-sa11)**2/ft11t1e**2
    xb11 = (ft11t2-sb11)**2/ft11t2e**2
    xc11 = (ft11t3-sc11)**2/ft11t3e**2
    xd11 = (ft11t4-sd11)**2/ft11t4e**2
    xe11 = (ft11t5-se11)**2/ft11t5e**2
    xg11 = (ft11t6-sg11)**2/ft11t6e**2
    xh11 = (ft11t7-sh11)**2/ft11t7e**2
    f13 = 0.08
do while (f13 <= 0.022)
sa12 = (S0-Slab*(gon1*f12+goff1*(f11+f13)+1-f11-f12-f13)-gnad1)/S0

```

```

sb12 = (S0-Slab*(gon2*f12+goff2*(f11+f13)+1-f11-f12-f13)-gnad2)/S0
sc12 = (S0-Slab*(gon3*f12+goff3*(f11+f13)+1-f11-f12-f13)-gnad3)/S0
sd12 = (S0-Slab*(gon4*f12+goff4*(f11+f13)+1-f11-f12-f13)-gnad4)/S0
se12 = (S0-Slab*(gon5*f12+goff5*(f11+f13)+1-f11-f12-f13)-gnad5)/S0
sg12 = (S0-Slab*(gon6*f12+goff6*(f11+f13)+1-f11-f12-f13)-gnad6)/S0
sh12 = (S0-Slab*(gon7*f12+goff7*(f11+f13)+1-f11-f12-f13)-gnad7)/S0
  xa12 = (ft12t1-sa12)**2/ft12t1e**2
  xb12 = (ft12t2-sb12)**2/ft12t2e**2
  xc12 = (ft12t3-sc12)**2/ft12t3e**2
  xd12 = (ft12t4-sd12)**2/ft12t4e**2
  xe12 = (ft12t5-se12)**2/ft12t5e**2
  xg12 = (ft12t6-sg12)**2/ft12t6e**2
  xh12 = (ft12t7-sh12)**2/ft12t7e**2
  f14 = 0.00
do while (f14 <= 0.19)
  sa13 = (S0-Slab*(gon1*f13+goff1*(f12+f14)+1-f12-f13-f14)-gnad1)/S0
  sb13 = (S0-Slab*(gon2*f13+goff2*(f12+f14)+1-f12-f13-f14)-gnad2)/S0
  sc13 = (S0-Slab*(gon3*f13+goff3*(f12+f14)+1-f12-f13-f14)-gnad3)/S0
  sd13 = (S0-Slab*(gon4*f13+goff4*(f12+f14)+1-f12-f13-f14)-gnad4)/S0
  se13 = (S0-Slab*(gon5*f13+goff5*(f12+f14)+1-f12-f13-f14)-gnad5)/S0
  sg13 = (S0-Slab*(gon6*f13+goff6*(f12+f14)+1-f12-f13-f14)-gnad6)/S0
  sh13 = (S0-Slab*(gon7*f13+goff7*(f12+f14)+1-f12-f13-f14)-gnad7)/S0
    xa13 = (ft13t1-sa13)**2/ft13t1e**2
    xb13 = (ft13t2-sb13)**2/ft13t2e**2
    xc13 = (ft13t3-sc13)**2/ft13t3e**2
    xd13 = (ft13t4-sd13)**2/ft13t4e**2
    xe13 = (ft13t5-se13)**2/ft13t5e**2
    xg13 = (ft13t6-sg13)**2/ft13t6e**2
    xh13 = (ft13t7-sh13)**2/ft13t7e**2
    f15 = 0.06
do while (f15 <= 0.23)
  sa14 = (S0-Slab*(gon1*f14+goff1*(f13+f15)+1-f13-f14-f15)-gnad1)/S0
  sb14 = (S0-Slab*(gon2*f14+goff2*(f13+f15)+1-f13-f14-f15)-gnad2)/S0
  sc14 = (S0-Slab*(gon3*f14+goff3*(f13+f15)+1-f13-f14-f15)-gnad3)/S0
  sd14 = (S0-Slab*(gon4*f14+goff4*(f13+f15)+1-f13-f14-f15)-gnad4)/S0
  se14 = (S0-Slab*(gon5*f14+goff5*(f13+f15)+1-f13-f14-f15)-gnad5)/S0
  sg14 = (S0-Slab*(gon6*f14+goff6*(f13+f15)+1-f13-f14-f15)-gnad6)/S0
  sh14 = (S0-Slab*(gon7*f14+goff7*(f13+f15)+1-f13-f14-f15)-gnad7)/S0
    xa14 = (ft14t1-sa14)**2/ft14t1e**2
    xb14 = (ft14t2-sb14)**2/ft14t2e**2
    xc14 = (ft14t3-sc14)**2/ft14t3e**2
    xd14 = (ft14t4-sd14)**2/ft14t4e**2
    xe14 = (ft14t5-se14)**2/ft14t5e**2
    xg14 = (ft14t6-sg14)**2/ft14t6e**2
    xh14 = (ft14t7-sh14)**2/ft14t7e**2
* DIRK
  f16 = 0.04+0.01*i16
*
  f16 = 0.00
*
  do while (f16 <= 0.03)
    sa15 = (S0-Slab*(gon1*f15+goff1*(f14+f16)+1-f14-f15-f16)-gnad1)/S0
    sb15 = (S0-Slab*(gon2*f15+goff2*(f14+f16)+1-f14-f15-f16)-gnad2)/S0
    sc15 = (S0-Slab*(gon3*f15+goff3*(f14+f16)+1-f14-f15-f16)-gnad3)/S0
    sd15 = (S0-Slab*(gon4*f15+goff4*(f14+f16)+1-f14-f15-f16)-gnad4)/S0
    se15 = (S0-Slab*(gon5*f15+goff5*(f14+f16)+1-f14-f15-f16)-gnad5)/S0
    sg15 = (S0-Slab*(gon6*f15+goff6*(f14+f16)+1-f14-f15-f16)-gnad6)/S0
    sh15 = (S0-Slab*(gon7*f15+goff7*(f14+f16)+1-f14-f15-f16)-gnad7)/S0
      xa15 = (ft15t1-sa15)**2/ft15t1e**2

```

```

xb15 = (ft15t2-sb15)**2/ft15t2e**2
xc15 = (ft15t3-sc15)**2/ft15t3e**2
xd15 = (ft15t4-sd15)**2/ft15t4e**2
xe15 = (ft15t5-se15)**2/ft15t5e**2
xg15 = (ft15t6-sg15)**2/ft15t6e**2
xh15 = (ft15t7-sh15)**2/ft15t7e**2
f17 = 0.10
do while (f17 <= 0.28)
sa16 = (S0-Slab*(gon1*f16+goff1*(f15+f17)+1-f15-f16-f17)-gnad1)/S0
sb16 = (S0-Slab*(gon2*f16+goff2*(f15+f17)+1-f15-f16-f17)-gnad2)/S0
sc16 = (S0-Slab*(gon3*f16+goff3*(f15+f17)+1-f15-f16-f17)-gnad3)/S0
sd16 = (S0-Slab*(gon4*f16+goff4*(f15+f17)+1-f15-f16-f17)-gnad4)/S0
se16 = (S0-Slab*(gon5*f16+goff5*(f15+f17)+1-f15-f16-f17)-gnad5)/S0
sg16 = (S0-Slab*(gon6*f16+goff6*(f15+f17)+1-f15-f16-f17)-gnad6)/S0
sh16 = (S0-Slab*(gon7*f16+goff7*(f15+f17)+1-f15-f16-f17)-gnad7)/S0
xa16 = (ft16t1-sa16)**2/ft16t1e**2
xb16 = (ft16t2-sb16)**2/ft16t2e**2
xc16 = (ft16t3-sc16)**2/ft16t3e**2
xd16 = (ft16t4-sd16)**2/ft16t4e**2
xe16 = (ft16t5-se16)**2/ft16t5e**2
xg16 = (ft16t6-sg16)**2/ft16t6e**2
xh16 = (ft16t7-sh16)**2/ft16t7e**2
f18 = 0.05
do while (f18 <= 0.21)
sa17 = (S0-Slab*(gon1*f17+goff1*(f16+f18)+1-f16-f17-f18)-gnad1)/S0
sb17 = (S0-Slab*(gon2*f17+goff2*(f16+f18)+1-f16-f17-f18)-gnad2)/S0
sc17 = (S0-Slab*(gon3*f17+goff3*(f16+f18)+1-f16-f17-f18)-gnad3)/S0
sd17 = (S0-Slab*(gon4*f17+goff4*(f16+f18)+1-f16-f17-f18)-gnad4)/S0
se17 = (S0-Slab*(gon5*f17+goff5*(f16+f18)+1-f16-f17-f18)-gnad5)/S0
sg17 = (S0-Slab*(gon6*f17+goff6*(f16+f18)+1-f16-f17-f18)-gnad6)/S0
sh17 = (S0-Slab*(gon7*f17+goff7*(f16+f18)+1-f16-f17-f18)-gnad7)/S0
xa17 = (ft17t1-sa17)**2/ft17t1e**2
xb17 = (ft17t2-sb17)**2/ft17t2e**2
xc17 = (ft17t3-sc17)**2/ft17t3e**2
xd17 = (ft17t4-sd17)**2/ft17t4e**2
xe17 = (ft17t5-se17)**2/ft17t5e**2
xg17 = (ft17t6-sg17)**2/ft17t6e**2
xh17 = (ft17t7-sh17)**2/ft17t7e**2
f19 = 0.00
do while (f19 <= 0.14)
sa18 = (S0-Slab*(gon1*f18+goff1*(f17+f19)+1-f17-f18-f19)-gnad1)/S0
sb18 = (S0-Slab*(gon2*f18+goff2*(f17+f19)+1-f17-f18-f19)-gnad2)/S0
sc18 = (S0-Slab*(gon3*f18+goff3*(f17+f19)+1-f17-f18-f19)-gnad3)/S0
sd18 = (S0-Slab*(gon4*f18+goff4*(f17+f19)+1-f17-f18-f19)-gnad4)/S0
se18 = (S0-Slab*(gon5*f18+goff5*(f17+f19)+1-f17-f18-f19)-gnad5)/S0
sg18 = (S0-Slab*(gon6*f18+goff6*(f17+f19)+1-f17-f18-f19)-gnad6)/S0
sh18 = (S0-Slab*(gon7*f18+goff7*(f17+f19)+1-f17-f18-f19)-gnad7)/S0
xa18 = (ft18t1-sa18)**2/ft18t1e**2
xb18 = (ft18t2-sb18)**2/ft18t2e**2
xc18 = (ft18t3-sc18)**2/ft18t3e**2
xd18 = (ft18t4-sd18)**2/ft18t4e**2
xe18 = (ft18t5-se18)**2/ft18t5e**2
xg18 = (ft18t6-sg18)**2/ft18t6e**2
xh18 = (ft18t7-sh18)**2/ft18t7e**2
f20 = 0.11
do while (f20 <= 0.18)
sa19 = (S0-Slab*(gon1*f19+goff1*(f18+f20)+1-f18-f19-f20)-gnad1)/S0

```

```

sb19 = (S0-Slab*(gon2*f19+goff2*(f18+f20)+1-f18-f19-f20)-gnad2)/S0
sc19 = (S0-Slab*(gon3*f19+goff3*(f18+f20)+1-f18-f19-f20)-gnad3)/S0
sd19 = (S0-Slab*(gon4*f19+goff4*(f18+f20)+1-f18-f19-f20)-gnad4)/S0
se19 = (S0-Slab*(gon5*f19+goff5*(f18+f20)+1-f18-f19-f20)-gnad5)/S0
sg19 = (S0-Slab*(gon6*f19+goff6*(f18+f20)+1-f18-f19-f20)-gnad6)/S0
sh19 = (S0-Slab*(gon7*f19+goff7*(f18+f20)+1-f18-f19-f20)-gnad7)/S0
  xa19 = (ft19t1-sa19)**2/ft19t1e**2
  xb19 = (ft19t2-sb19)**2/ft19t2e**2
  xc19 = (ft19t3-sc19)**2/ft19t3e**2
  xd19 = (ft19t4-sd19)**2/ft19t4e**2
  xe19 = (ft19t5-se19)**2/ft19t5e**2
  xg19 = (ft19t6-sg19)**2/ft19t6e**2
  xh19 = (ft19t7-sh19)**2/ft19t7e**2
  f21 = 0.00
do while (f21 <= 0.04)
sa20 = (S0-Slab*(gon1*f20+goff1*(f19+f21)+1-f19-f20-f21)-gnad1)/S0
sb20 = (S0-Slab*(gon2*f20+goff2*(f19+f21)+1-f19-f20-f21)-gnad2)/S0
sc20 = (S0-Slab*(gon3*f20+goff3*(f19+f21)+1-f19-f20-f21)-gnad3)/S0
sd20 = (S0-Slab*(gon4*f20+goff4*(f19+f21)+1-f19-f20-f21)-gnad4)/S0
se20 = (S0-Slab*(gon5*f20+goff5*(f19+f21)+1-f19-f20-f21)-gnad5)/S0
sg20 = (S0-Slab*(gon6*f20+goff6*(f19+f21)+1-f19-f20-f21)-gnad6)/S0
sh20 = (S0-Slab*(gon7*f20+goff7*(f19+f21)+1-f19-f20-f21)-gnad7)/S0
  xa20 = (ft20t1-sa20)**2/ft20t1e**2
  xb20 = (ft20t2-sb20)**2/ft20t2e**2
  xc20 = (ft20t3-sc20)**2/ft20t3e**2
  xd20 = (ft20t4-sd20)**2/ft20t4e**2
  xe20 = (ft20t5-se20)**2/ft20t5e**2
  xg20 = (ft20t6-sg20)**2/ft20t6e**2
  xh20 = (ft20t7-sh20)**2/ft20t7e**2
* DIRK
  f22 = 0.00+0.01*i22
*   f22 = 0.00
*   do while (f22 <= 0.03)
sa21 = (S0-Slab*(gon1*f21+goff1*(f20+f22)+1-f20-f21-f22)-gnad1)/S0
sb21 = (S0-Slab*(gon2*f21+goff2*(f20+f22)+1-f20-f21-f22)-gnad2)/S0
sc21 = (S0-Slab*(gon3*f21+goff3*(f20+f22)+1-f20-f21-f22)-gnad3)/S0
sd21 = (S0-Slab*(gon4*f21+goff4*(f20+f22)+1-f20-f21-f22)-gnad4)/S0
se21 = (S0-Slab*(gon5*f21+goff5*(f20+f22)+1-f20-f21-f22)-gnad5)/S0
sg21 = (S0-Slab*(gon6*f21+goff6*(f20+f22)+1-f20-f21-f22)-gnad6)/S0
sh21 = (S0-Slab*(gon7*f21+goff7*(f20+f22)+1-f20-f21-f22)-gnad7)/S0
  xa21 = (ft21t1-sa21)**2/ft21t1e**2
  xb21 = (ft21t2-sb21)**2/ft21t2e**2
  xc21 = (ft21t3-sc21)**2/ft21t3e**2
  xd21 = (ft21t4-sd21)**2/ft21t4e**2
  xe21 = (ft21t5-se21)**2/ft21t5e**2
  xg21 = (ft21t6-sg21)**2/ft21t6e**2
  xh21 = (ft21t7-sh21)**2/ft21t7e**2
  f23 = 0.04
do while (f23 <= 0.05)
sa22 = (S0-Slab*(gon1*f22+goff1*(f21+f23)+1-f21-f22-f23)-gnad1)/S0
sb22 = (S0-Slab*(gon2*f22+goff2*(f21+f23)+1-f21-f22-f23)-gnad2)/S0
sc22 = (S0-Slab*(gon3*f22+goff3*(f21+f23)+1-f21-f22-f23)-gnad3)/S0
sd22 = (S0-Slab*(gon4*f22+goff4*(f21+f23)+1-f21-f22-f23)-gnad4)/S0
se22 = (S0-Slab*(gon5*f22+goff5*(f21+f23)+1-f21-f22-f23)-gnad5)/S0
sg22 = (S0-Slab*(gon6*f22+goff6*(f21+f23)+1-f21-f22-f23)-gnad6)/S0
sh22 = (S0-Slab*(gon7*f22+goff7*(f21+f23)+1-f21-f22-f23)-gnad7)/S0
  xa22 = (ft22t1-sa22)**2/ft22t1e**2

```

```

xb22 = (ft22t2-sb22)**2/ft22t2e**2
xc22 = (ft22t3-sc22)**2/ft22t3e**2
xd22 = (ft22t4-sd22)**2/ft22t4e**2
xe22 = (ft22t5-se22)**2/ft22t5e**2
xg22 = (ft22t6-sg22)**2/ft22t6e**2
xh22 = (ft22t7-sh22)**2/ft22t7e**2
f24 = 0.00
do while (f24 <= 0.00)
sa23 = (S0-Slab*(gon1*f23+goff1*(f22+f24)+1-f22-f23-f24)-gnad1)/S0
sb23 = (S0-Slab*(gon2*f23+goff2*(f22+f24)+1-f22-f23-f24)-gnad2)/S0
sc23 = (S0-Slab*(gon3*f23+goff3*(f22+f24)+1-f22-f23-f24)-gnad3)/S0
sd23 = (S0-Slab*(gon4*f23+goff4*(f22+f24)+1-f22-f23-f24)-gnad4)/S0
se23 = (S0-Slab*(gon5*f23+goff5*(f22+f24)+1-f22-f23-f24)-gnad5)/S0
sg23 = (S0-Slab*(gon6*f23+goff6*(f22+f24)+1-f22-f23-f24)-gnad6)/S0
sh23 = (S0-Slab*(gon7*f23+goff7*(f22+f24)+1-f22-f23-f24)-gnad7)/S0

sa24 = (S0-Slab*(gon1*f24+goff1*(f23+f25)+1-f23-f24-f25)-gnad1)/S0
sb24 = (S0-Slab*(gon2*f24+goff2*(f23+f25)+1-f23-f24-f25)-gnad2)/S0
sc24 = (S0-Slab*(gon3*f24+goff3*(f23+f25)+1-f23-f24-f25)-gnad3)/S0
sd24 = (S0-Slab*(gon4*f24+goff4*(f23+f25)+1-f23-f24-f25)-gnad4)/S0
se24 = (S0-Slab*(gon5*f24+goff5*(f23+f25)+1-f23-f24-f25)-gnad5)/S0
sg24 = (S0-Slab*(gon6*f24+goff6*(f23+f25)+1-f23-f24-f25)-gnad6)/S0
sh24 = (S0-Slab*(gon7*f24+goff7*(f23+f25)+1-f23-f24-f25)-gnad7)/S0

xa23 = (ft23t1-sa23)**2/ft23t1e**2
xb23 = (ft23t2-sb23)**2/ft23t2e**2
xc23 = (ft23t3-sc23)**2/ft23t3e**2
xd23 = (ft23t4-sd23)**2/ft23t4e**2
xe23 = (ft23t5-se23)**2/ft23t5e**2
xg23 = (ft23t6-sg23)**2/ft23t6e**2
xh23 = (ft23t7-sh23)**2/ft23t7e**2

xa24 = (ft24t1-sa24)**2/ft24t1e**2
xb24 = (ft24t2-sb24)**2/ft24t2e**2
xc24 = (ft24t3-sc24)**2/ft24t3e**2
xd24 = (ft24t4-sd24)**2/ft24t4e**2
xe24 = (ft24t5-se24)**2/ft24t5e**2
xg24 = (ft24t6-sg24)**2/ft24t6e**2
xh24 = (ft24t7-sh24)**2/ft24t7e**2

kaisq = xa8+xb8+xc8+xd8+xe8+xg8+xh8+
& xa9+xb9+xc9+xd9+xe9+xg9+xh9+
& xa10+xb10+xc10+xd10+xe10+xg10+xh10+
& xa11+xb11+xc11+xd11+xe11+xg11+xh11+
& xa12+xb12+xc12+xd12+xe12+xg12+xh12+
& xa13+xb13+xc13+xd13+xe13+xg13+xh13+
& xa14+xb14+xc14+xd14+xe14+xg14+xh14+
& xa15+xb15+xc15+xd15+xe15+xg15+xh15+
& xa16+xb16+xc16+xd16+xe16+xg16+xh16+
& xa17+xb17+xc17+xd17+xe17+xg17+xh17+
& xa18+xb18+xc18+xd18+xe18+xg18+xh18+
& xa19+xb19+xc19+xd19+xe19+xg19+xh19+
& xa20+xb20+xc20+xd20+xe20+xg20+xh20+
& xa21+xb21+xc21+xd21+xe21+xg21+xh21+
& xa22+xb22+xc22+xd22+xe22+xg22+xh22+
& xa23+xb23+xc23+xd23+xe23+xg23+xh23+
& xa24+xb24+xc24+xd24+xe24+xg24+xh24

```

```

IF (kaisq < bestkaisq) THEN
    bestkaisq = kaisq
    bf8 = f8
    bf9 = f9
    bf10 = f10
    bf11 = f11
    bf12 = f12
    bf13 = f13
    bf14 = f14
    bf15 = f15
    bf16 = f16
    bf17 = f17
    bf18 = f18
    bf19 = f19
    bf20 = f20
    bf21 = f21
    bf22 = f22
    bf23 = f23
    bf24 = f24

endif

    f24 = f24 + 0.01
enddo
    f23 = f23 + 0.01
enddo

* DIRK
*
*
    f22 = f22 + 0.01
enddo
    f21 = f21 + 0.01
enddo
    f20 = f20 + 0.01
enddo
    f19 = f19 + 0.01
enddo
    f18 = f18 + 0.01
enddo
    f17 = f17 + 0.01
enddo

* DIRK
*
*
    f16 = f16 + 0.01
enddo
    f15 = f15 + 0.01
enddo
    f14 = f14 + 0.01
enddo
    f13 = f13 + 0.01
enddo
    f12 = f12 + 0.01
enddo
    f11 = f11 + 0.01
enddo
    f10 = f10 + 0.01
enddo

```

```

        f9 = f9 + 0.01
    enddo
        f8 = f8 + 0.01

    enddo

* DIRK
*      OPEN(UNIT = 12, FILE = 'output', STATUS = 'NEW')
*      WRITE(12,*) bestkaisq,
*      WRITE(*,*) bestkaisq,
*      &          bf8,
*      &          bf9,
*      &          bf10, bf11, bf12, bf13, bf14, bf15,
*      &          bf16, bf17, bf18, bf19, bf20, bf21, bf22,
*      &          bf23, bf24
*      end

```

3.3 HFP 5 Registry Fitting qsub Script, “x2_HFP”

For HFP 5 registry fittings, two main script files were created (see **Appendix I**). In the file below, $f_9 = 0.00$. The second main script file used $f_9 = 0.01$. Splitting the main script file into two separate jobs was required to complete each script file’s computations in less than 168 hours (i.e. the maximum time allowed to occupy a node). This time limit is set by the High Performance Computing Center and jobs that run longer than 168 hours are automatically terminated.

```

#!/bin/bash
#PBS -l nodes=1:ppn=1,walltime=168:00:00,mem=2gb,feature=gbe
#PBS -j oe
#PBS -t 0-139

#change to the original working directory
cd ${PBS_O_WORKDIR}

# Define number of loops for i12
cols=13
# Define number of loops for i20 (not used)
rows=9

# Number of jobs = cols*rows (i.e. -t 0-139)

#math to figure out the variable values based on the array id
i12=`echo "${PBS_ARRAYID} % ( ${cols} + 1 )" | bc`
i20=`echo "${PBS_ARRAYID} / ( ${cols} + 1 )" | bc`

#display the command we are going to run
echo "./x2 ${i12} ${i20} > ${i12}_${i20}.txt"

```

```
#run the command with the input variables
./x2 ${i12} ${i20} > ${i12}_${i20}.txt
```

```
# Calculate the runtiem for the job
qstat -f ${PBS_JOBID}
```

3.4 HFP 5 Registry Fitting Main Script, “HFP_5var.f”

```
* This is a comment
* This program was written by Scott Schmick 030111
*
* t values (1 = a = 482, 2 = b = 402, etc.)
  real f7 f8 f9 f10 f11 f12 f13 f14 f15 f16
  real f17 f18 f19 f20 f21 f22 f23 f24 f25
  real sa8 sb8 sc8 sd8 se8 sg8 sh8
  real sa9 sb9 sc9 sd9 se9 sg9 sh9
  real sa10 sb10 sc10 sd10 se10 sg10 sh10
  real sa11 sb11 sc11 sd11 se11 sg11 sh11
  real sa12 sb12 sc12 sd12 se12 sg12 sh12
  real sa13 sb13 sc13 sd13 se13 sg13 sh13
  real sa14 sb14 sc14 sd14 se14 sg14 sh14
  real sa15 sb15 sc15 sd15 se15 sg15 sh15
  real sa16 sb16 sc16 sd16 se16 sg16 sh16
  real sa17 sb17 sc17 sd17 se17 sg17 sh17
  real sa18 sb18 sc18 sd18 se18 sg18 sh18
  real sa19 sb19 sc19 sd19 se19 sg19 sh19
  real sa20 sb20 sc20 sd20 se20 sg20 sh20
  real sa21 sb21 sc21 sd21 se21 sg21 sh21
  real sa22 sb22 sc22 sd22 se22 sg22 sh22
  real sa23 sb23 sc23 sd23 se23 sg23 sh23
  real sa24 sb24 sc24 sd24 se24 sg24 sh24

  real S0
  real xa8 xb8 xc8 xd8 xe8 xg8 xh8
  real xa9 xb9 xc9 xd9 xe9 xg9 xh9
  real xa10 xb10 xc10 xd10 xe10 xg10 xh10
  real xa11 xb11 xc11 xd11 xe11 xg11 xh11
  real xa12 xb12 xc12 xd12 xe12 xg12 xh12
  real xa13 xb13 xc13 xd13 xe13 xg13 xh13
  real xa14 xb14 xc14 xd14 xe14 xg14 xh14
  real xa15 xb15 xc15 xd15 xe15 xg15 xh15
  real xa16 xb16 xc16 xd16 xe16 xg16 xh16
  real xa17 xb17 xc17 xd17 xe17 xg17 xh17
  real xa18 xb18 xc18 xd18 xe18 xg18 xh18
  real xa19 xb19 xc19 xd19 xe19 xg19 xh19
  real xa20 xb20 xc20 xd20 xe20 xg20 xh20
  real xa21 xb21 xc21 xd21 xe21 xg21 xh21
  real xa22 xb22 xc22 xd22 xe22 xg22 xh22
  real xa23 xb23 xc23 xd23 xe23 xg23 xh23
  real xa24 xb24 xc24 xd24 xe24 xg24 xh24

  real goff1 goff2 goff3 goff4 goff5 goff6 goff7
  real gon1 gon2 gon3 gon4 gon5 gon6 gon7
  real gnad1 gnad2 gnad3 gnad4 gnad5 gnad6 gnad7
  real ft8t1
```


real	ft8t2	
real	ft8t3	
real	ft8t4	
real	ft8t5	
real	ft8t6	
real	ft8t7	
real	ft8t1e	
real	ft8t2e	
real	ft8t3e	
real	ft8t4e	
real	ft8t5e	
real	ft8t6e	
real	ft8t7e	
real	ft9t1	
real	ft9t2	
real	ft9t3	
real	ft9t4	
real	ft9t5	
real	ft9t6	
real	ft9t7	
real	ft9t1e	
real	ft9t2e	
real	ft9t3e	
real	ft9t4e	
real	ft9t5e	
real	ft9t6e	
real	ft9t7e	
real	ft10t1	
real	ft10t2	
real	ft10t3	
real	ft10t4	
real	ft10t5	
real	ft10t6	
real	ft10t7	
real	ft10t1e	
real	ft10t2e	
real	ft10t3e	
real	ft10t4e	
real	ft10t5e	
real	ft10t6e	
real	ft10t7e	
real		ft11t1
real		ft11t2
real		ft11t3
real		ft11t4
real		ft11t5
real		ft11t6
real		ft11t7
real		ft11t1e
real		ft11t2e
real		ft11t3e
real		ft11t4e
real		ft11t5e
real		ft11t6e
real		ft11t7e
real		ft12t1
real		ft12t2

real	ft12t3
real	ft12t4
real	ft12t5
real	ft12t6
real	ft12t7
real	ft12t1e
real	ft12t2e
real	ft12t3e
real	ft12t4e
real	ft12t5e
real	ft12t6e
real	ft12t7e
real	ft13t1
real	ft13t2
real	ft13t3
real	ft13t4
real	ft13t5
real	ft13t6
real	ft13t7
real	ft13t1e
real	ft13t2e
real	ft13t3e
real	ft13t4e
real	ft13t5e
real	ft13t6e
real	ft13t7e
real	ft14t1
real	ft14t2
real	ft14t3
real	ft14t4
real	ft14t5
real	ft14t6
real	ft14t7
real	ft14t1e
real	ft14t2e
real	ft14t3e
real	ft14t4e
real	ft14t5e
real	ft14t6e
real	ft14t7e
real	ft15t1
real	ft15t2
real	ft15t3
real	ft15t4
real	ft15t5
real	ft15t6
real	ft15t7
real	ft15t1e
real	ft15t2e
real	ft15t3e
real	ft15t4e
real	ft15t5e
real	ft15t6e
real	ft15t7e
real	ft16t1
real	ft16t2
real	ft16t3

real	ft16t4
real	ft16t5
real	ft16t6
real	ft16t7
real	ft16t1e
real	ft16t2e
real	ft16t3e
real	ft16t4e
real	ft16t5e
real	ft16t6e
real	ft16t7e
real	ft17t1
real	ft17t2
real	ft17t3
real	ft17t4
real	ft17t5
real	ft17t6
real	ft17t7
real	ft17t1e
real	ft17t2e
real	ft17t3e
real	ft17t4e
real	ft17t5e
real	ft17t6e
real	ft17t7e
real	ft18t1
real	ft18t2
real	ft18t3
real	ft18t4
real	ft18t5
real	ft18t6
real	ft18t7
real	ft18t1e
real	ft18t2e
real	ft18t3e
real	ft18t4e
real	ft18t5e
real	ft18t6e
real	ft18t7e
real	ft19t1
real	ft19t2
real	ft19t3
real	ft19t4
real	ft19t5
real	ft19t6
real	ft19t7
real	ft19t1e
real	ft19t2e
real	ft19t3e
real	ft19t4e
real	ft19t5e
real	ft19t6e
real	ft19t7e
real	ft20t1
real	ft20t2
real	ft20t3
real	ft20t4

real	ft20t5
real	ft20t6
real	ft20t7
real	ft20t1e
real	ft20t2e
real	ft20t3e
real	ft20t4e
real	ft20t5e
real	ft20t6e
real	ft20t7e
real	ft21t1
real	ft21t2
real	ft21t3
real	ft21t4
real	ft21t5
real	ft21t6
real	ft21t7
real	ft21t1e
real	ft21t2e
real	ft21t3e
real	ft21t4e
real	ft21t5e
real	ft21t6e
real	ft21t7e
real	ft22t1
real	ft22t2
real	ft22t3
real	ft22t4
real	ft22t5
real	ft22t6
real	ft22t7
real	ft22t1e
real	ft22t2e
real	ft22t3e
real	ft22t4e
real	ft22t5e
real	ft22t6e
real	ft22t7e
real	ft23t1
real	ft23t2
real	ft23t3
real	ft23t4
real	ft23t5
real	ft23t6
real	ft23t7
real	ft23t1e
real	ft23t2e
real	ft23t3e
real	ft23t4e
real	ft23t5e
real	ft23t6e
real	ft23t7e
real	ft24t1
real	ft24t2
real	ft24t3
real	ft24t4
real	ft24t5

```

real          ft24t6
real          ft24t7
real          ft24t1e
real          ft24t2e
real          ft24t3e
real          ft24t4e
real          ft24t5e
real          ft24t6e
real          ft24t7e
real kaisq
real bestkaisq
real bf8
real bf9
real bf10
real bf11
real bf12
real bf13
real bf14
real bf15
real bf16
real bf17
real bf18
real bf19
real bf20
real bf21
real bf22
real bf23
real bf24
real a8 aa8 a9 aa9 a10 aa10 a11 aa11 a12 aa12 a13 aa13 a14 aa14
real b8 b9 b10 b11 b12 b13 b14 b15 b16 b17 b18 b19
real b20 b21 b22 b23 b24
real z1 z2 z3 z4 z5 z6 z7 f6 f26
real a15 aa15 a16 aa16 a17 aa17 a18 aa18 a19 aa19 a20 aa20
real a21 aa21 a22 aa22 a23 aa23 a24 aa24

* DIRK

integer i12 i20
character inputarg*128

CALL getarg(1,inputarg)
read(inputarg,*) i12

CALL getarg(2,inputarg)
read(inputarg,*) i20

goff1 = 0.45011191
goff2 = 0.581357776
goff3 = 0.710329562
goff4 = 0.826254485
goff5 = 0.918588976
goff6 = 0.978470252
goff7 = 0.998385086
gon1  = 0.093897695
gon2  = 0.118554369
gon3  = 0.196398598
gon4  = 0.378629871
gon5  = 0.645290442

```

```

gon6    =      0.893821899
gon7    =      0.991709633
z1 = 0.9052388602
z2 = 0.9334319587
z3 = 0.95693731265
z4 = 0.97551074985
z5 = 0.98895970015
z6 = 0.99714805251
z7 = 0.9997874594365
gnad1   =      0.281406262
gnad2   =      0.283440096
gnad3   =      0.298896559
gnad4   =      0.309714762
gnad5   =      0.320805841
gnad6   =      0.325732761
gnad7   =      0.342143946
ft8t1   =      0.05708
ft8t2   =      0.05172
ft8t3   =      0.03724
ft8t4   =      0.03827
ft8t5   =      0.02986
ft8t6   =      0.01733
ft8t7   =      0.00602
ft8t1e  =      0.01557
ft8t2e  =      0.01152
ft8t3e  =      0.01107
ft8t4e  =      0.01
ft8t5e  =      0.01
ft8t6e  =      0.01
ft8t7e  =      0.01
ft9t1   =      0.06324
ft9t2   =      0.06789
ft9t3   =      0.04697
ft9t4   =      0.03341
ft9t5   =      0.03188
ft9t6   =      0.00943
ft9t7   =      0.01236
ft9t1e  =      0.01814
ft9t2e  =      0.01219
ft9t3e  =      0.01063
ft9t4e  =      0.01
ft9t5e  =      0.01
ft9t6e  =      0.01
ft9t7e  =      0.01
ft10t1  =      0.1108
ft10t2  =      0.06207
ft10t3  =      0.03901
ft10t4  =      0.04643
ft10t5  =      0.03261
ft10t6  =      0.02239
ft10t7  =      0.01478
ft10t1e =      0.02111
ft10t2e =      0.01698
ft10t3e =      0.01879
ft10t4e =      0.01561
ft10t5e =      0.01157
ft10t6e =      0.01

```

ft10t7e =	0.01
ft11t1 =	0.1414
ft11t2 =	0.09741
ft11t3 =	0.08076
ft11t4 =	0.06556
ft11t5 =	0.04604
ft11t6 =	0.02563
ft11t7 =	0.01426
ft11t1e =	0.02165
ft11t2e =	0.01693
ft11t3e =	0.01136
ft11t4e =	0.012
ft11t5e =	0.01
ft11t6e =	0.01
ft11t7e =	0.01
ft12t1 =	0.21467
ft12t2 =	0.16973
ft12t3 =	0.1133
ft12t4 =	0.09493
ft12t5 =	0.05973
ft12t6 =	0.01603
ft12t7 =	0.01078
ft12t1e =	0.023
ft12t2e =	0.0164
ft12t3e =	0.0113
ft12t4e =	0.01236
ft12t5e =	0.01
ft12t6e =	0.01
ft12t7e =	0.01
ft13t1 =	0.25555
ft13t2 =	0.21818
ft13t3 =	0.17189
ft13t4 =	0.10223
ft13t5 =	0.06738
ft13t6 =	0.03356
ft13t7 =	0.01049
ft13t1e =	0.02456
ft13t2e =	0.01572
ft13t3e =	0.01381
ft13t4e =	0.01466
ft13t5e =	0.01151
ft13t6e =	0.01
ft13t7e =	0.01
ft14t1 =	0.23479
ft14t2 =	0.17063
ft14t3 =	0.13751
ft14t4 =	0.10895
ft14t5 =	0.08805
ft14t6 =	0.03327
ft14t7 =	0.00304
ft14t1e =	0.02056
ft14t2e =	0.01094
ft14t3e =	0.01379
ft14t4e =	0.01283
ft14t5e =	0.012
ft14t6e =	0.01
ft14t7e =	0.01

```

ft15t1 = 0.24363
ft15t2 = 0.21509
ft15t3 = 0.17268
ft15t4 = 0.12293
ft15t5 = 0.09325
ft15t6 = 0.04324
ft15t7 = 0.00788
ft15t1e = 0.01916
ft15t2e = 0.01912
ft15t3e = 0.01382
ft15t4e = 0.01091
ft15t5e = 0.01227
ft15t6e = 0.01
ft15t7e = 0.01
ft16t1 = 0.25327
ft16t2 = 0.23779
ft16t3 = 0.17889
ft16t4 = 0.12801
ft16t5 = 0.09009
ft16t6 = 0.0444
ft16t7 = 0.0119
ft16t1e = 0.0148
ft16t2e = 0.01449
ft16t3e = 0.01081
ft16t4e = 0.01
ft16t5e = 0.01
ft16t6e = 0.01
ft16t7e = 0.01
ft17t1 = 0.27503
ft17t2 = 0.24656
ft17t3 = 0.19215
ft17t4 = 0.15509
ft17t5 = 0.0985
ft17t6 = 0.05809
ft17t7 = 0.00448
ft17t1e = 0.02104
ft17t2e = 0.01588
ft17t3e = 0.01137
ft17t4e = 0.01298
ft17t5e = 0.01
ft17t6e = 0.01
ft17t7e = 0.01
ft18t1 = 0.20074
ft18t2 = 0.18812
ft18t3 = 0.17372
ft18t4 = 0.12562
ft18t5 = 0.08545
ft18t6 = 0.05519
ft18t7 = 0.01112
ft18t1e = 0.0208
ft18t2e = 0.02035
ft18t3e = 0.01215
ft18t4e = 0.01004
ft18t5e = 0.01077
ft18t6e = 0.01061
ft18t7e = 0.01
ft19t1 = 0.15663

```


ft19t2 =	0.14513
ft19t3 =	0.13131
ft19t4 =	0.08177
ft19t5 =	0.06441
ft19t6 =	0.02196
ft19t7 =	0.01039
ft19t1e =	0.01323
ft19t2e =	0.01279
ft19t3e =	0.01142
ft19t4e =	0.01
ft19t5e =	0.01
ft19t6e =	0.01
ft19t7e =	0.01
ft20t1 =	0.17509
ft20t2 =	0.17716
ft20t3 =	0.16072
ft20t4 =	0.11588
ft20t5 =	0.06816
ft20t6 =	0.01721
ft20t7 =	0.02198
ft20t1e =	0.01534
ft20t2e =	0.02447
ft20t3e =	0.01242
ft20t4e =	0.0147
ft20t5e =	0.01
ft20t6e =	0.01226
ft20t7e =	0.01231
ft21t1 =	0.11176
ft21t2 =	0.07229
ft21t3 =	0.07415
ft21t4 =	0.05246
ft21t5 =	0.02774
ft21t6 =	0.00517
ft21t7 =	0.0098
ft21t1e =	0.01604
ft21t2e =	0.01604
ft21t3e =	0.01327
ft21t4e =	0.01275
ft21t5e =	0.01133
ft21t6e =	0.01032
ft21t7e =	0.01
ft22t1 =	0.09585
ft22t2 =	0.08392
ft22t3 =	0.06998
ft22t4 =	0.02077
ft22t5 =	0.041
ft22t6 =	0.04159
ft22t7 =	0.01122
ft22t1e =	0.01307
ft22t2e =	0.01559
ft22t3e =	0.01072
ft22t4e =	0.01187
ft22t5e =	0.01
ft22t6e =	0.01
ft22t7e =	0.01
ft23t1 =	0.11271
ft23t2 =	0.08928

```

ft23t3 =      0.05727
ft23t4 =      0.05885
ft23t5 =      0.04872
ft23t6 =      0.0138
ft23t7 =      0.02577
ft23t1e =      0.01934
ft23t2e =      0.01941
ft23t3e =      0.01625
ft23t4e =      0.01762
ft23t5e =      0.01068
ft23t6e =      0.01402
ft23t7e =      0.01024
ft24t1 =      0.0458
ft24t2 =      0.04996
ft24t3 =      0.0152
ft24t4 =      0.02364
ft24t5 =      0.03095
ft24t6 =      0.01628
ft24t7 =      0.00624
ft24t1e =      0.01433
ft24t2e =      0.0138
ft24t3e =      0.01858
ft24t4e =      0.01
ft24t5e =      0.01
ft24t6e =      0.01
ft24t7e =      0.01

```

```
bestkaisq = 100000000
```

S0 = 1.32

Slab = 0.9625275

```

f6 = 0.00
f7 = 0.00
f25 = 0.00
f26 = 0.00

```

```

      f8 = 0.00
do while (f8 <= 0.00)
      f9 = 0.01
do while (f9 <= 0.01)
      f10 = 0.00
      a8 = f7+f9
      b8 = f6+f10
      aa8 = 1-f6-f7-f8-f9-f10
sa8 = (S0-Slab*(gon1*f8+goff1*a8+z1*b8+aa8)-gnad1)/S0
sb8 = (S0-Slab*(gon2*f8+goff2*a8+z2*b8+aa8)-gnad2)/S0
sc8 = (S0-Slab*(gon3*f8+goff3*a8+z3*b8+aa8)-gnad3)/S0
sd8 = (S0-Slab*(gon4*f8+goff4*a8+z4*b8+aa8)-gnad4)/S0
se8 = (S0-Slab*(gon5*f8+goff5*a8+z5*b8+aa8)-gnad5)/S0
sg8 = (S0-Slab*(gon6*f8+goff6*a8+z6*b8+aa8)-gnad6)/S0
sh8 = (S0-Slab*(gon7*f8+goff7*a8+z7*b8+aa8)-gnad7)/S0
      xa8 = (ft8t1-sa8)**2/ft8t1e**2
      xb8 = (ft8t2-sb8)**2/ft8t2e**2
      xc8 = (ft8t3-sc8)**2/ft8t3e**2
      xd8 = (ft8t4-sd8)**2/ft8t4e**2
      xe8 = (ft8t5-se8)**2/ft8t5e**2

```

```

        xg8 = (ft8t6-sg8)**2/ft8t6e**2
        xh8 = (ft8t7-sh8)**2/ft8t7e**2
do while (f10 <= 0.02)
    a9 = f8+f10
    b9 = f7+f11
    aa9 = 1-f7-f8-f9-f10-f11
sa9 = (S0-Slab*(gon1*f9+goff1*a9+b9*z1+aa9)-gnad1)/S0
sb9 = (S0-Slab*(gon2*f9+goff2*a9+b9*z2+aa9)-gnad2)/S0
sc9 = (S0-Slab*(gon3*f9+goff3*a9+b9*z3+aa9)-gnad3)/S0
sd9 = (S0-Slab*(gon4*f9+goff4*a9+b9*z4+aa9)-gnad4)/S0
se9 = (S0-Slab*(gon5*f9+goff5*a9+b9*z5+aa9)-gnad5)/S0
sg9 = (S0-Slab*(gon6*f9+goff6*a9+b9*z6+aa9)-gnad6)/S0
sh9 = (S0-Slab*(gon7*f9+goff7*a9+b9*z7+aa9)-gnad7)/S0
    xa9 = (ft9t1-sa9)**2/ft9t1e**2
    xb9 = (ft9t2-sb9)**2/ft9t2e**2
    xc9 = (ft9t3-sc9)**2/ft9t3e**2
    xd9 = (ft9t4-sd9)**2/ft9t4e**2
    xe9 = (ft9t5-se9)**2/ft9t5e**2
    xg9 = (ft9t6-sg9)**2/ft9t6e**2
    xh9 = (ft9t7-sh9)**2/ft9t7e**2
    f11 = 0.00
do while (f11 <= 0.07)
    a10 = f9+f11
    b10 = f8+f12
    aa10 = 1-f8-f9-f10-f11-f12
sa10 = (S0-Slab*(gon1*f10+goff1*a10+z1*b10+aa10)-gnad1)/S0
sb10 = (S0-Slab*(gon2*f10+goff2*a10+z2*b10+aa10)-gnad2)/S0
sc10 = (S0-Slab*(gon3*f10+goff3*a10+z3*b10+aa10)-gnad3)/S0
sd10 = (S0-Slab*(gon4*f10+goff4*a10+z4*b10+aa10)-gnad4)/S0
se10 = (S0-Slab*(gon5*f10+goff5*a10+z5*b10+aa10)-gnad5)/S0
sg10 = (S0-Slab*(gon6*f10+goff6*a10+z6*b10+aa10)-gnad6)/S0
sh10 = (S0-Slab*(gon7*f10+goff7*a10+z7*b10+aa10)-gnad7)/S0
    xa10 = (ft10t1-sa10)**2/ft10t1e**2
    xb10 = (ft10t2-sb10)**2/ft10t2e**2
    xc10 = (ft10t3-sc10)**2/ft10t3e**2
    xd10 = (ft10t4-sd10)**2/ft10t4e**2
    xe10 = (ft10t5-se10)**2/ft10t5e**2
    xg10 = (ft10t6-sg10)**2/ft10t6e**2
    xh10 = (ft10t7-sh10)**2/ft10t7e**2
    f12 = 0.02
*
*Dirk
    f12 = 0.02+0.01*i12
*
do while (f12 <= 0.15)
    a11 = f10+f12
    b11 = f9+f13
    aa11 = 1-f9-f10-f11-f12-f13
sa11 = (S0-Slab*(gon1*f11+goff1*a11+z1*b11+aa11)-gnad1)/S0
sb11 = (S0-Slab*(gon2*f11+goff2*a11+z2*b11+aa11)-gnad2)/S0
sc11 = (S0-Slab*(gon3*f11+goff3*a11+z3*b11+aa11)-gnad3)/S0
sd11 = (S0-Slab*(gon4*f11+goff4*a11+z4*b11+aa11)-gnad4)/S0
se11 = (S0-Slab*(gon5*f11+goff5*a11+z5*b11+aa11)-gnad5)/S0
sg11 = (S0-Slab*(gon6*f11+goff6*a11+z6*b11+aa11)-gnad6)/S0
sh11 = (S0-Slab*(gon7*f11+goff7*a11+z7*b11+aa11)-gnad7)/S0
    xa11 = (ft11t1-sa11)**2/ft11t1e**2
    xb11 = (ft11t2-sb11)**2/ft11t2e**2
    xc11 = (ft11t3-sc11)**2/ft11t3e**2

```

```

xd11 = (ft11t4-sd11)**2/ft11t4e**2
xe11 = (ft11t5-se11)**2/ft11t5e**2
xg11 = (ft11t6-sg11)**2/ft11t6e**2
xh11 = (ft11t7-sh11)**2/ft11t7e**2
f13 = 0.07

do while (f13 <= 0.22)
  a12 = f11+f13
  b12 = f10+f14
  aa12 = 1-f10-f11-f12-f13-f14
sa12 = (S0-Slab*(gon1*f12+goff1*a12+z1*b12+aa12)-gnad1)/S0
sb12 = (S0-Slab*(gon2*f12+goff2*a12+z2*b12+aa12)-gnad2)/S0
sc12 = (S0-Slab*(gon3*f12+goff3*a12+z3*b12+aa12)-gnad3)/S0
sd12 = (S0-Slab*(gon4*f12+goff4*a12+z4*b12+aa12)-gnad4)/S0
se12 = (S0-Slab*(gon5*f12+goff5*a12+z5*b12+aa12)-gnad5)/S0
sg12 = (S0-Slab*(gon6*f12+goff6*a12+z6*b12+aa12)-gnad6)/S0
sh12 = (S0-Slab*(gon7*f12+goff7*a12+z7*b12+aa12)-gnad7)/S0
xa12 = (ft12t1-sa12)**2/ft12t1e**2
xb12 = (ft12t2-sb12)**2/ft12t2e**2
xc12 = (ft12t3-sc12)**2/ft12t3e**2
xd12 = (ft12t4-sd12)**2/ft12t4e**2
xe12 = (ft12t5-se12)**2/ft12t5e**2
xg12 = (ft12t6-sg12)**2/ft12t6e**2
xh12 = (ft12t7-sh12)**2/ft12t7e**2
f14 = 0.00
do while (f14 <= 0.19)
  a13 = f12+f14
  b13 = f11+f15
  aa13 = 1-f11-f12-f13-f14-f15
sa13 = (S0-Slab*(gon1*f13+goff1*a13+z1*b13+aa13)-gnad1)/S0
sb13 = (S0-Slab*(gon2*f13+goff2*a13+z2*b13+aa13)-gnad2)/S0
sc13 = (S0-Slab*(gon3*f13+goff3*a13+z3*b13+aa13)-gnad3)/S0
sd13 = (S0-Slab*(gon4*f13+goff4*a13+z4*b13+aa13)-gnad4)/S0
se13 = (S0-Slab*(gon5*f13+goff5*a13+z5*b13+aa13)-gnad5)/S0
sg13 = (S0-Slab*(gon6*f13+goff6*a13+z6*b13+aa13)-gnad6)/S0
sh13 = (S0-Slab*(gon7*f13+goff7*a13+z7*b13+aa13)-gnad7)/S0
xa13 = (ft13t1-sa13)**2/ft13t1e**2
xb13 = (ft13t2-sb13)**2/ft13t2e**2
xc13 = (ft13t3-sc13)**2/ft13t3e**2
xd13 = (ft13t4-sd13)**2/ft13t4e**2
xe13 = (ft13t5-se13)**2/ft13t5e**2
xg13 = (ft13t6-sg13)**2/ft13t6e**2
xh13 = (ft13t7-sh13)**2/ft13t7e**2
f15 = 0.03

do while (f15 <= 0.23)
  a14 = f13+f15
  b14 = f12+f16
  aa14 = 1-f12-f13-f14-f15-f16
sa14 = (S0-Slab*(gon1*f14+goff1*a14+z1*b14+aa14)-gnad1)/S0
sb14 = (S0-Slab*(gon2*f14+goff2*a14+z2*b14+aa14)-gnad2)/S0
sc14 = (S0-Slab*(gon3*f14+goff3*a14+z3*b14+aa14)-gnad3)/S0
sd14 = (S0-Slab*(gon4*f14+goff4*a14+z4*b14+aa14)-gnad4)/S0
se14 = (S0-Slab*(gon5*f14+goff5*a14+z5*b14+aa14)-gnad5)/S0
sg14 = (S0-Slab*(gon6*f14+goff6*a14+z6*b14+aa14)-gnad6)/S0
sh14 = (S0-Slab*(gon7*f14+goff7*a14+z7*b14+aa14)-gnad7)/S0
xa14 = (ft14t1-sa14)**2/ft14t1e**2

```

```

xb14 = (ft14t2-sb14)**2/ft14t2e**2
xc14 = (ft14t3-sc14)**2/ft14t3e**2
xd14 = (ft14t4-sd14)**2/ft14t4e**2
xe14 = (ft14t5-se14)**2/ft14t5e**2
xg14 = (ft14t6-sg14)**2/ft14t6e**2
xh14 = (ft14t7-sh14)**2/ft14t7e**2

f16 = 0.02
do while (f16 <= 0.25)
  a15 = f14+f16
  b15 = f13+f17
  aa15 = 1-f13-f14-f15-f16-f17
  sa15 = (S0-Slab*(gon1*f15+goff1*a15+z1*b15+aa15)-gnad1)/S0
  sb15 = (S0-Slab*(gon2*f15+goff2*a15+z2*b15+aa15)-gnad2)/S0
  sc15 = (S0-Slab*(gon3*f15+goff3*a15+z3*b15+aa15)-gnad3)/S0
  sd15 = (S0-Slab*(gon4*f15+goff4*a15+z4*b15+aa15)-gnad4)/S0
  se15 = (S0-Slab*(gon5*f15+goff5*a15+z5*b15+aa15)-gnad5)/S0
  sg15 = (S0-Slab*(gon6*f15+goff6*a15+z6*b15+aa15)-gnad6)/S0
  sh15 = (S0-Slab*(gon7*f15+goff7*a15+z7*b15+aa15)-gnad7)/S0
  xa15 = (ft15t1-sa15)**2/ft15t1e**2
  xb15 = (ft15t2-sb15)**2/ft15t2e**2
  xc15 = (ft15t3-sc15)**2/ft15t3e**2
  xd15 = (ft15t4-sd15)**2/ft15t4e**2
  xe15 = (ft15t5-se15)**2/ft15t5e**2
  xg15 = (ft15t6-sg15)**2/ft15t6e**2
  xh15 = (ft15t7-sh15)**2/ft15t7e**2
  f17 = 0.07
do while (f17 <= 0.28)
  a16 = f15+f17
  b16 = f14+f18
  aa16 = 1-f14-f15-f16-f17-f18
  sa16 = (S0-Slab*(gon1*f16+goff1*a16+z1*b16+aa16)-gnad1)/S0
  sb16 = (S0-Slab*(gon2*f16+goff2*a16+z2*b16+aa16)-gnad2)/S0
  sc16 = (S0-Slab*(gon3*f16+goff3*a16+z3*b16+aa16)-gnad3)/S0
  sd16 = (S0-Slab*(gon4*f16+goff4*a16+z4*b16+aa16)-gnad4)/S0
  se16 = (S0-Slab*(gon5*f16+goff5*a16+z5*b16+aa16)-gnad5)/S0
  sg16 = (S0-Slab*(gon6*f16+goff6*a16+z6*b16+aa16)-gnad6)/S0
  sh16 = (S0-Slab*(gon7*f16+goff7*a16+z7*b16+aa16)-gnad7)/S0
  xa16 = (ft16t1-sa16)**2/ft16t1e**2
  xb16 = (ft16t2-sb16)**2/ft16t2e**2
  xc16 = (ft16t3-sc16)**2/ft16t3e**2
  xd16 = (ft16t4-sd16)**2/ft16t4e**2
  xe16 = (ft16t5-se16)**2/ft16t5e**2
  xg16 = (ft16t6-sg16)**2/ft16t6e**2
  xh16 = (ft16t7-sh16)**2/ft16t7e**2
  f18 = 0.03
do while (f18 <= 0.21)
  a17 = f16+f18
  b17 = f15+f19
  aa17 = 1-f15-f16-f17-f18-f19
  sa17 = (S0-Slab*(gon1*f17+goff1*a17+z1*b17+aa17)-gnad1)/S0
  sb17 = (S0-Slab*(gon2*f17+goff2*a17+z2*b17+aa17)-gnad2)/S0
  sc17 = (S0-Slab*(gon3*f17+goff3*a17+z3*b17+aa17)-gnad3)/S0
  sd17 = (S0-Slab*(gon4*f17+goff4*a17+z4*b17+aa17)-gnad4)/S0
  se17 = (S0-Slab*(gon5*f17+goff5*a17+z5*b17+aa17)-gnad5)/S0
  sg17 = (S0-Slab*(gon6*f17+goff6*a17+z6*b17+aa17)-gnad6)/S0

```

```

sh17 = (S0-Slab*(gon7*f17+goff7*a17+z7*b17+aa17)-gnad7)/S0
xa17 = (ft17t1-sa17)**2/ft17t1e**2
xb17 = (ft17t2-sb17)**2/ft17t2e**2
xc17 = (ft17t3-sc17)**2/ft17t3e**2
xd17 = (ft17t4-sd17)**2/ft17t4e**2
xe17 = (ft17t5-se17)**2/ft17t5e**2
xg17 = (ft17t6-sg17)**2/ft17t6e**2
xh17 = (ft17t7-sh17)**2/ft17t7e**2
f19 = 0.00
do while (f19 <= 0.14)
  a18 = f17+f19
  b18 = f16+f20
  aa18 = 1-f16-f17-f18-f19-f20
sa18 = (S0-Slab*(gon1*f18+goff1*a18+z1*b18+aa18)-gnad1)/S0
sb18 = (S0-Slab*(gon2*f18+goff2*a18+z2*b18+aa18)-gnad2)/S0
sc18 = (S0-Slab*(gon3*f18+goff3*a18+z3*b18+aa18)-gnad3)/S0
sd18 = (S0-Slab*(gon4*f18+goff4*a18+z4*b18+aa18)-gnad4)/S0
se18 = (S0-Slab*(gon5*f18+goff5*a18+z5*b18+aa18)-gnad5)/S0
sg18 = (S0-Slab*(gon6*f18+goff6*a18+z6*b18+aa18)-gnad6)/S0
sh18 = (S0-Slab*(gon7*f18+goff7*a18+z7*b18+aa18)-gnad7)/S0
xa18 = (ft18t1-sa18)**2/ft18t1e**2
xb18 = (ft18t2-sb18)**2/ft18t2e**2
xc18 = (ft18t3-sc18)**2/ft18t3e**2
xd18 = (ft18t4-sd18)**2/ft18t4e**2
xe18 = (ft18t5-se18)**2/ft18t5e**2
xg18 = (ft18t6-sg18)**2/ft18t6e**2
xh18 = (ft18t7-sh18)**2/ft18t7e**2
  f20 = 0.24
*
*Dirk
  f20 = 0.09+0.01*i20
*
  do while (f20 <= 0.49)
    a19 = f18+f20
    b19 = f17+f21
    aa19 = 1-f17-f18-f19-f20-f21
sa19 = (S0-Slab*(gon1*f19+goff1*a19+z1*b19+aa19)-gnad1)/S0
sb19 = (S0-Slab*(gon2*f19+goff2*a19+z2*b19+aa19)-gnad2)/S0
sc19 = (S0-Slab*(gon3*f19+goff3*a19+z3*b19+aa19)-gnad3)/S0
sd19 = (S0-Slab*(gon4*f19+goff4*a19+z4*b19+aa19)-gnad4)/S0
se19 = (S0-Slab*(gon5*f19+goff5*a19+z5*b19+aa19)-gnad5)/S0
sg19 = (S0-Slab*(gon6*f19+goff6*a19+z6*b19+aa19)-gnad6)/S0
sh19 = (S0-Slab*(gon7*f19+goff7*a19+z7*b19+aa19)-gnad7)/S0
xa19 = (ft19t1-sa19)**2/ft19t1e**2
xb19 = (ft19t2-sb19)**2/ft19t2e**2
xc19 = (ft19t3-sc19)**2/ft19t3e**2
xd19 = (ft19t4-sd19)**2/ft19t4e**2
xe19 = (ft19t5-se19)**2/ft19t5e**2
xg19 = (ft19t6-sg19)**2/ft19t6e**2
xh19 = (ft19t7-sh19)**2/ft19t7e**2
    f21 = 0.00
  do while (f21 <= 0.04)
    a20 = f19+f21
    b20 = f18+f22
    aa20 = 1-f18-f19-f20-f21-f22
sa20 = (S0-Slab*(gon1*f20+goff1*a20+z1*b20+aa20)-gnad1)/S0
sb20 = (S0-Slab*(gon2*f20+goff2*a20+z2*b20+aa20)-gnad2)/S0
sc20 = (S0-Slab*(gon3*f20+goff3*a20+z3*b20+aa20)-gnad3)/S0

```

```

sd20 = (S0-Slab*(gon4*f20+goff4*a20+z4*b20+aa20)-gnad4)/S0
se20 = (S0-Slab*(gon5*f20+goff5*a20+z5*b20+aa20)-gnad5)/S0
sg20 = (S0-Slab*(gon6*f20+goff6*a20+z6*b20+aa20)-gnad6)/S0
sh20 = (S0-Slab*(gon7*f20+goff7*a20+z7*b20+aa20)-gnad7)/S0
  xa20 = (ft20t1-sa20)**2/ft20t1e**2
  xb20 = (ft20t2-sb20)**2/ft20t2e**2
  xc20 = (ft20t3-sc20)**2/ft20t3e**2
  xd20 = (ft20t4-sd20)**2/ft20t4e**2
  xe20 = (ft20t5-se20)**2/ft20t5e**2
  xg20 = (ft20t6-sg20)**2/ft20t6e**2
  xh20 = (ft20t7-sh20)**2/ft20t7e**2
  f22 = 0.00
do while (f22 <= 0.04)
  a21 = f20+f22
  b21 = f19+f23
  aa21 = 1-f19-f20-f21-f22-f23
sa21 = (S0-Slab*(gon1*f21+goff1*a21+z1*b21+aa21)-gnad1)/S0
sb21 = (S0-Slab*(gon2*f21+goff2*a21+z2*b21+aa21)-gnad2)/S0
sc21 = (S0-Slab*(gon3*f21+goff3*a21+z3*b21+aa21)-gnad3)/S0
sd21 = (S0-Slab*(gon4*f21+goff4*a21+z4*b21+aa21)-gnad4)/S0
se21 = (S0-Slab*(gon5*f21+goff5*a21+z5*b21+aa21)-gnad5)/S0
sg21 = (S0-Slab*(gon6*f21+goff6*a21+z6*b21+aa21)-gnad6)/S0
sh21 = (S0-Slab*(gon7*f21+goff7*a21+z7*b21+aa21)-gnad7)/S0
  xa21 = (ft21t1-sa21)**2/ft21t1e**2
  xb21 = (ft21t2-sb21)**2/ft21t2e**2
  xc21 = (ft21t3-sc21)**2/ft21t3e**2
  xd21 = (ft21t4-sd21)**2/ft21t4e**2
  xe21 = (ft21t5-se21)**2/ft21t5e**2
  xg21 = (ft21t6-sg21)**2/ft21t6e**2
  xh21 = (ft21t7-sh21)**2/ft21t7e**2
  f23 = 0.03
do while (f23 <= 0.05)
  a22 = f21+f23
  b22 = f20+f24
  aa22 = 1-f20-f21-f22-f23-f24
sa22 = (S0-Slab*(gon1*f22+goff1*a22+z1*b22+aa22)-gnad1)/S0
sb22 = (S0-Slab*(gon2*f22+goff2*a22+z2*b22+aa22)-gnad2)/S0
sc22 = (S0-Slab*(gon3*f22+goff3*a22+z3*b22+aa22)-gnad3)/S0
sd22 = (S0-Slab*(gon4*f22+goff4*a22+z4*b22+aa22)-gnad4)/S0
se22 = (S0-Slab*(gon5*f22+goff5*a22+z5*b22+aa22)-gnad5)/S0
sg22 = (S0-Slab*(gon6*f22+goff6*a22+z6*b22+aa22)-gnad6)/S0
sh22 = (S0-Slab*(gon7*f22+goff7*a22+z7*b22+aa22)-gnad7)/S0
  xa22 = (ft22t1-sa22)**2/ft22t1e**2
  xb22 = (ft22t2-sb22)**2/ft22t2e**2
  xc22 = (ft22t3-sc22)**2/ft22t3e**2
  xd22 = (ft22t4-sd22)**2/ft22t4e**2
  xe22 = (ft22t5-se22)**2/ft22t5e**2
  xg22 = (ft22t6-sg22)**2/ft22t6e**2
  xh22 = (ft22t7-sh22)**2/ft22t7e**2
  f24 = 0.00
do while (f24 <= 0.00)
  a23 = f22+f24
  b23 = f21+f25
  aa23 = 1-f21-f22-f23-f24-f25
sa23 = (S0-Slab*(gon1*f23+goff1*a23+z1*b23+aa23)-gnad1)/S0
sb23 = (S0-Slab*(gon2*f23+goff2*a23+z2*b23+aa23)-gnad2)/S0
sc23 = (S0-Slab*(gon3*f23+goff3*a23+z3*b23+aa23)-gnad3)/S0

```

```

sd23 = (S0-Slab*(gon4*f23+goff4*a23+z4*b23+aa23)-gnad4)/S0
se23 = (S0-Slab*(gon5*f23+goff5*a23+z5*b23+aa23)-gnad5)/S0
sg23 = (S0-Slab*(gon6*f23+goff6*a23+z6*b23+aa23)-gnad6)/S0
sh23 = (S0-Slab*(gon7*f23+goff7*a23+z7*b23+aa23)-gnad7)/S0
a24 = f23+f25
b24 = f22+f26
aa24 = 1-f22-f23-f24-f25-f26
sa24 = (S0-Slab*(gon1*f24+goff1*a24+z1*b24+aa24)-gnad1)/S0
sb24 = (S0-Slab*(gon2*f24+goff2*a24+z2*b24+aa24)-gnad2)/S0
sc24 = (S0-Slab*(gon3*f24+goff3*a24+z3*b24+aa24)-gnad3)/S0
sd24 = (S0-Slab*(gon4*f24+goff4*a24+z4*b24+aa24)-gnad4)/S0
se24 = (S0-Slab*(gon5*f24+goff5*a24+z5*b24+aa24)-gnad5)/S0
sg24 = (S0-Slab*(gon6*f24+goff6*a24+z6*b24+aa24)-gnad6)/S0
sh24 = (S0-Slab*(gon7*f24+goff7*a24+z7*b24+aa24)-gnad7)/S0

xa23 = (ft23t1-sa23)**2/ft23t1e**2
xb23 = (ft23t2-sb23)**2/ft23t2e**2
xc23 = (ft23t3-sc23)**2/ft23t3e**2
xd23 = (ft23t4-sd23)**2/ft23t4e**2
xe23 = (ft23t5-se23)**2/ft23t5e**2
xg23 = (ft23t6-sg23)**2/ft23t6e**2
xh23 = (ft23t7-sh23)**2/ft23t7e**2

xa24 = (ft24t1-sa24)**2/ft24t1e**2
xb24 = (ft24t2-sb24)**2/ft24t2e**2
xc24 = (ft24t3-sc24)**2/ft24t3e**2
xd24 = (ft24t4-sd24)**2/ft24t4e**2
xe24 = (ft24t5-se24)**2/ft24t5e**2
xg24 = (ft24t6-sg24)**2/ft24t6e**2
xh24 = (ft24t7-sh24)**2/ft24t7e**2

kaisq = xa8+xb8+xc8+xd8+xe8+xg8+xh8+
& xa9+xb9+xc9+xd9+xe9+xg9+xh9+
& xa10+xb10+xc10+xd10+xe10+xg10+xh10+
& xa11+xb11+xc11+xd11+xe11+xg11+xh11+
& xa12+xb12+xc12+xd12+xe12+xg12+xh12+
& xa13+xb13+xc13+xd13+xe13+xg13+xh13+
& xa14+xb14+xc14+xd14+xe14+xg14+xh14+
& xa15+xb15+xc15+xd15+xe15+xg15+xh15+
& xa16+xb16+xc16+xd16+xe16+xg16+xh16+
& xa17+xb17+xc17+xd17+xe17+xg17+xh17+
& xa18+xb18+xc18+xd18+xe18+xg18+xh18+
& xa19+xb19+xc19+xd19+xe19+xg19+xh19+
& xa20+xb20+xc20+xd20+xe20+xg20+xh20+
& xa21+xb21+xc21+xd21+xe21+xg21+xh21+
& xa22+xb22+xc22+xd22+xe22+xg22+xh22+
& xa23+xb23+xc23+xd23+xe23+xg23+xh23+
& xa24+xb24+xc24+xd24+xe24+xg24+xh24

```

```

IF (kaisq < bestkaisq) THEN
    bestkaisq = kaisq
    bf8 = f8
    bf9 = f9
    bf10 = f10
    bf11 = f11

```



```

        bf12 = f12
        bf13 = f13
        bf14 = f14
        bf15 = f15
        bf16 = f16
        bf17 = f17
        bf18 = f18
        bf19 = f19
        bf20 = f20
        bf21 = f21
        bf22 = f22
        bf23 = f23
        bf24 = f24

endif
    f24 = f24 + 0.01
enddo
    f23 = f23 + 0.01
enddo
    f22 = f22 + 0.01
enddo
    f21 = f21 + 0.01
enddo
*Dirk
*        f20 = f20 + 0.01
*    enddo
*        f19 = f19 + 0.01
enddo
    f18 = f18 + 0.01
enddo
    f17 = f17 + 0.01
enddo
    f16 = f16 + 0.01
enddo
    f15 = f15 + 0.01
enddo
    f14 = f14 + 0.01
enddo

    f13 = f13 + 0.01
enddo
*Dirk
*        f12 = f12 + 0.01
*    enddo
*        f11 = f11 + 0.01
enddo
    f10 = f10 + 0.01
enddo
    f9 = f9 + 0.01
enddo
    f8 = f8 + 0.01

enddo
* Dirk
*    OPEN(UNIT = 12, FILE = 'values5', STATUS = 'NEW')

```

```

*          WRITE(12,*) bestkaisq,
          WRITE(*,*) bestkaisq,
&          bf8,
&          bf9,
&          bf10, bf11, bf12, bf13, bf14, bf15,
&          bf16, bf17, bf18, bf19, bf20, bf21, bf22,
&          bf23, bf24
end

```

3.5 V2E-HFP 3 Registry Fitting qsub Script, “x2_V2E”

```

#!/bin/bash
#PBS -l nodes=1:ppn=1,walltime=144:00:00,mem=2gb,feature=gbe
#PBS -j oe
#PBS -t 0-137

#change to the original working directory
cd ${PBS_O_WORKDIR}

# Define number of loops for i16
cols=22
# Define number of loops for i13 (not used)
rows=5

# Number of jobs = cols*rows (i.e. -t 0-137)

#math to figure out the variable values based on the array id
i16=`echo "${PBS_ARRAYID} % ( ${cols} + 1 )" | bc`
i13=`echo "${PBS_ARRAYID} / ( ${cols} + 1 )" | bc`

#display the command we are going to run
echo "./x2 ${i16} ${i13} > ${i16}_${i13}.txt"

#run the command with the input variables
./x2 ${i16} ${i13} > ${i16}_${i13}.txt

# Calculate the runtiem for the job
qstat -f ${PBS_JOBID}

```

3.6 V2E-HFP 3 Registry Fitting Main Script, “V2E.f”

```

* This is a comment
* This program was written by Scott Schmick 030111
*
* t values (1 = a = 482, 2 = b = 402, etc.)
  real f7 f8 f9 f10 f11 f12 f13 f14 f15 f16
  real f17 f18 f19 f20 f21 f22 f23 f24 f25
  real sa8 sb8 sc8 sd8 se8 sg8 sh8
  real sa9 sb9 sc9 sd9 se9 sg9 sh9
  real sa10 sb10 sc10 sd10 se10 sg10 sh10
  real sa11 sb11 sc11 sd11 se11 sg11 sh11
  real sa12 sb12 sc12 sd12 se12 sg12 sh12
  real sa13 sb13 sc13 sd13 se13 sg13 sh13
  real sa14 sb14 sc14 sd14 se14 sg14 sh14

```

```

real sa15 sb15 sc15 sd15 se15 sg15 sh15
real sa16 sb16 sc16 sd16 se16 sg16 sh16
real sa17 sb17 sc17 sd17 se17 sg17 sh17
real sa18 sb18 sc18 sd18 se18 sg18 sh18
real sa19 sb19 sc19 sd19 se19 sg19 sh19
real sa20 sb20 sc20 sd20 se20 sg20 sh20
real sa21 sb21 sc21 sd21 se21 sg21 sh21
real sa22 sb22 sc22 sd22 se22 sg22 sh22
real sa23 sb23 sc23 sd23 se23 sg23 sh23
real sa24 sb24 sc24 sd24 se24 sg24 sh24

real S0
real xa8 xb8 xc8 xd8 xe8 xg8 xh8
real xa9 xb9 xc9 xd9 xe9 xg9 xh9
real xa10 xb10 xc10 xd10 xe10 xg10 xh10
real xa11 xb11 xc11 xd11 xe11 xg11 xh11
real xa12 xb12 xc12 xd12 xe12 xg12 xh12
real xa13 xb13 xc13 xd13 xe13 xg13 xh13
real xa14 xb14 xc14 xd14 xe14 xg14 xh14
real xa15 xb15 xc15 xd15 xe15 xg15 xh15
real xa16 xb16 xc16 xd16 xe16 xg16 xh16
real xa17 xb17 xc17 xd17 xe17 xg17 xh17
real xa18 xb18 xc18 xd18 xe18 xg18 xh18
real xa19 xb19 xc19 xd19 xe19 xg19 xh19
real xa20 xb20 xc20 xd20 xe20 xg20 xh20
real xa21 xb21 xc21 xd21 xe21 xg21 xh21
real xa22 xb22 xc22 xd22 xe22 xg22 xh22
real xa23 xb23 xc23 xd23 xe23 xg23 xh23
real xa24 xb24 xc24 xd24 xe24 xg24 xh24

real goff1 goff2 goff3 goff4 goff5 goff6 goff7
  real gon1 gon2 gon3 gon4 gon5 gon6 gon7
  real gnad1 gnad2 gnad3 gnad4 gnad5 gnad6 gnad7
real ft8t1
real ft8t2
real ft8t3
real ft8t4
real ft8t5
real ft8t6
real ft8t7
real ft8t1e
real ft8t2e
real ft8t3e
real ft8t4e
real ft8t5e
real ft8t6e
real ft8t7e
real ft9t1
real ft9t2
real ft9t3
real ft9t4
real ft9t5
real ft9t6
real ft9t7
real ft9t1e
real ft9t2e
real ft9t3e

```

real	ft9t4e
real	ft9t5e
real	ft9t6e
real	ft9t7e
real	ft10t1
real	ft10t2
real	ft10t3
real	ft10t4
real	ft10t5
real	ft10t6
real	ft10t7
real	ft10t1e
real	ft10t2e
real	ft10t3e
real	ft10t4e
real	ft10t5e
real	ft10t6e
real	ft10t7e
real	ft11t1
real	ft11t2
real	ft11t3
real	ft11t4
real	ft11t5
real	ft11t6
real	ft11t7
real	ft11t1e
real	ft11t2e
real	ft11t3e
real	ft11t4e
real	ft11t5e
real	ft11t6e
real	ft11t7e
real	ft12t1
real	ft12t2
real	ft12t3
real	ft12t4
real	ft12t5
real	ft12t6
real	ft12t7
real	ft12t1e
real	ft12t2e
real	ft12t3e
real	ft12t4e
real	ft12t5e
real	ft12t6e
real	ft12t7e
real	ft13t1
real	ft13t2
real	ft13t3
real	ft13t4
real	ft13t5
real	ft13t6
real	ft13t7
real	ft13t1e
real	ft13t2e
real	ft13t3e
real	ft13t4e

real	ft13t5e
real	ft13t6e
real	ft13t7e
real	ft14t1
real	ft14t2
real	ft14t3
real	ft14t4
real	ft14t5
real	ft14t6
real	ft14t7
real	ft14t1e
real	ft14t2e
real	ft14t3e
real	ft14t4e
real	ft14t5e
real	ft14t6e
real	ft14t7e
real	ft15t1
real	ft15t2
real	ft15t3
real	ft15t4
real	ft15t5
real	ft15t6
real	ft15t7
real	ft15t1e
real	ft15t2e
real	ft15t3e
real	ft15t4e
real	ft15t5e
real	ft15t6e
real	ft15t7e
real	ft16t1
real	ft16t2
real	ft16t3
real	ft16t4
real	ft16t5
real	ft16t6
real	ft16t7
real	ft16t1e
real	ft16t2e
real	ft16t3e
real	ft16t4e
real	ft16t5e
real	ft16t6e
real	ft16t7e
real	ft17t1
real	ft17t2
real	ft17t3
real	ft17t4
real	ft17t5
real	ft17t6
real	ft17t7
real	ft17t1e
real	ft17t2e
real	ft17t3e
real	ft17t4e
real	ft17t5e

real	ft17t6e
real	ft17t7e
real	ft18t1
real	ft18t2
real	ft18t3
real	ft18t4
real	ft18t5
real	ft18t6
real	ft18t7
real	ft18t1e
real	ft18t2e
real	ft18t3e
real	ft18t4e
real	ft18t5e
real	ft18t6e
real	ft18t7e
real	ft19t1
real	ft19t2
real	ft19t3
real	ft19t4
real	ft19t5
real	ft19t6
real	ft19t7
real	ft19t1e
real	ft19t2e
real	ft19t3e
real	ft19t4e
real	ft19t5e
real	ft19t6e
real	ft19t7e
real	ft20t1
real	ft20t2
real	ft20t3
real	ft20t4
real	ft20t5
real	ft20t6
real	ft20t7
real	ft20t1e
real	ft20t2e
real	ft20t3e
real	ft20t4e
real	ft20t5e
real	ft20t6e
real	ft20t7e
real	ft21t1
real	ft21t2
real	ft21t3
real	ft21t4
real	ft21t5
real	ft21t6
real	ft21t7
real	ft21t1e
real	ft21t2e
real	ft21t3e
real	ft21t4e
real	ft21t5e
real	ft21t6e

real	ft21t7e
real	ft22t1
real	ft22t2
real	ft22t3
real	ft22t4
real	ft22t5
real	ft22t6
real	ft22t7
real	ft22t1e
real	ft22t2e
real	ft22t3e
real	ft22t4e
real	ft22t5e
real	ft22t6e
real	ft22t7e
real	ft23t1
real	ft23t2
real	ft23t3
real	ft23t4
real	ft23t5
real	ft23t6
real	ft23t7
real	ft23t1e
real	ft23t2e
real	ft23t3e
real	ft23t4e
real	ft23t5e
real	ft23t6e
real	ft23t7e
real	ft24t1
real	ft24t2
real	ft24t3
real	ft24t4
real	ft24t5
real	ft24t6
real	ft24t7
real	ft24t1e
real	ft24t2e
real	ft24t3e
real	ft24t4e
real	ft24t5e
real	ft24t6e
real	ft24t7e
real	kaisq
real	bestkaisq
real	bf8
real	bf9
real	bf10
real	bf11
real	bf12
real	bf13
real	bf14
real	bf15
real	bf16
real	bf17
real	bf18
real	bf19

```

real bf20
real bf21
real bf22
real bf23
real bf24
real a8 aa8 a9 aa9 a10 aa10 a11 aa11 a12 aa12 a13 aa13 a14 aa14
real a15 aa15 a16 aa16 a17 aa17 a18 aa18 a19 aa19 a20 aa20
real a21 aa21 a22 aa22 a23 aa23 a24 aa24

* DIRK

integer il6 il3
character inputarg*128

CALL getarg(1,inputarg)
read(inputarg,*) il6

CALL getarg(2,inputarg)
read(inputarg,*) il3

goff1 = 0.45011191
goff2 = 0.581357776
goff3 = 0.710329562
goff4 = 0.826254485
goff5 = 0.918588976
goff6 = 0.978470252
goff7 = 0.998385086
gon1 = 0.093897695
gon2 = 0.118554369
gon3 = 0.196398598
gon4 = 0.378629871
gon5 = 0.645290442
gon6 = 0.893821899
gon7 = 0.991709633
gnad1 = 0.281406262
gnad2 = 0.283440096
gnad3 = 0.298896559
gnad4 = 0.309714762
gnad5 = 0.320805841
gnad6 = 0.325732761
gnad7 = 0.342143946
ft8t1 = 0.05455
ft8t2 = 0.05178
ft8t3 = 0.05255
ft8t4 = 0.01784
ft8t5 = 0.02752
ft8t6 = 0.02554
ft8t7 = 0.00864
ft8t1e = 0.01171
ft8t2e = 0.01136
ft8t3e = 0.01377
ft8t4e = 0.01
ft8t5e = 0.01
ft8t6e = 0.01
ft8t7e = 0.01
ft9t1 = 0.0615
ft9t2 = 0.05812
ft9t3 = 0.04719

```


ft9t4	=	0.03191
ft9t5	=	0.02722
ft9t6	=	0.00512
ft9t7	=	0.01355
ft9t1e	=	0.01671
ft9t2e	=	0.01235
ft9t3e	=	0.01
ft9t4e	=	0.01
ft9t5e	=	0.01
ft9t6e	=	0.01
ft9t7e	=	0.01
ft10t1	=	0.06781
ft10t2	=	0.05011
ft10t3	=	0.04724
ft10t4	=	0.0408
ft10t5	=	0.0239
ft10t6	=	0.01835
ft10t7	=	0.02391
ft10t1e	=	0.01606
ft10t2e	=	0.01294
ft10t3e	=	0.01184
ft10t4e	=	0.01233
ft10t5e	=	0.01013
ft10t6e	=	0.01
ft10t7e	=	0.0104
ft11t1	=	0.06249
ft11t2	=	0.05293
ft11t3	=	0.04992
ft11t4	=	0.03957
ft11t5	=	0.031
ft11t6	=	0.00575
ft11t7	=	-0.01026
ft11t1e	=	0.01341
ft11t2e	=	0.01711
ft11t3e	=	0.01107
ft11t4e	=	0.01247
ft11t5e	=	0.01
ft11t6e	=	0.0106
ft11t7e	=	0.01
ft12t1	=	0.08531
ft12t2	=	0.08566
ft12t3	=	0.0591
ft12t4	=	0.03005
ft12t5	=	0.02208
ft12t6	=	0.0117
ft12t7	=	6.49E-04
ft12t1e	=	0.01586
ft12t2e	=	0.01
ft12t3e	=	0.01
ft12t4e	=	0.01418
ft12t5e	=	0.01
ft12t6e	=	0.01039
ft12t7e	=	0.0106
ft13t1	=	0.12877
ft13t2	=	0.07773
ft13t3	=	0.06296
ft13t4	=	0.04703

ft13t5	=	0.03825
ft13t6	=	0.04439
ft13t7	=	0.0017
ft13t1e	=	0.02489
ft13t2e	=	0.0254
ft13t3e	=	0.02
ft13t4e	=	0.0155
ft13t5e	=	0.01
ft13t6e	=	0.01673
ft13t7e	=	0.01285
ft14t1	=	0.15352
ft14t2	=	0.11403
ft14t3	=	0.07306
ft14t4	=	0.06083
ft14t5	=	0.03403
ft14t6	=	0.04299
ft14t7	=	0.00683
ft14t1e	=	0.0119
ft14t2e	=	0.01368
ft14t3e	=	0.01192
ft14t4e	=	0.01
ft14t5e	=	0.01019
ft14t6e	=	0.01263
ft14t7e	=	0.01659
ft15t1	=	0.19465
ft15t2	=	0.1749
ft15t3	=	0.14289
ft15t4	=	0.09534
ft15t5	=	0.04554
ft15t6	=	0.03886
ft15t7	=	0.00352
ft15t1e	=	0.01573
ft15t2e	=	0.01096
ft15t3e	=	0.01
ft15t4e	=	0.01
ft15t5e	=	0.01565
ft15t6e	=	0.01
ft15t7e	=	0.0131
ft16t1	=	0.28671
ft16t2	=	0.24526
ft16t3	=	0.19745
ft16t4	=	0.12897
ft16t5	=	0.07304
ft16t6	=	0.03423
ft16t7	=	0.01909
ft16t1e	=	0.02292
ft16t2e	=	0.01424
ft16t3e	=	0.01478
ft16t4e	=	0.01189
ft16t5e	=	0.01
ft16t6e	=	0.01226
ft16t7e	=	0.01049
ft17t1	=	0.31006
ft17t2	=	0.27143
ft17t3	=	0.23816
ft17t4	=	0.17499
ft17t5	=	0.07795

ft17t6	=	0.0353
ft17t7	=	0.01897
ft17t1e	=	0.02074
ft17t2e	=	0.01444
ft17t3e	=	0.01598
ft17t4e	=	0.01853
ft17t5e	=	0.01963
ft17t6e	=	0.0198
ft17t7e	=	0.01164
ft18t1	=	0.30283
ft18t2	=	0.3016
ft18t3	=	0.25411
ft18t4	=	0.19389
ft18t5	=	0.11342
ft18t6	=	0.05735
ft18t7	=	0.02475
ft18t1e	=	0.02213
ft18t2e	=	0.01776
ft18t3e	=	0.01885
ft18t4e	=	0.0122
ft18t5e	=	0.01153
ft18t6e	=	0.01299
ft18t7e	=	0.01682
ft19t1	=	0.34606
ft19t2	=	0.27994
ft19t3	=	0.21297
ft19t4	=	0.1443
ft19t5	=	0.06939
ft19t6	=	0.04606
ft19t7	=	0.01298
ft19t1e	=	0.01175
ft19t2e	=	0.01249
ft19t3e	=	0.01
ft19t4e	=	0.01044
ft19t5e	=	0.01414
ft19t6e	=	0.01
ft19t7e	=	0.0107
ft20t1	=	0.39766
ft20t2	=	0.37867
ft20t3	=	0.32995
ft20t4	=	0.26239
ft20t5	=	0.14554
ft20t6	=	0.05559
ft20t7	=	0.00911
ft20t1e	=	0.01651
ft20t2e	=	0.01394
ft20t3e	=	0.01
ft20t4e	=	0.01
ft20t5e	=	0.01
ft20t6e	=	0.01
ft20t7e	=	0.01
ft21t1	=	0.25706
ft21t2	=	0.19786
ft21t3	=	0.17866
ft21t4	=	0.1297
ft21t5	=	0.0752
ft21t6	=	0.04328

ft21t7	=	0.00665
ft21t1e	=	0.01116
ft21t2e	=	0.01489
ft21t3e	=	0.01
ft21t4e	=	0.01
ft21t5e	=	0.01
ft21t6e	=	0.01
ft21t7e	=	0.01
ft22t1	=	0.15497
ft22t2	=	0.10269
ft22t3	=	0.07505
ft22t4	=	0.05987
ft22t5	=	0.02889
ft22t6	=	0.02218
ft22t7	=	0.0051
ft22t1e	=	0.01455
ft22t2e	=	0.01357
ft22t3e	=	0.01179
ft22t4e	=	0.01182
ft22t5e	=	0.01011
ft22t6e	=	0.01134
ft22t7e	=	0.01
ft23t1	=	0.11272
ft23t2	=	0.10263
ft23t3	=	0.08713
ft23t4	=	0.07678
ft23t5	=	0.04061
ft23t6	=	0.01599
ft23t7	=	0.01091
ft23t1e	=	0.01183
ft23t2e	=	0.01
ft23t3e	=	0.01055
ft23t4e	=	0.01
ft23t5e	=	0.01
ft23t6e	=	0.01081
ft23t7e	=	0.01
ft24t1	=	0.10148
ft24t2	=	0.09025
ft24t3	=	0.05784
ft24t4	=	0.05364
ft24t5	=	0.029
ft24t6	=	0.01416
ft24t7	=	4.40E-04
ft24t1e	=	0.01731
ft24t2e	=	0.01521
ft24t3e	=	0.01928
ft24t4e	=	0.01275
ft24t5e	=	0.01
ft24t6e	=	0.01139
ft24t7e	=	0.01256

bestkaisq = 100000000

S0 = 1.32

Slab = 0.9625275

```
f7 = 0.00
f25 = 0.00

f8 = 0.00
do while (f8 <= 0.00)
  f9 = 0.00
do while (f9 <= 0.00)
  f10 = 0.00
  a8 = f7+f9
  aa8 = 1-f7-f8-f9
  sa8 = (S0-Slab*(gon1*f8+goff1*(a8)+aa8)-gnad1)/S0
  sb8 = (S0-Slab*(gon2*f8+goff2*(a8)+aa8)-gnad2)/S0
  sc8 = (S0-Slab*(gon3*f8+goff3*(a8)+aa8)-gnad3)/S0
  sd8 = (S0-Slab*(gon4*f8+goff4*(a8)+aa8)-gnad4)/S0
  se8 = (S0-Slab*(gon5*f8+goff5*(a8)+aa8)-gnad5)/S0
  sg8 = (S0-Slab*(gon6*f8+goff6*(a8)+aa8)-gnad6)/S0
  sh8 = (S0-Slab*(gon7*f8+goff7*(a8)+aa8)-gnad7)/S0
  xa8 = (ft8t1-sa8)**2/ft8t1e**2
  xb8 = (ft8t2-sb8)**2/ft8t2e**2
  xc8 = (ft8t3-sc8)**2/ft8t3e**2
  xd8 = (ft8t4-sd8)**2/ft8t4e**2
  xe8 = (ft8t5-se8)**2/ft8t5e**2
  xg8 = (ft8t6-sg8)**2/ft8t6e**2
  xh8 = (ft8t7-sh8)**2/ft8t7e**2
do while (f10 <= 0.00)
  a9 = f8+f10
  aa9 = 1-f8-f9-f10
  sa9 = (S0-Slab*(gon1*f9+goff1*(a9)+aa9)-gnad1)/S0
  sb9 = (S0-Slab*(gon2*f9+goff2*(a9)+aa9)-gnad2)/S0
  sc9 = (S0-Slab*(gon3*f9+goff3*(a9)+aa9)-gnad3)/S0
  sd9 = (S0-Slab*(gon4*f9+goff4*(a9)+aa9)-gnad4)/S0
  se9 = (S0-Slab*(gon5*f9+goff5*(a9)+aa9)-gnad5)/S0
  sg9 = (S0-Slab*(gon6*f9+goff6*(a9)+aa9)-gnad6)/S0
  sh9 = (S0-Slab*(gon7*f9+goff7*(a9)+aa9)-gnad7)/S0
  xa9 = (ft9t1-sa9)**2/ft9t1e**2
  xb9 = (ft9t2-sb9)**2/ft9t2e**2
  xc9 = (ft9t3-sc9)**2/ft9t3e**2
  xd9 = (ft9t4-sd9)**2/ft9t4e**2
  xe9 = (ft9t5-se9)**2/ft9t5e**2
  xg9 = (ft9t6-sg9)**2/ft9t6e**2
  xh9 = (ft9t7-sh9)**2/ft9t7e**2
  f11 = 0.00
do while (f11 <= 0.00)
  a10 = f9+f11
  aa10 = 1-f9-f10-f11
  sa10 = (S0-Slab*(gon1*f10+goff1*(a10)+aa10)-gnad1)/S0
  sb10 = (S0-Slab*(gon2*f10+goff2*(a10)+aa10)-gnad2)/S0
  sc10 = (S0-Slab*(gon3*f10+goff3*(a10)+aa10)-gnad3)/S0
  sd10 = (S0-Slab*(gon4*f10+goff4*(a10)+aa10)-gnad4)/S0
  se10 = (S0-Slab*(gon5*f10+goff5*(a10)+aa10)-gnad5)/S0
  sg10 = (S0-Slab*(gon6*f10+goff6*(a10)+aa10)-gnad6)/S0
  sh10 = (S0-Slab*(gon7*f10+goff7*(a10)+aa10)-gnad7)/S0
  xa10 = (ft10t1-sa10)**2/ft10t1e**2
  xb10 = (ft10t2-sb10)**2/ft10t2e**2
```

```

        xc10 = (ft10t3-sc10)**2/ft10t3e**2
        xd10 = (ft10t4-sd10)**2/ft10t4e**2
        xe10 = (ft10t5-se10)**2/ft10t5e**2
        xg10 = (ft10t6-sg10)**2/ft10t6e**2
        xh10 = (ft10t7-sh10)**2/ft10t7e**2
        f12 = 0.00
do while (f12 <= 0.03)
    a11 = f10+f12
    aa11 = 1-f10-f11-f12
    sa11 = (S0-Slab*(gon1*f11+goff1*(a11)+aa11)-gnad1)/S0
    sb11 = (S0-Slab*(gon2*f11+goff2*(a11)+aa11)-gnad2)/S0
    sc11 = (S0-Slab*(gon3*f11+goff3*(a11)+aa11)-gnad3)/S0
    sd11 = (S0-Slab*(gon4*f11+goff4*(a11)+aa11)-gnad4)/S0
    se11 = (S0-Slab*(gon5*f11+goff5*(a11)+aa11)-gnad5)/S0
    sg11 = (S0-Slab*(gon6*f11+goff6*(a11)+aa11)-gnad6)/S0
    sh11 = (S0-Slab*(gon7*f11+goff7*(a11)+aa11)-gnad7)/S0
    xa11 = (ft11t1-sa11)**2/ft11t1e**2
    xb11 = (ft11t2-sb11)**2/ft11t2e**2
    xc11 = (ft11t3-sc11)**2/ft11t3e**2
    xd11 = (ft11t4-sd11)**2/ft11t4e**2
    xe11 = (ft11t5-se11)**2/ft11t5e**2
    xg11 = (ft11t6-sg11)**2/ft11t6e**2
    xh11 = (ft11t7-sh11)**2/ft11t7e**2
*       f13 = 0.08
* DIRK
        f13 = 0.00+0.01*i13
*       do while (f13 <= 0.22)
            a12 = f11+f13
            aa12 = 1-f11-f12-f13
            sa12 = (S0-Slab*(gon1*f12+goff1*(a12)+aa12)-gnad1)/S0
            sb12 = (S0-Slab*(gon2*f12+goff2*(a12)+aa12)-gnad2)/S0
            sc12 = (S0-Slab*(gon3*f12+goff3*(a12)+aa12)-gnad3)/S0
            sd12 = (S0-Slab*(gon4*f12+goff4*(a12)+aa12)-gnad4)/S0
            se12 = (S0-Slab*(gon5*f12+goff5*(a12)+aa12)-gnad5)/S0
            sg12 = (S0-Slab*(gon6*f12+goff6*(a12)+aa12)-gnad6)/S0
            sh12 = (S0-Slab*(gon7*f12+goff7*(a12)+aa12)-gnad7)/S0
            xa12 = (ft12t1-sa12)**2/ft12t1e**2
            xb12 = (ft12t2-sb12)**2/ft12t2e**2
            xc12 = (ft12t3-sc12)**2/ft12t3e**2
            xd12 = (ft12t4-sd12)**2/ft12t4e**2
            xe12 = (ft12t5-se12)**2/ft12t5e**2
            xg12 = (ft12t6-sg12)**2/ft12t6e**2
            xh12 = (ft12t7-sh12)**2/ft12t7e**2
            f14 = 0.00
do while (f14 <= 0.09)
    a13 = f12+f14
    aa13 = 1-f12-f13-f14
    sa13 = (S0-Slab*(gon1*f13+goff1*(a13)+aa13)-gnad1)/S0
    sb13 = (S0-Slab*(gon2*f13+goff2*(a13)+aa13)-gnad2)/S0
    sc13 = (S0-Slab*(gon3*f13+goff3*(a13)+aa13)-gnad3)/S0
    sd13 = (S0-Slab*(gon4*f13+goff4*(a13)+aa13)-gnad4)/S0
    se13 = (S0-Slab*(gon5*f13+goff5*(a13)+aa13)-gnad5)/S0
    sg13 = (S0-Slab*(gon6*f13+goff6*(a13)+aa13)-gnad6)/S0
    sh13 = (S0-Slab*(gon7*f13+goff7*(a13)+aa13)-gnad7)/S0
    xa13 = (ft13t1-sa13)**2/ft13t1e**2
    xb13 = (ft13t2-sb13)**2/ft13t2e**2
    xc13 = (ft13t3-sc13)**2/ft13t3e**2

```

```

        xd13 = (ft13t4-sd13)**2/ft13t4e**2
        xe13 = (ft13t5-se13)**2/ft13t5e**2
        xg13 = (ft13t6-sg13)**2/ft13t6e**2
        xh13 = (ft13t7-sh13)**2/ft13t7e**2
        f15 = 0.00
do while (f15 <= 0.17)
    a14 = f13+f15
    aa14 = 1-f13-f14-f15
    sa14 = (S0-Slab*(gon1*f14+goff1*(a14)+aa14)-gnad1)/S0
    sb14 = (S0-Slab*(gon2*f14+goff2*(a14)+aa14)-gnad2)/S0
    sc14 = (S0-Slab*(gon3*f14+goff3*(a14)+aa14)-gnad3)/S0
    sd14 = (S0-Slab*(gon4*f14+goff4*(a14)+aa14)-gnad4)/S0
    se14 = (S0-Slab*(gon5*f14+goff5*(a14)+aa14)-gnad5)/S0
    sg14 = (S0-Slab*(gon6*f14+goff6*(a14)+aa14)-gnad6)/S0
    sh14 = (S0-Slab*(gon7*f14+goff7*(a14)+aa14)-gnad7)/S0
    xa14 = (ft14t1-sa14)**2/ft14t1e**2
    xb14 = (ft14t2-sb14)**2/ft14t2e**2
    xc14 = (ft14t3-sc14)**2/ft14t3e**2
    xd14 = (ft14t4-sd14)**2/ft14t4e**2
    xe14 = (ft14t5-se14)**2/ft14t5e**2
    xg14 = (ft14t6-sg14)**2/ft14t6e**2
    xh14 = (ft14t7-sh14)**2/ft14t7e**2

* DIRK
        f16 = 0.04+0.01*i16
*
*       f16 = 0.00
*       do while (f16 <= 0.03)
            a15 = f14+f16
            aa15 = 1-f14-f15-f16
            sa15 = (S0-Slab*(gon1*f15+goff1*(a15)+aa15)-gnad1)/S0
            sb15 = (S0-Slab*(gon2*f15+goff2*(a15)+aa15)-gnad2)/S0
            sc15 = (S0-Slab*(gon3*f15+goff3*(a15)+aa15)-gnad3)/S0
            sd15 = (S0-Slab*(gon4*f15+goff4*(a15)+aa15)-gnad4)/S0
            se15 = (S0-Slab*(gon5*f15+goff5*(a15)+aa15)-gnad5)/S0
            sg15 = (S0-Slab*(gon6*f15+goff6*(a15)+aa15)-gnad6)/S0
            sh15 = (S0-Slab*(gon7*f15+goff7*(a15)+aa15)-gnad7)/S0
            xa15 = (ft15t1-sa15)**2/ft15t1e**2
            xb15 = (ft15t2-sb15)**2/ft15t2e**2
            xc15 = (ft15t3-sc15)**2/ft15t3e**2
            xd15 = (ft15t4-sd15)**2/ft15t4e**2
            xe15 = (ft15t5-se15)**2/ft15t5e**2
            xg15 = (ft15t6-sg15)**2/ft15t6e**2
            xh15 = (ft15t7-sh15)**2/ft15t7e**2
            f17 = 0.05
do while (f17 <= 0.33)
    a16 = f15+f17
    aa16 = 1-f15-f16-f17
    sa16 = (S0-Slab*(gon1*f16+goff1*(a16)+aa16)-gnad1)/S0
    sb16 = (S0-Slab*(gon2*f16+goff2*(a16)+aa16)-gnad2)/S0
    sc16 = (S0-Slab*(gon3*f16+goff3*(a16)+aa16)-gnad3)/S0
    sd16 = (S0-Slab*(gon4*f16+goff4*(a16)+aa16)-gnad4)/S0
    se16 = (S0-Slab*(gon5*f16+goff5*(a16)+aa16)-gnad5)/S0
    sg16 = (S0-Slab*(gon6*f16+goff6*(a16)+aa16)-gnad6)/S0
    sh16 = (S0-Slab*(gon7*f16+goff7*(a16)+aa16)-gnad7)/S0
    xa16 = (ft16t1-sa16)**2/ft16t1e**2
    xb16 = (ft16t2-sb16)**2/ft16t2e**2
    xc16 = (ft16t3-sc16)**2/ft16t3e**2
    xd16 = (ft16t4-sd16)**2/ft16t4e**2

```

```

    xe16 = (ft16t5-se16)**2/ft16t5e**2
    xg16 = (ft16t6-sg16)**2/ft16t6e**2
    xh16 = (ft16t7-sh16)**2/ft16t7e**2
    f18 = 0.10
do while (f18 <= 0.36)
    a17 = f16+f18
    aa17 = 1-f16-f17-f18
    sa17 = (S0-Slab*(gon1*f17+goff1*(a17)+aa17)-gnad1)/S0
    sb17 = (S0-Slab*(gon2*f17+goff2*(a17)+aa17)-gnad2)/S0
    sc17 = (S0-Slab*(gon3*f17+goff3*(a17)+aa17)-gnad3)/S0
    sd17 = (S0-Slab*(gon4*f17+goff4*(a17)+aa17)-gnad4)/S0
    se17 = (S0-Slab*(gon5*f17+goff5*(a17)+aa17)-gnad5)/S0
    sg17 = (S0-Slab*(gon6*f17+goff6*(a17)+aa17)-gnad6)/S0
    sh17 = (S0-Slab*(gon7*f17+goff7*(a17)+aa17)-gnad7)/S0
    xa17 = (ft17t1-sa17)**2/ft17t1e**2
    xb17 = (ft17t2-sb17)**2/ft17t2e**2
    xc17 = (ft17t3-sc17)**2/ft17t3e**2
    xd17 = (ft17t4-sd17)**2/ft17t4e**2
    xe17 = (ft17t5-se17)**2/ft17t5e**2
    xg17 = (ft17t6-sg17)**2/ft17t6e**2
    xh17 = (ft17t7-sh17)**2/ft17t7e**2
    f19 = 0.00
do while (f19 <= 0.33)
    a18 = f17+f19
    aa18 = 1-f17-f18-f19
    sa18 = (S0-Slab*(gon1*f18+goff1*(a18)+aa18)-gnad1)/S0
    sb18 = (S0-Slab*(gon2*f18+goff2*(a18)+aa18)-gnad2)/S0
    sc18 = (S0-Slab*(gon3*f18+goff3*(a18)+aa18)-gnad3)/S0
    sd18 = (S0-Slab*(gon4*f18+goff4*(a18)+aa18)-gnad4)/S0
    se18 = (S0-Slab*(gon5*f18+goff5*(a18)+aa18)-gnad5)/S0
    sg18 = (S0-Slab*(gon6*f18+goff6*(a18)+aa18)-gnad6)/S0
    sh18 = (S0-Slab*(gon7*f18+goff7*(a18)+aa18)-gnad7)/S0
    xa18 = (ft18t1-sa18)**2/ft18t1e**2
    xb18 = (ft18t2-sb18)**2/ft18t2e**2
    xc18 = (ft18t3-sc18)**2/ft18t3e**2
    xd18 = (ft18t4-sd18)**2/ft18t4e**2
    xe18 = (ft18t5-se18)**2/ft18t5e**2
    xg18 = (ft18t6-sg18)**2/ft18t6e**2
    xh18 = (ft18t7-sh18)**2/ft18t7e**2
    f20 = 0.28
do while (f20 <= 0.49)
    a19 = f18+f20
    aa19 = 1-f18-f19-f20
    sa19 = (S0-Slab*(gon1*f19+goff1*(a19)+aa19)-gnad1)/S0
    sb19 = (S0-Slab*(gon2*f19+goff2*(a19)+aa19)-gnad2)/S0
    sc19 = (S0-Slab*(gon3*f19+goff3*(a19)+aa19)-gnad3)/S0
    sd19 = (S0-Slab*(gon4*f19+goff4*(a19)+aa19)-gnad4)/S0
    se19 = (S0-Slab*(gon5*f19+goff5*(a19)+aa19)-gnad5)/S0
    sg19 = (S0-Slab*(gon6*f19+goff6*(a19)+aa19)-gnad6)/S0
    sh19 = (S0-Slab*(gon7*f19+goff7*(a19)+aa19)-gnad7)/S0
    xa19 = (ft19t1-sa19)**2/ft19t1e**2
    xb19 = (ft19t2-sb19)**2/ft19t2e**2
    xc19 = (ft19t3-sc19)**2/ft19t3e**2
    xd19 = (ft19t4-sd19)**2/ft19t4e**2
    xe19 = (ft19t5-se19)**2/ft19t5e**2
    xg19 = (ft19t6-sg19)**2/ft19t6e**2
    xh19 = (ft19t7-sh19)**2/ft19t7e**2

```



```

    f21 = 0.05
do while (f21 <= 0.24)
    a20 = f19+f21
    aa20 = 1-f19-f20-f21
    sa20 = (S0-Slab*(gon1*f20+goff1*(a20)+aa20)-gnad1)/S0
    sb20 = (S0-Slab*(gon2*f20+goff2*(a20)+aa20)-gnad2)/S0
    sc20 = (S0-Slab*(gon3*f20+goff3*(a20)+aa20)-gnad3)/S0
    sd20 = (S0-Slab*(gon4*f20+goff4*(a20)+aa20)-gnad4)/S0
    se20 = (S0-Slab*(gon5*f20+goff5*(a20)+aa20)-gnad5)/S0
    sg20 = (S0-Slab*(gon6*f20+goff6*(a20)+aa20)-gnad6)/S0
    sh20 = (S0-Slab*(gon7*f20+goff7*(a20)+aa20)-gnad7)/S0
    xa20 = (ft20t1-sa20)**2/ft20t1e**2
    xb20 = (ft20t2-sb20)**2/ft20t2e**2
    xc20 = (ft20t3-sc20)**2/ft20t3e**2
    xd20 = (ft20t4-sd20)**2/ft20t4e**2
    xe20 = (ft20t5-se20)**2/ft20t5e**2
    xg20 = (ft20t6-sg20)**2/ft20t6e**2
    xh20 = (ft20t7-sh20)**2/ft20t7e**2
    f22 = 0.00
do while (f22 <= 0.08)
    a21 = f20+f22
    aa21 = 1-f20-f21-f22
    sa21 = (S0-Slab*(gon1*f21+goff1*(a21)+aa21)-gnad1)/S0
    sb21 = (S0-Slab*(gon2*f21+goff2*(a21)+aa21)-gnad2)/S0
    sc21 = (S0-Slab*(gon3*f21+goff3*(a21)+aa21)-gnad3)/S0
    sd21 = (S0-Slab*(gon4*f21+goff4*(a21)+aa21)-gnad4)/S0
    se21 = (S0-Slab*(gon5*f21+goff5*(a21)+aa21)-gnad5)/S0
    sg21 = (S0-Slab*(gon6*f21+goff6*(a21)+aa21)-gnad6)/S0
    sh21 = (S0-Slab*(gon7*f21+goff7*(a21)+aa21)-gnad7)/S0
    xa21 = (ft21t1-sa21)**2/ft21t1e**2
    xb21 = (ft21t2-sb21)**2/ft21t2e**2
    xc21 = (ft21t3-sc21)**2/ft21t3e**2
    xd21 = (ft21t4-sd21)**2/ft21t4e**2
    xe21 = (ft21t5-se21)**2/ft21t5e**2
    xg21 = (ft21t6-sg21)**2/ft21t6e**2
    xh21 = (ft21t7-sh21)**2/ft21t7e**2
    f23 = 0.03
do while (f23 <= 0.08)
    a22 = f21+f23
    aa22 = 1-f21-f22-f23
    sa22 = (S0-Slab*(gon1*f22+goff1*(a22)+aa22)-gnad1)/S0
    sb22 = (S0-Slab*(gon2*f22+goff2*(a22)+aa22)-gnad2)/S0
    sc22 = (S0-Slab*(gon3*f22+goff3*(a22)+aa22)-gnad3)/S0
    sd22 = (S0-Slab*(gon4*f22+goff4*(a22)+aa22)-gnad4)/S0
    se22 = (S0-Slab*(gon5*f22+goff5*(a22)+aa22)-gnad5)/S0
    sg22 = (S0-Slab*(gon6*f22+goff6*(a22)+aa22)-gnad6)/S0
    sh22 = (S0-Slab*(gon7*f22+goff7*(a22)+aa22)-gnad7)/S0
    xa22 = (ft22t1-sa22)**2/ft22t1e**2
    xb22 = (ft22t2-sb22)**2/ft22t2e**2
    xc22 = (ft22t3-sc22)**2/ft22t3e**2
    xd22 = (ft22t4-sd22)**2/ft22t4e**2
    xe22 = (ft22t5-se22)**2/ft22t5e**2
    xg22 = (ft22t6-sg22)**2/ft22t6e**2
    xh22 = (ft22t7-sh22)**2/ft22t7e**2
    f24 = 0.02
do while (f24 <= 0.04)
    a23 = f22+f24

```

```

aa23 = 1-f22-f23-f24
sa23 = (S0-Slab*(gon1*f23+goff1*(a23)+aa23)-gnad1)/S0
sb23 = (S0-Slab*(gon2*f23+goff2*(a23)+aa23)-gnad2)/S0
sc23 = (S0-Slab*(gon3*f23+goff3*(a23)+aa23)-gnad3)/S0
sd23 = (S0-Slab*(gon4*f23+goff4*(a23)+aa23)-gnad4)/S0
se23 = (S0-Slab*(gon5*f23+goff5*(a23)+aa23)-gnad5)/S0
sg23 = (S0-Slab*(gon6*f23+goff6*(a23)+aa23)-gnad6)/S0
sh23 = (S0-Slab*(gon7*f23+goff7*(a23)+aa23)-gnad7)/S0
a24 = f23+f25
aa24 = 1-f23-f24-f25
sa24 = (S0-Slab*(gon1*f24+goff1*(a24)+aa24)-gnad1)/S0
sb24 = (S0-Slab*(gon2*f24+goff2*(a24)+aa24)-gnad2)/S0
sc24 = (S0-Slab*(gon3*f24+goff3*(a24)+aa24)-gnad3)/S0
sd24 = (S0-Slab*(gon4*f24+goff4*(a24)+aa24)-gnad4)/S0
se24 = (S0-Slab*(gon5*f24+goff5*(a24)+aa24)-gnad5)/S0
sg24 = (S0-Slab*(gon6*f24+goff6*(a24)+aa24)-gnad6)/S0
sh24 = (S0-Slab*(gon7*f24+goff7*(a24)+aa24)-gnad7)/S0

xa23 = (ft23t1-sa23)**2/ft23t1e**2
xb23 = (ft23t2-sb23)**2/ft23t2e**2
xc23 = (ft23t3-sc23)**2/ft23t3e**2
xd23 = (ft23t4-sd23)**2/ft23t4e**2
xe23 = (ft23t5-se23)**2/ft23t5e**2
xg23 = (ft23t6-sg23)**2/ft23t6e**2
xh23 = (ft23t7-sh23)**2/ft23t7e**2

xa24 = (ft24t1-sa24)**2/ft24t1e**2
xb24 = (ft24t2-sb24)**2/ft24t2e**2
xc24 = (ft24t3-sc24)**2/ft24t3e**2
xd24 = (ft24t4-sd24)**2/ft24t4e**2
xe24 = (ft24t5-se24)**2/ft24t5e**2
xg24 = (ft24t6-sg24)**2/ft24t6e**2
xh24 = (ft24t7-sh24)**2/ft24t7e**2

kaisq = xa8+xb8+xc8+xd8+xe8+xg8+xh8+
& xa9+xb9+xc9+xd9+xe9+xg9+xh9+
& xa10+xb10+xc10+xd10+xe10+xg10+xh10+
& xa11+xb11+xc11+xd11+xe11+xg11+xh11+
& xa12+xb12+xc12+xd12+xe12+xg12+xh12+
& xa13+xb13+xc13+xd13+xe13+xg13+xh13+
& xa14+xb14+xc14+xd14+xe14+xg14+xh14+
& xa15+xb15+xc15+xd15+xe15+xg15+xh15+
& xa16+xb16+xc16+xd16+xe16+xg16+xh16+
& xa17+xb17+xc17+xd17+xe17+xg17+xh17+
& xa18+xb18+xc18+xd18+xe18+xg18+xh18+
& xa19+xb19+xc19+xd19+xe19+xg19+xh19+
& xa20+xb20+xc20+xd20+xe20+xg20+xh20+
& xa21+xb21+xc21+xd21+xe21+xg21+xh21+
& xa22+xb22+xc22+xd22+xe22+xg22+xh22+
& xa23+xb23+xc23+xd23+xe23+xg23+xh23+
& xa24+xb24+xc24+xd24+xe24+xg24+xh24

```

```

IF (kaisq < bestkaisq) THEN
    bestkaisq = kaisq
    bf8 = f8

```

```

        bf9 = f9
        bf10 = f10
        bf11 = f11
        bf12 = f12
        bf13 = f13
        bf14 = f14
        bf15 = f15
        bf16 = f16
        bf17 = f17
        bf18 = f18
        bf19 = f19
        bf20 = f20
        bf21 = f21
        bf22 = f22
        bf23 = f23
        bf24 = f24

    endif
    f24 = f24 + 0.01
enddo
    f23 = f23 + 0.01
enddo
    f22 = f22 + 0.01
enddo
    f21 = f21 + 0.01
enddo
    f20 = f20 + 0.01
enddo
    f19 = f19 + 0.01
enddo
    f18 = f18 + 0.01
enddo
    f17 = f17 + 0.01
enddo
* Dirk
*
*     f16 = f16 + 0.01
*     enddo
*     f15 = f15 + 0.01
*     enddo
*     f14 = f14 + 0.01
*     enddo
*Dirk
*
*     f13 = f13 + 0.01
*     enddo
*     f12 = f12 + 0.01
*     enddo
*     f11 = f11 + 0.01
*     enddo
*     f10 = f10 + 0.01
*     enddo
*     f9 = f9 + 0.01
*     enddo
*     f8 = f8 + 0.01
*
enddo

```

```

* Dirk
*      OPEN(UNIT = 12, FILE = 'values5', STATUS = 'NEW')
*      WRITE(12,*) bestkaisq,
*      WRITE(*,*) bestkaisq,
*      &          bf8,
*      &          bf9,
*      &          bf10, bf11, bf12, bf13, bf14, bf15,
*      &          bf16, bf17, bf18, bf19, bf20, bf21, bf22,
*      &          bf23, bf24
*      end

```

3.7 V2E-HFP 5 Registry Fitting qsub Script, “x2_V2E”

For V2E-HFP 5 registry fittings, four main script files were created (see **Appendix I**). In the file below, $f_{12} = 0.00$. The other main script files used set $f_{12} = 0.01$, $f_{12} = 0.02$ and $f_{12} = 0.03$. Splitting the main script file into four separate jobs was required to complete each script file’s computations in less than 168 hours (i.e. the maximum time allowed to occupy a node). This time limit is set by the High Performance Computing Center and jobs that run longer than 168 hours are automatically terminated.

```

#!/bin/bash
#PBS -l nodes=1:ppn=1,walltime=168:00:00,mem=2gb,feature=gbe
#PBS -j oe
#PBS -t 0-143

#change to the original working directory
cd ${PBS_O_WORKDIR}

# Define number of loops for i16
cols=17
# Define number of loops for i13 (not used)
rows=7

# Number of jobs = cols*rows (i.e. -t 0-137)

#math to figure out the variable values based on the array id
i15=`echo "${PBS_ARRAYID} % ( ${cols} + 1 )" | bc`
i23=`echo "${PBS_ARRAYID} / ( ${cols} + 1 )" | bc`

#display the command we are going to run
echo "./x2 ${i15} ${i23} > ${i15}_${i23}.txt"

#run the command with the input variables
./x2 ${i15} ${i23} > ${i15}_${i23}.txt

# Calculate the runtiem for the job
qstat -f ${PBS_JOBID}

```

3.8 V2E-HFP 5 Registry Fitting Main Script, “V2E_5var.f”

```
* This is a comment
* This program was written by Scott Schmick 030111
*
* t values (1 = a = 482, 2 = b = 402, etc.)
  real f7 f8 f9 f10 f11 f12 f13 f14 f15 f16
  real f17 f18 f19 f20 f21 f22 f23 f24 f25
  real sa8 sb8 sc8 sd8 se8 sg8 sh8
  real sa9 sb9 sc9 sd9 se9 sg9 sh9
  real sa10 sb10 sc10 sd10 se10 sg10 sh10
  real sa11 sb11 sc11 sd11 se11 sg11 sh11
  real sa12 sb12 sc12 sd12 se12 sg12 sh12
  real sa13 sb13 sc13 sd13 se13 sg13 sh13
  real sa14 sb14 sc14 sd14 se14 sg14 sh14
  real sa15 sb15 sc15 sd15 se15 sg15 sh15
  real sa16 sb16 sc16 sd16 se16 sg16 sh16
  real sa17 sb17 sc17 sd17 se17 sg17 sh17
  real sa18 sb18 sc18 sd18 se18 sg18 sh18
  real sa19 sb19 sc19 sd19 se19 sg19 sh19
  real sa20 sb20 sc20 sd20 se20 sg20 sh20
  real sa21 sb21 sc21 sd21 se21 sg21 sh21
  real sa22 sb22 sc22 sd22 se22 sg22 sh22
  real sa23 sb23 sc23 sd23 se23 sg23 sh23
  real sa24 sb24 sc24 sd24 se24 sg24 sh24

  real S0
  real xa8 xb8 xc8 xd8 xe8 xg8 xh8
  real xa9 xb9 xc9 xd9 xe9 xg9 xh9
  real xa10 xb10 xc10 xd10 xe10 xg10 xh10
  real xa11 xb11 xc11 xd11 xe11 xg11 xh11
  real xa12 xb12 xc12 xd12 xe12 xg12 xh12
  real xa13 xb13 xc13 xd13 xe13 xg13 xh13
  real xa14 xb14 xc14 xd14 xe14 xg14 xh14
  real xa15 xb15 xc15 xd15 xe15 xg15 xh15
  real xa16 xb16 xc16 xd16 xe16 xg16 xh16
  real xa17 xb17 xc17 xd17 xe17 xg17 xh17
  real xa18 xb18 xc18 xd18 xe18 xg18 xh18
  real xa19 xb19 xc19 xd19 xe19 xg19 xh19
  real xa20 xb20 xc20 xd20 xe20 xg20 xh20
  real xa21 xb21 xc21 xd21 xe21 xg21 xh21
  real xa22 xb22 xc22 xd22 xe22 xg22 xh22
  real xa23 xb23 xc23 xd23 xe23 xg23 xh23
  real xa24 xb24 xc24 xd24 xe24 xg24 xh24

  real goff1 goff2 goff3 goff4 goff5 goff6 goff7
  real gon1 gon2 gon3 gon4 gon5 gon6 gon7
  real gnad1 gnad2 gnad3 gnad4 gnad5 gnad6 gnad7
  real ft8t1
  real ft8t2
  real ft8t3
  real ft8t4
  real ft8t5
  real ft8t6
  real ft8t7
```

real	ft8t1e	
real	ft8t2e	
real	ft8t3e	
real	ft8t4e	
real	ft8t5e	
real	ft8t6e	
real	ft8t7e	
real	ft9t1	
real	ft9t2	
real	ft9t3	
real	ft9t4	
real	ft9t5	
real	ft9t6	
real	ft9t7	
real	ft9t1e	
real	ft9t2e	
real	ft9t3e	
real	ft9t4e	
real	ft9t5e	
real	ft9t6e	
real	ft9t7e	
real	ft10t1	
real	ft10t2	
real	ft10t3	
real	ft10t4	
real	ft10t5	
real	ft10t6	
real	ft10t7	
real	ft10t1e	
real	ft10t2e	
real	ft10t3e	
real	ft10t4e	
real	ft10t5e	
real	ft10t6e	
real	ft10t7e	
real		ft11t1
real		ft11t2
real		ft11t3
real		ft11t4
real		ft11t5
real		ft11t6
real		ft11t7
real		ft11t1e
real		ft11t2e
real		ft11t3e
real		ft11t4e
real		ft11t5e
real		ft11t6e
real		ft11t7e
real		ft12t1
real		ft12t2
real		ft12t3
real		ft12t4
real		ft12t5
real		ft12t6
real		ft12t7
real		ft12t1e

real	ft12t2e
real	ft12t3e
real	ft12t4e
real	ft12t5e
real	ft12t6e
real	ft12t7e
real	ft13t1
real	ft13t2
real	ft13t3
real	ft13t4
real	ft13t5
real	ft13t6
real	ft13t7
real	ft13t1e
real	ft13t2e
real	ft13t3e
real	ft13t4e
real	ft13t5e
real	ft13t6e
real	ft13t7e
real	ft14t1
real	ft14t2
real	ft14t3
real	ft14t4
real	ft14t5
real	ft14t6
real	ft14t7
real	ft14t1e
real	ft14t2e
real	ft14t3e
real	ft14t4e
real	ft14t5e
real	ft14t6e
real	ft14t7e
real	ft15t1
real	ft15t2
real	ft15t3
real	ft15t4
real	ft15t5
real	ft15t6
real	ft15t7
real	ft15t1e
real	ft15t2e
real	ft15t3e
real	ft15t4e
real	ft15t5e
real	ft15t6e
real	ft15t7e
real	ft16t1
real	ft16t2
real	ft16t3
real	ft16t4
real	ft16t5
real	ft16t6
real	ft16t7
real	ft16t1e
real	ft16t2e

real	ft16t3e
real	ft16t4e
real	ft16t5e
real	ft16t6e
real	ft16t7e
real	ft17t1
real	ft17t2
real	ft17t3
real	ft17t4
real	ft17t5
real	ft17t6
real	ft17t7
real	ft17t1e
real	ft17t2e
real	ft17t3e
real	ft17t4e
real	ft17t5e
real	ft17t6e
real	ft17t7e
real	ft18t1
real	ft18t2
real	ft18t3
real	ft18t4
real	ft18t5
real	ft18t6
real	ft18t7
real	ft18t1e
real	ft18t2e
real	ft18t3e
real	ft18t4e
real	ft18t5e
real	ft18t6e
real	ft18t7e
real	ft19t1
real	ft19t2
real	ft19t3
real	ft19t4
real	ft19t5
real	ft19t6
real	ft19t7
real	ft19t1e
real	ft19t2e
real	ft19t3e
real	ft19t4e
real	ft19t5e
real	ft19t6e
real	ft19t7e
real	ft20t1
real	ft20t2
real	ft20t3
real	ft20t4
real	ft20t5
real	ft20t6
real	ft20t7
real	ft20t1e
real	ft20t2e
real	ft20t3e

real	ft20t4e
real	ft20t5e
real	ft20t6e
real	ft20t7e
real	ft21t1
real	ft21t2
real	ft21t3
real	ft21t4
real	ft21t5
real	ft21t6
real	ft21t7
real	ft21t1e
real	ft21t2e
real	ft21t3e
real	ft21t4e
real	ft21t5e
real	ft21t6e
real	ft21t7e
real	ft22t1
real	ft22t2
real	ft22t3
real	ft22t4
real	ft22t5
real	ft22t6
real	ft22t7
real	ft22t1e
real	ft22t2e
real	ft22t3e
real	ft22t4e
real	ft22t5e
real	ft22t6e
real	ft22t7e
real	ft23t1
real	ft23t2
real	ft23t3
real	ft23t4
real	ft23t5
real	ft23t6
real	ft23t7
real	ft23t1e
real	ft23t2e
real	ft23t3e
real	ft23t4e
real	ft23t5e
real	ft23t6e
real	ft23t7e
real	ft24t1
real	ft24t2
real	ft24t3
real	ft24t4
real	ft24t5
real	ft24t6
real	ft24t7
real	ft24t1e
real	ft24t2e
real	ft24t3e
real	ft24t4e

```

real          ft24t5e
real          ft24t6e
real          ft24t7e
real kaisq
real bestkaisq
real bf8
real bf9
real bf10
real bf11
real bf12
real bf13
real bf14
real bf15
real bf16
real bf17
real bf18
real bf19
real bf20
real bf21
real bf22
real bf23
real bf24
real a8 aa8 a9 aa9 a10 aa10 a11 aa11 a12 aa12 a13 aa13 a14 aa14
real b8 b9 b10 b11 b12 b13 b14 b15 b16 b17 b18 b19
real b20 b21 b22 b23 b24
real z1 z2 z3 z4 z5 z6 z7 f6 f26
real a15 aa15 a16 aa16 a17 aa17 a18 aa18 a19 aa19 a20 aa20
real a21 aa21 a22 aa22 a23 aa23 a24 aa24

* DIRK

integer i15 i23
character inputarg*128

CALL getarg(1,inputarg)
read(inputarg,*) i15

CALL getarg(2,inputarg)
read(inputarg,*) i23

goff1 = 0.45011191
goff2 =      0.581357776
goff3 =      0.710329562
goff4 =      0.826254485
goff5 =      0.918588976
goff6 =      0.978470252
goff7 =      0.998385086
gon1  =      0.093897695
gon2  =      0.118554369
gon3  =      0.196398598
gon4  =      0.378629871
gon5  =      0.645290442
gon6  =      0.893821899
gon7  =      0.991709633
z1 = 0.9052388602
z2 = 0.9334319587
z3 = 0.95693731265
z4 = 0.97551074985

```

```

z5 = 0.98895970015
z6 = 0.99714805251
z7 = 0.9997874594365

gnad1    =    0.281406262
gnad2    =    0.283440096
gnad3    =    0.298896559
gnad4    =    0.309714762
gnad5    =    0.320805841
gnad6    =    0.325732761
gnad7    =    0.342143946
ft8t1    =    0.05455
ft8t2    =    0.05178
ft8t3    =    0.05255
ft8t4    =    0.01784
ft8t5    =    0.02752
ft8t6    =    0.02554
ft8t7    =    0.00864
ft8t1e   =    0.01171
ft8t2e   =    0.01136
ft8t3e   =    0.01377
ft8t4e   =    0.01
ft8t5e   =    0.01
ft8t6e   =    0.01
ft8t7e   =    0.01
ft9t1    =    0.0615
ft9t2    =    0.05812
ft9t3    =    0.04719
ft9t4    =    0.03191
ft9t5    =    0.02722
ft9t6    =    0.00512
ft9t7    =    0.01355
ft9t1e   =    0.01671
ft9t2e   =    0.01235
ft9t3e   =    0.01
ft9t4e   =    0.01
ft9t5e   =    0.01
ft9t6e   =    0.01
ft9t7e   =    0.01
ft10t1   =    0.06781
ft10t2   =    0.05011
ft10t3   =    0.04724
ft10t4   =    0.0408
ft10t5   =    0.0239
ft10t6   =    0.01835
ft10t7   =    0.02391
ft10t1e  =    0.01606
ft10t2e  =    0.01294
ft10t3e  =    0.01184
ft10t4e  =    0.01233
ft10t5e  =    0.01013
ft10t6e  =    0.01
ft10t7e  =    0.0104
ft11t1   =    0.06249
ft11t2   =    0.05293
ft11t3   =    0.04992
ft11t4   =    0.03957

```

ft11t5	=	0.031
ft11t6	=	0.00575
ft11t7	=	-0.01026
ft11t1e	=	0.01341
ft11t2e	=	0.01711
ft11t3e	=	0.01107
ft11t4e	=	0.01247
ft11t5e	=	0.01
ft11t6e	=	0.0106
ft11t7e	=	0.01
ft12t1	=	0.08531
ft12t2	=	0.08566
ft12t3	=	0.0591
ft12t4	=	0.03005
ft12t5	=	0.02208
ft12t6	=	0.0117
ft12t7	=	6.49E-04
ft12t1e	=	0.01586
ft12t2e	=	0.01
ft12t3e	=	0.01
ft12t4e	=	0.01418
ft12t5e	=	0.01
ft12t6e	=	0.01039
ft12t7e	=	0.0106
ft13t1	=	0.12877
ft13t2	=	0.07773
ft13t3	=	0.06296
ft13t4	=	0.04703
ft13t5	=	0.03825
ft13t6	=	0.04439
ft13t7	=	0.0017
ft13t1e	=	0.02489
ft13t2e	=	0.0254
ft13t3e	=	0.02
ft13t4e	=	0.0155
ft13t5e	=	0.01
ft13t6e	=	0.01673
ft13t7e	=	0.01285
ft14t1	=	0.15352
ft14t2	=	0.11403
ft14t3	=	0.07306
ft14t4	=	0.06083
ft14t5	=	0.03403
ft14t6	=	0.04299
ft14t7	=	0.00683
ft14t1e	=	0.0119
ft14t2e	=	0.01368
ft14t3e	=	0.01192
ft14t4e	=	0.01
ft14t5e	=	0.01019
ft14t6e	=	0.01263
ft14t7e	=	0.01659
ft15t1	=	0.19465
ft15t2	=	0.1749
ft15t3	=	0.14289
ft15t4	=	0.09534
ft15t5	=	0.04554

ft15t6	=	0.03886
ft15t7	=	0.00352
ft15t1e	=	0.01573
ft15t2e	=	0.01096
ft15t3e	=	0.01
ft15t4e	=	0.01
ft15t5e	=	0.01565
ft15t6e	=	0.01
ft15t7e	=	0.0131
ft16t1	=	0.28671
ft16t2	=	0.24526
ft16t3	=	0.19745
ft16t4	=	0.12897
ft16t5	=	0.07304
ft16t6	=	0.03423
ft16t7	=	0.01909
ft16t1e	=	0.02292
ft16t2e	=	0.01424
ft16t3e	=	0.01478
ft16t4e	=	0.01189
ft16t5e	=	0.01
ft16t6e	=	0.01226
ft16t7e	=	0.01049
ft17t1	=	0.31006
ft17t2	=	0.27143
ft17t3	=	0.23816
ft17t4	=	0.17499
ft17t5	=	0.07795
ft17t6	=	0.0353
ft17t7	=	0.01897
ft17t1e	=	0.02074
ft17t2e	=	0.01444
ft17t3e	=	0.01598
ft17t4e	=	0.01853
ft17t5e	=	0.01963
ft17t6e	=	0.0198
ft17t7e	=	0.01164
ft18t1	=	0.30283
ft18t2	=	0.3016
ft18t3	=	0.25411
ft18t4	=	0.19389
ft18t5	=	0.11342
ft18t6	=	0.05735
ft18t7	=	0.02475
ft18t1e	=	0.02213
ft18t2e	=	0.01776
ft18t3e	=	0.01885
ft18t4e	=	0.0122
ft18t5e	=	0.01153
ft18t6e	=	0.01299
ft18t7e	=	0.01682
ft19t1	=	0.34606
ft19t2	=	0.27994
ft19t3	=	0.21297
ft19t4	=	0.1443
ft19t5	=	0.06939
ft19t6	=	0.04606

ft19t7	=	0.01298
ft19t1e	=	0.01175
ft19t2e	=	0.01249
ft19t3e	=	0.01
ft19t4e	=	0.01044
ft19t5e	=	0.01414
ft19t6e	=	0.01
ft19t7e	=	0.0107
ft20t1	=	0.39766
ft20t2	=	0.37867
ft20t3	=	0.32995
ft20t4	=	0.26239
ft20t5	=	0.14554
ft20t6	=	0.05559
ft20t7	=	0.00911
ft20t1e	=	0.01651
ft20t2e	=	0.01394
ft20t3e	=	0.01
ft20t4e	=	0.01
ft20t5e	=	0.01
ft20t6e	=	0.01
ft20t7e	=	0.01
ft21t1	=	0.25706
ft21t2	=	0.19786
ft21t3	=	0.17866
ft21t4	=	0.1297
ft21t5	=	0.0752
ft21t6	=	0.04328
ft21t7	=	0.00665
ft21t1e	=	0.01116
ft21t2e	=	0.01489
ft21t3e	=	0.01
ft21t4e	=	0.01
ft21t5e	=	0.01
ft21t6e	=	0.01
ft21t7e	=	0.01
ft22t1	=	0.15497
ft22t2	=	0.10269
ft22t3	=	0.07505
ft22t4	=	0.05987
ft22t5	=	0.02889
ft22t6	=	0.02218
ft22t7	=	0.0051
ft22t1e	=	0.01455
ft22t2e	=	0.01357
ft22t3e	=	0.01179
ft22t4e	=	0.01182
ft22t5e	=	0.01011
ft22t6e	=	0.01134
ft22t7e	=	0.01
ft23t1	=	0.11272
ft23t2	=	0.10263
ft23t3	=	0.08713
ft23t4	=	0.07678
ft23t5	=	0.04061
ft23t6	=	0.01599
ft23t7	=	0.01091

```

ft23t1e = 0.01183
ft23t2e = 0.01
ft23t3e = 0.01055
ft23t4e = 0.01
ft23t5e = 0.01
ft23t6e = 0.01081
ft23t7e = 0.01
ft24t1 = 0.10148
ft24t2 = 0.09025
ft24t3 = 0.05784
ft24t4 = 0.05364
ft24t5 = 0.029
ft24t6 = 0.01416
ft24t7 = 4.40E-04
ft24t1e = 0.01731
ft24t2e = 0.01521
ft24t3e = 0.01928
ft24t4e = 0.01275
ft24t5e = 0.01
ft24t6e = 0.01139
ft24t7e = 0.01256

```

```
bestkaisq = 100000000
```

S0 = 1.32

Slab = 0.9625275

```

f6 = 0.00
f7 = 0.00
f25 = 0.00
f26 = 0.00

```

```

      f8 = 0.00
do while (f8 <= 0.00)
      f9 = 0.00
do while (f9 <= 0.00)
      f10 = 0.00
      a8 = f7+f9
      b8 = f6+f10
      aa8 = 1-f6-f7-f8-f9-f10
sa8 = (S0-Slab*(gon1*f8+goff1*a8+z1*b8+aa8)-gnad1)/S0
sb8 = (S0-Slab*(gon2*f8+goff2*a8+z2*b8+aa8)-gnad2)/S0
sc8 = (S0-Slab*(gon3*f8+goff3*a8+z3*b8+aa8)-gnad3)/S0
sd8 = (S0-Slab*(gon4*f8+goff4*a8+z4*b8+aa8)-gnad4)/S0
se8 = (S0-Slab*(gon5*f8+goff5*a8+z5*b8+aa8)-gnad5)/S0
sg8 = (S0-Slab*(gon6*f8+goff6*a8+z6*b8+aa8)-gnad6)/S0
sh8 = (S0-Slab*(gon7*f8+goff7*a8+z7*b8+aa8)-gnad7)/S0
      xa8 = (ft8t1-sa8)**2/ft8t1e**2
      xb8 = (ft8t2-sb8)**2/ft8t2e**2
      xc8 = (ft8t3-sc8)**2/ft8t3e**2
      xd8 = (ft8t4-sd8)**2/ft8t4e**2
      xe8 = (ft8t5-se8)**2/ft8t5e**2
      xg8 = (ft8t6-sg8)**2/ft8t6e**2
      xh8 = (ft8t7-sh8)**2/ft8t7e**2

```

```

do while (f10 <= 0.00)
    a9 = f8+f10
    b9 = f7+f11
    aa9 = 1-f7-f8-f9-f10-f11
    sa9 = (S0-Slab*(gon1*f9+goff1*a9+b9*z1+aa9)-gnad1)/S0
    sb9 = (S0-Slab*(gon2*f9+goff2*a9+b9*z2+aa9)-gnad2)/S0
    sc9 = (S0-Slab*(gon3*f9+goff3*a9+b9*z3+aa9)-gnad3)/S0
    sd9 = (S0-Slab*(gon4*f9+goff4*a9+b9*z4+aa9)-gnad4)/S0
    se9 = (S0-Slab*(gon5*f9+goff5*a9+b9*z5+aa9)-gnad5)/S0
    sg9 = (S0-Slab*(gon6*f9+goff6*a9+b9*z6+aa9)-gnad6)/S0
    sh9 = (S0-Slab*(gon7*f9+goff7*a9+b9*z7+aa9)-gnad7)/S0
    xa9 = (ft9t1-sa9)**2/ft9t1e**2
    xb9 = (ft9t2-sb9)**2/ft9t2e**2
    xc9 = (ft9t3-sc9)**2/ft9t3e**2
    xd9 = (ft9t4-sd9)**2/ft9t4e**2
    xe9 = (ft9t5-se9)**2/ft9t5e**2
    xg9 = (ft9t6-sg9)**2/ft9t6e**2
    xh9 = (ft9t7-sh9)**2/ft9t7e**2
    f11 = 0.00
do while (f11 <= 0.00)
    a10 = f9+f11
    b10 = f8+f12
    aa10 = 1-f8-f9-f10-f11-f12
    sa10 = (S0-Slab*(gon1*f10+goff1*a10+z1*b10+aa10)-gnad1)/S0
    sb10 = (S0-Slab*(gon2*f10+goff2*a10+z2*b10+aa10)-gnad2)/S0
    sc10 = (S0-Slab*(gon3*f10+goff3*a10+z3*b10+aa10)-gnad3)/S0
    sd10 = (S0-Slab*(gon4*f10+goff4*a10+z4*b10+aa10)-gnad4)/S0
    se10 = (S0-Slab*(gon5*f10+goff5*a10+z5*b10+aa10)-gnad5)/S0
    sg10 = (S0-Slab*(gon6*f10+goff6*a10+z6*b10+aa10)-gnad6)/S0
    sh10 = (S0-Slab*(gon7*f10+goff7*a10+z7*b10+aa10)-gnad7)/S0
    xa10 = (ft10t1-sa10)**2/ft10t1e**2
    xb10 = (ft10t2-sb10)**2/ft10t2e**2
    xc10 = (ft10t3-sc10)**2/ft10t3e**2
    xd10 = (ft10t4-sd10)**2/ft10t4e**2
    xe10 = (ft10t5-se10)**2/ft10t5e**2
    xg10 = (ft10t6-sg10)**2/ft10t6e**2
    xh10 = (ft10t7-sh10)**2/ft10t7e**2
    f12 = 0.00
do while (f12 <= 0.00)
    a11 = f10+f12
    b11 = f9+f13
    aa11 = 1-f9-f10-f11-f12-f13
    sa11 = (S0-Slab*(gon1*f11+goff1*a11+z1*b11+aa11)-gnad1)/S0
    sb11 = (S0-Slab*(gon2*f11+goff2*a11+z2*b11+aa11)-gnad2)/S0
    sc11 = (S0-Slab*(gon3*f11+goff3*a11+z3*b11+aa11)-gnad3)/S0
    sd11 = (S0-Slab*(gon4*f11+goff4*a11+z4*b11+aa11)-gnad4)/S0
    se11 = (S0-Slab*(gon5*f11+goff5*a11+z5*b11+aa11)-gnad5)/S0
    sg11 = (S0-Slab*(gon6*f11+goff6*a11+z6*b11+aa11)-gnad6)/S0
    sh11 = (S0-Slab*(gon7*f11+goff7*a11+z7*b11+aa11)-gnad7)/S0
    xa11 = (ft11t1-sa11)**2/ft11t1e**2
    xb11 = (ft11t2-sb11)**2/ft11t2e**2
    xc11 = (ft11t3-sc11)**2/ft11t3e**2
    xd11 = (ft11t4-sd11)**2/ft11t4e**2
    xe11 = (ft11t5-se11)**2/ft11t5e**2
    xg11 = (ft11t6-sg11)**2/ft11t6e**2
    xh11 = (ft11t7-sh11)**2/ft11t7e**2
    f13 = 0.00

```



```

do while (f13 <= 0.05)
  a12 = f11+f13
  b12 = f10+f14
  aa12 = 1-f10-f11-f12-f13-f14
sa12 = (S0-Slab*(gon1*f12+goff1*a12+z1*b12+aa12)-gnad1)/S0
sb12 = (S0-Slab*(gon2*f12+goff2*a12+z2*b12+aa12)-gnad2)/S0
sc12 = (S0-Slab*(gon3*f12+goff3*a12+z3*b12+aa12)-gnad3)/S0
sd12 = (S0-Slab*(gon4*f12+goff4*a12+z4*b12+aa12)-gnad4)/S0
se12 = (S0-Slab*(gon5*f12+goff5*a12+z5*b12+aa12)-gnad5)/S0
sg12 = (S0-Slab*(gon6*f12+goff6*a12+z6*b12+aa12)-gnad6)/S0
sh12 = (S0-Slab*(gon7*f12+goff7*a12+z7*b12+aa12)-gnad7)/S0
  xa12 = (ft12t1-sa12)**2/ft12t1e**2
  xb12 = (ft12t2-sb12)**2/ft12t2e**2
  xc12 = (ft12t3-sc12)**2/ft12t3e**2
  xd12 = (ft12t4-sd12)**2/ft12t4e**2
  xe12 = (ft12t5-se12)**2/ft12t5e**2
  xg12 = (ft12t6-sg12)**2/ft12t6e**2
  xh12 = (ft12t7-sh12)**2/ft12t7e**2
  f14 = 0.00
do while (f14 <= 0.09)
  a13 = f12+f14
  b13 = f11+f15
  aa13 = 1-f11-f12-f13-f14-f15
sa13 = (S0-Slab*(gon1*f13+goff1*a13+z1*b13+aa13)-gnad1)/S0
sb13 = (S0-Slab*(gon2*f13+goff2*a13+z2*b13+aa13)-gnad2)/S0
sc13 = (S0-Slab*(gon3*f13+goff3*a13+z3*b13+aa13)-gnad3)/S0
sd13 = (S0-Slab*(gon4*f13+goff4*a13+z4*b13+aa13)-gnad4)/S0
se13 = (S0-Slab*(gon5*f13+goff5*a13+z5*b13+aa13)-gnad5)/S0
sg13 = (S0-Slab*(gon6*f13+goff6*a13+z6*b13+aa13)-gnad6)/S0
sh13 = (S0-Slab*(gon7*f13+goff7*a13+z7*b13+aa13)-gnad7)/S0
  xa13 = (ft13t1-sa13)**2/ft13t1e**2
  xb13 = (ft13t2-sb13)**2/ft13t2e**2
  xc13 = (ft13t3-sc13)**2/ft13t3e**2
  xd13 = (ft13t4-sd13)**2/ft13t4e**2
  xe13 = (ft13t5-se13)**2/ft13t5e**2
  xg13 = (ft13t6-sg13)**2/ft13t6e**2
  xh13 = (ft13t7-sh13)**2/ft13t7e**2
  f15 = 0.00
*
* Dirk
  f15 = 0.00+0.01*i15
*
do while (f15 <= 0.17)
  a14 = f13+f15
  b14 = f12+f16
  aa14 = 1-f12-f13-f14-f15-f16
sa14 = (S0-Slab*(gon1*f14+goff1*a14+z1*b14+aa14)-gnad1)/S0
sb14 = (S0-Slab*(gon2*f14+goff2*a14+z2*b14+aa14)-gnad2)/S0
sc14 = (S0-Slab*(gon3*f14+goff3*a14+z3*b14+aa14)-gnad3)/S0
sd14 = (S0-Slab*(gon4*f14+goff4*a14+z4*b14+aa14)-gnad4)/S0
se14 = (S0-Slab*(gon5*f14+goff5*a14+z5*b14+aa14)-gnad5)/S0
sg14 = (S0-Slab*(gon6*f14+goff6*a14+z6*b14+aa14)-gnad6)/S0
sh14 = (S0-Slab*(gon7*f14+goff7*a14+z7*b14+aa14)-gnad7)/S0
  xa14 = (ft14t1-sa14)**2/ft14t1e**2
  xb14 = (ft14t2-sb14)**2/ft14t2e**2
  xc14 = (ft14t3-sc14)**2/ft14t3e**2
  xd14 = (ft14t4-sd14)**2/ft14t4e**2
  xe14 = (ft14t5-se14)**2/ft14t5e**2

```

```

xg14 = (ft14t6-sg14)**2/ft14t6e**2
xh14 = (ft14t7-sh14)**2/ft14t7e**2

f16 = 0.04
do while (f16 <= 0.26)
  a15 = f14+f16
  b15 = f13+f17
  aa15 = 1-f13-f14-f15-f16-f17
  sa15 = (S0-Slab*(gon1*f15+goff1*a15+z1*b15+aa15)-gnad1)/S0
  sb15 = (S0-Slab*(gon2*f15+goff2*a15+z2*b15+aa15)-gnad2)/S0
  sc15 = (S0-Slab*(gon3*f15+goff3*a15+z3*b15+aa15)-gnad3)/S0
  sd15 = (S0-Slab*(gon4*f15+goff4*a15+z4*b15+aa15)-gnad4)/S0
  se15 = (S0-Slab*(gon5*f15+goff5*a15+z5*b15+aa15)-gnad5)/S0
  sg15 = (S0-Slab*(gon6*f15+goff6*a15+z6*b15+aa15)-gnad6)/S0
  sh15 = (S0-Slab*(gon7*f15+goff7*a15+z7*b15+aa15)-gnad7)/S0
  xa15 = (ft15t1-sa15)**2/ft15t1e**2
  xb15 = (ft15t2-sb15)**2/ft15t2e**2
  xc15 = (ft15t3-sc15)**2/ft15t3e**2
  xd15 = (ft15t4-sd15)**2/ft15t4e**2
  xe15 = (ft15t5-se15)**2/ft15t5e**2
  xg15 = (ft15t6-sg15)**2/ft15t6e**2
  xh15 = (ft15t7-sh15)**2/ft15t7e**2
  f17 = 0.03
do while (f17 <= 0.33)
  a16 = f15+f17
  b16 = f14+f18
  aa16 = 1-f14-f15-f16-f17-f18
  sa16 = (S0-Slab*(gon1*f16+goff1*a16+z1*b16+aa16)-gnad1)/S0
  sb16 = (S0-Slab*(gon2*f16+goff2*a16+z2*b16+aa16)-gnad2)/S0
  sc16 = (S0-Slab*(gon3*f16+goff3*a16+z3*b16+aa16)-gnad3)/S0
  sd16 = (S0-Slab*(gon4*f16+goff4*a16+z4*b16+aa16)-gnad4)/S0
  se16 = (S0-Slab*(gon5*f16+goff5*a16+z5*b16+aa16)-gnad5)/S0
  sg16 = (S0-Slab*(gon6*f16+goff6*a16+z6*b16+aa16)-gnad6)/S0
  sh16 = (S0-Slab*(gon7*f16+goff7*a16+z7*b16+aa16)-gnad7)/S0
  xa16 = (ft16t1-sa16)**2/ft16t1e**2
  xb16 = (ft16t2-sb16)**2/ft16t2e**2
  xc16 = (ft16t3-sc16)**2/ft16t3e**2
  xd16 = (ft16t4-sd16)**2/ft16t4e**2
  xe16 = (ft16t5-se16)**2/ft16t5e**2
  xg16 = (ft16t6-sg16)**2/ft16t6e**2
  xh16 = (ft16t7-sh16)**2/ft16t7e**2
  f18 = 0.05
do while (f18 <= 0.36)
  a17 = f16+f18
  b17 = f15+f19
  aa17 = 1-f15-f16-f17-f18-f19
  sa17 = (S0-Slab*(gon1*f17+goff1*a17+z1*b17+aa17)-gnad1)/S0
  sb17 = (S0-Slab*(gon2*f17+goff2*a17+z2*b17+aa17)-gnad2)/S0
  sc17 = (S0-Slab*(gon3*f17+goff3*a17+z3*b17+aa17)-gnad3)/S0
  sd17 = (S0-Slab*(gon4*f17+goff4*a17+z4*b17+aa17)-gnad4)/S0
  se17 = (S0-Slab*(gon5*f17+goff5*a17+z5*b17+aa17)-gnad5)/S0
  sg17 = (S0-Slab*(gon6*f17+goff6*a17+z6*b17+aa17)-gnad6)/S0
  sh17 = (S0-Slab*(gon7*f17+goff7*a17+z7*b17+aa17)-gnad7)/S0
  xa17 = (ft17t1-sa17)**2/ft17t1e**2
  xb17 = (ft17t2-sb17)**2/ft17t2e**2
  xc17 = (ft17t3-sc17)**2/ft17t3e**2

```

```

xd17 = (ft17t4-sd17)**2/ft17t4e**2
xe17 = (ft17t5-se17)**2/ft17t5e**2
xg17 = (ft17t6-sg17)**2/ft17t6e**2
xh17 = (ft17t7-sh17)**2/ft17t7e**2
f19 = 0.00
do while (f19 <= 0.33)
  a18 = f17+f19
  b18 = f16+f20
  aa18 = 1-f16-f17-f18-f19-f20
  sa18 = (S0-Slab*(gon1*f18+goff1*a18+z1*b18+aa18)-gnad1)/S0
  sb18 = (S0-Slab*(gon2*f18+goff2*a18+z2*b18+aa18)-gnad2)/S0
  sc18 = (S0-Slab*(gon3*f18+goff3*a18+z3*b18+aa18)-gnad3)/S0
  sd18 = (S0-Slab*(gon4*f18+goff4*a18+z4*b18+aa18)-gnad4)/S0
  se18 = (S0-Slab*(gon5*f18+goff5*a18+z5*b18+aa18)-gnad5)/S0
  sg18 = (S0-Slab*(gon6*f18+goff6*a18+z6*b18+aa18)-gnad6)/S0
  sh18 = (S0-Slab*(gon7*f18+goff7*a18+z7*b18+aa18)-gnad7)/S0
  xa18 = (ft18t1-sa18)**2/ft18t1e**2
  xb18 = (ft18t2-sb18)**2/ft18t2e**2
  xc18 = (ft18t3-sc18)**2/ft18t3e**2
  xd18 = (ft18t4-sd18)**2/ft18t4e**2
  xe18 = (ft18t5-se18)**2/ft18t5e**2
  xg18 = (ft18t6-sg18)**2/ft18t6e**2
  xh18 = (ft18t7-sh18)**2/ft18t7e**2
  f20 = 0.24
do while (f20 <= 0.49)
  a19 = f18+f20
  b19 = f17+f21
  aa19 = 1-f17-f18-f19-f20-f21
  sa19 = (S0-Slab*(gon1*f19+goff1*a19+z1*b19+aa19)-gnad1)/S0
  sb19 = (S0-Slab*(gon2*f19+goff2*a19+z2*b19+aa19)-gnad2)/S0
  sc19 = (S0-Slab*(gon3*f19+goff3*a19+z3*b19+aa19)-gnad3)/S0
  sd19 = (S0-Slab*(gon4*f19+goff4*a19+z4*b19+aa19)-gnad4)/S0
  se19 = (S0-Slab*(gon5*f19+goff5*a19+z5*b19+aa19)-gnad5)/S0
  sg19 = (S0-Slab*(gon6*f19+goff6*a19+z6*b19+aa19)-gnad6)/S0
  sh19 = (S0-Slab*(gon7*f19+goff7*a19+z7*b19+aa19)-gnad7)/S0
  xa19 = (ft19t1-sa19)**2/ft19t1e**2
  xb19 = (ft19t2-sb19)**2/ft19t2e**2
  xc19 = (ft19t3-sc19)**2/ft19t3e**2
  xd19 = (ft19t4-sd19)**2/ft19t4e**2
  xe19 = (ft19t5-se19)**2/ft19t5e**2
  xg19 = (ft19t6-sg19)**2/ft19t6e**2
  xh19 = (ft19t7-sh19)**2/ft19t7e**2
  f21 = 0.00
do while (f21 <= 0.24)
  a20 = f19+f21
  b20 = f18+f22
  aa20 = 1-f18-f19-f20-f21-f22
  sa20 = (S0-Slab*(gon1*f20+goff1*a20+z1*b20+aa20)-gnad1)/S0
  sb20 = (S0-Slab*(gon2*f20+goff2*a20+z2*b20+aa20)-gnad2)/S0
  sc20 = (S0-Slab*(gon3*f20+goff3*a20+z3*b20+aa20)-gnad3)/S0
  sd20 = (S0-Slab*(gon4*f20+goff4*a20+z4*b20+aa20)-gnad4)/S0
  se20 = (S0-Slab*(gon5*f20+goff5*a20+z5*b20+aa20)-gnad5)/S0
  sg20 = (S0-Slab*(gon6*f20+goff6*a20+z6*b20+aa20)-gnad6)/S0
  sh20 = (S0-Slab*(gon7*f20+goff7*a20+z7*b20+aa20)-gnad7)/S0
  xa20 = (ft20t1-sa20)**2/ft20t1e**2
  xb20 = (ft20t2-sb20)**2/ft20t2e**2
  xc20 = (ft20t3-sc20)**2/ft20t3e**2

```

```

        xd20 = (ft20t4-sd20)**2/ft20t4e**2
        xe20 = (ft20t5-se20)**2/ft20t5e**2
        xg20 = (ft20t6-sg20)**2/ft20t6e**2
        xh20 = (ft20t7-sh20)**2/ft20t7e**2
        f22 = 0.00
do while (f22 <= 0.08)
    a21 = f20+f22
    b21 = f19+f23
    aa21 = 1-f19-f20-f21-f22-f23
sa21 = (S0-Slab*(gon1*f21+goff1*a21+z1*b21+aa21)-gnad1)/S0
sb21 = (S0-Slab*(gon2*f21+goff2*a21+z2*b21+aa21)-gnad2)/S0
sc21 = (S0-Slab*(gon3*f21+goff3*a21+z3*b21+aa21)-gnad3)/S0
sd21 = (S0-Slab*(gon4*f21+goff4*a21+z4*b21+aa21)-gnad4)/S0
se21 = (S0-Slab*(gon5*f21+goff5*a21+z5*b21+aa21)-gnad5)/S0
sg21 = (S0-Slab*(gon6*f21+goff6*a21+z6*b21+aa21)-gnad6)/S0
sh21 = (S0-Slab*(gon7*f21+goff7*a21+z7*b21+aa21)-gnad7)/S0
    xa21 = (ft21t1-sa21)**2/ft21t1e**2
    xb21 = (ft21t2-sb21)**2/ft21t2e**2
    xc21 = (ft21t3-sc21)**2/ft21t3e**2
    xd21 = (ft21t4-sd21)**2/ft21t4e**2
    xe21 = (ft21t5-se21)**2/ft21t5e**2
    xg21 = (ft21t6-sg21)**2/ft21t6e**2
    xh21 = (ft21t7-sh21)**2/ft21t7e**2
*
    f23 = 0.01
*Dirk
        f23 = 0.00+0.01*i23
*
do while (f23 <= 0.08)
    a22 = f21+f23
    b22 = f20+f24
    aa22 = 1-f20-f21-f22-f23-f24
sa22 = (S0-Slab*(gon1*f22+goff1*a22+z1*b22+aa22)-gnad1)/S0
sb22 = (S0-Slab*(gon2*f22+goff2*a22+z2*b22+aa22)-gnad2)/S0
sc22 = (S0-Slab*(gon3*f22+goff3*a22+z3*b22+aa22)-gnad3)/S0
sd22 = (S0-Slab*(gon4*f22+goff4*a22+z4*b22+aa22)-gnad4)/S0
se22 = (S0-Slab*(gon5*f22+goff5*a22+z5*b22+aa22)-gnad5)/S0
sg22 = (S0-Slab*(gon6*f22+goff6*a22+z6*b22+aa22)-gnad6)/S0
sh22 = (S0-Slab*(gon7*f22+goff7*a22+z7*b22+aa22)-gnad7)/S0
    xa22 = (ft22t1-sa22)**2/ft22t1e**2
    xb22 = (ft22t2-sb22)**2/ft22t2e**2
    xc22 = (ft22t3-sc22)**2/ft22t3e**2
    xd22 = (ft22t4-sd22)**2/ft22t4e**2
    xe22 = (ft22t5-se22)**2/ft22t5e**2
    xg22 = (ft22t6-sg22)**2/ft22t6e**2
    xh22 = (ft22t7-sh22)**2/ft22t7e**2
        f24 = 0.00
do while (f24 <= 0.04)
    a23 = f22+f24
    b23 = f21+f25
    aa23 = 1-f21-f22-f23-f24-f25
sa23 = (S0-Slab*(gon1*f23+goff1*a23+z1*b23+aa23)-gnad1)/S0
sb23 = (S0-Slab*(gon2*f23+goff2*a23+z2*b23+aa23)-gnad2)/S0
sc23 = (S0-Slab*(gon3*f23+goff3*a23+z3*b23+aa23)-gnad3)/S0
sd23 = (S0-Slab*(gon4*f23+goff4*a23+z4*b23+aa23)-gnad4)/S0
se23 = (S0-Slab*(gon5*f23+goff5*a23+z5*b23+aa23)-gnad5)/S0
sg23 = (S0-Slab*(gon6*f23+goff6*a23+z6*b23+aa23)-gnad6)/S0
sh23 = (S0-Slab*(gon7*f23+goff7*a23+z7*b23+aa23)-gnad7)/S0

```

```

a24 = f23+f25
b24 = f22+f26
aa24 = 1-f22-f23-f24-f25-f26
sa24 = (S0-Slab*(gon1*f24+goff1*a24+z1*b24+aa24)-gnad1)/S0
sb24 = (S0-Slab*(gon2*f24+goff2*a24+z2*b24+aa24)-gnad2)/S0
sc24 = (S0-Slab*(gon3*f24+goff3*a24+z3*b24+aa24)-gnad3)/S0
sd24 = (S0-Slab*(gon4*f24+goff4*a24+z4*b24+aa24)-gnad4)/S0
se24 = (S0-Slab*(gon5*f24+goff5*a24+z5*b24+aa24)-gnad5)/S0
sg24 = (S0-Slab*(gon6*f24+goff6*a24+z6*b24+aa24)-gnad6)/S0
sh24 = (S0-Slab*(gon7*f24+goff7*a24+z7*b24+aa24)-gnad7)/S0

xa23 = (ft23t1-sa23)**2/ft23t1e**2
xb23 = (ft23t2-sb23)**2/ft23t2e**2
xc23 = (ft23t3-sc23)**2/ft23t3e**2
xd23 = (ft23t4-sd23)**2/ft23t4e**2
xe23 = (ft23t5-se23)**2/ft23t5e**2
xg23 = (ft23t6-sg23)**2/ft23t6e**2
xh23 = (ft23t7-sh23)**2/ft23t7e**2

xa24 = (ft24t1-sa24)**2/ft24t1e**2
xb24 = (ft24t2-sb24)**2/ft24t2e**2
xc24 = (ft24t3-sc24)**2/ft24t3e**2
xd24 = (ft24t4-sd24)**2/ft24t4e**2
xe24 = (ft24t5-se24)**2/ft24t5e**2
xg24 = (ft24t6-sg24)**2/ft24t6e**2
xh24 = (ft24t7-sh24)**2/ft24t7e**2

kaisq = xa8+xb8+xc8+xd8+xe8+xg8+xh8+
&      xa9+xb9+xc9+xd9+xe9+xg9+xh9+
&      xa10+xb10+xc10+xd10+xe10+xg10+xh10+
&      xa11+xb11+xc11+xd11+xe11+xg11+xh11+
&      xa12+xb12+xc12+xd12+xe12+xg12+xh12+
&      xa13+xb13+xc13+xd13+xe13+xg13+xh13+
&      xa14+xb14+xc14+xd14+xe14+xg14+xh14+
&      xa15+xb15+xc15+xd15+xe15+xg15+xh15+
&      xa16+xb16+xc16+xd16+xe16+xg16+xh16+
&      xa17+xb17+xc17+xd17+xe17+xg17+xh17+
&      xa18+xb18+xc18+xd18+xe18+xg18+xh18+
&      xa19+xb19+xc19+xd19+xe19+xg19+xh19+
&      xa20+xb20+xc20+xd20+xe20+xg20+xh20+
&      xa21+xb21+xc21+xd21+xe21+xg21+xh21+
&      xa22+xb22+xc22+xd22+xe22+xg22+xh22+
&      xa23+xb23+xc23+xd23+xe23+xg23+xh23+
&      xa24+xb24+xc24+xd24+xe24+xg24+xh24

```

IF (kaisq < bestkaisq) THEN

```

bestkaisq = kaisq
bf8 = f8
bf9 = f9
bf10 = f10
bf11 = f11
bf12 = f12
bf13 = f13
bf14 = f14
bf15 = f15

```

```

        bf16 = f16
        bf17 = f17
        bf18 = f18
        bf19 = f19
        bf20 = f20
        bf21 = f21
        bf22 = f22
        bf23 = f23
        bf24 = f24

    endif
    f24 = f24 + 0.01
enddo

*Dirk
*      f23 = f23 + 0.01
*      enddo
*      f22 = f22 + 0.01
enddo
    f21 = f21 + 0.01
enddo
    f20 = f20 + 0.01
enddo
    f19 = f19 + 0.01
enddo
    f18 = f18 + 0.01
enddo
    f17 = f17 + 0.01
enddo

    f16 = f16 + 0.01
enddo

*Dirk
*      f15 = f15 + 0.01
*      enddo
*      f14 = f14 + 0.01
enddo

    f13 = f13 + 0.01
enddo
    f12 = f12 + 0.01
enddo
    f11 = f11 + 0.01
enddo
    f10 = f10 + 0.01
enddo
    f9 = f9 + 0.01
enddo
    f8 = f8 + 0.01

enddo

* Dirk
*      OPEN(UNIT = 12, FILE = 'values5', STATUS = 'NEW')
*      WRITE(12,*) bestkaisq,
*      WRITE(*,*) bestkaisq,
&          bf8,

```

```
&          bf9,  
&          bf10, bf11, bf12, bf13, bf14, bf15,  
&          bf16, bf17, bf18, bf19, bf20, bf21, bf22,  
&          bf23, bf24  
end
```

Appendix VII. Chapter III Table of $(\Delta S/S_0)^{exp}$ values and $\gamma_{1tuv}^{lab}(\tau)$ or $\gamma_{1t_1t_2uv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON.

Table 16. Chapter III $(\Delta S/S_0)^{exp}$ and rms error.

Dephasing time (ms)	$(\Delta S/S_0)^{exp} + \sigma^{exp}$ in parentheses			
	(HFP-NC) ^{exp}	(HFP-P) ^{exp}	(HFP-A) ^{exp}	(HFP-AP) ^{exp}
2.2	0.019 (.013)	0.027 (.013)	0.012 (.023)	0.017 (.028)
8.2	0.056 (.020)	0.053 (.013)	0.077 (.036)	0.055 (.026)
16.2	0.071 (.020)	0.061 (.015)	0.107 (.021)	0.128 (.022)
24.2	0.077 (.023)	0.108 (.013)	0.202 (.026)	0.219 (.016)
32.2	0.086 (.017)	0.087 (.018)	0.244 (.025)	0.265 (.020)
40.2	0.093 (.018)	0.126 (.023)	0.306 (.031)	0.294 (.023)
48.2	0.089 (.028)	0.162 (.025)	0.325 (.060)	0.344 (.020)

Figure 49 illustrates the spin geometries and associated calculated $\gamma_{1tuv}^{lab}(\tau)$ for specific t , u , and v in the fully constrained model or calculated $\gamma_{1t_1t_2uv}^{lab}(\tau)$ for specific t_1 , t_2 , u , and v in the unconstrained model. In each spin geometry schematic, N or C respectively denote a ^{15}N nucleus or a ^{13}CO nucleus included in the simulation, and Y denotes either X registry or a ^{13}CO HFP. Each arrow denotes ^{13}CO - ^{15}N dipolar coupling considered in the simulation. For unconstrained model geometries, the top/middle strand registry is at the top of the schematic and the middle/bottom strand registry is at the bottom of the schematic. When more than one spin geometries are shown, each $\gamma(\tau)$ is the average for the displayed geometries.

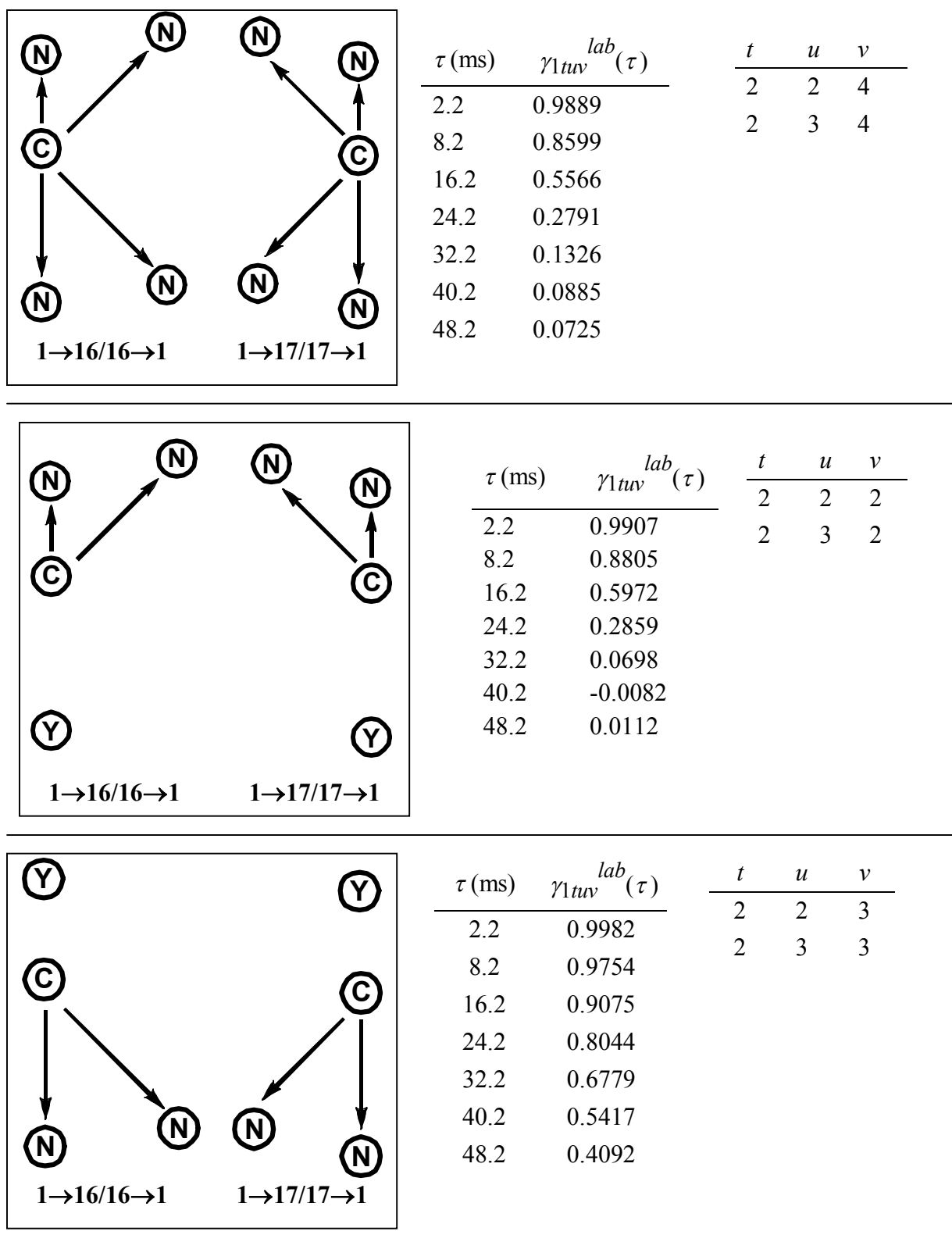


Figure 49. Chapter III spin geometries and simulated data.

Figure 49 (cont'd).

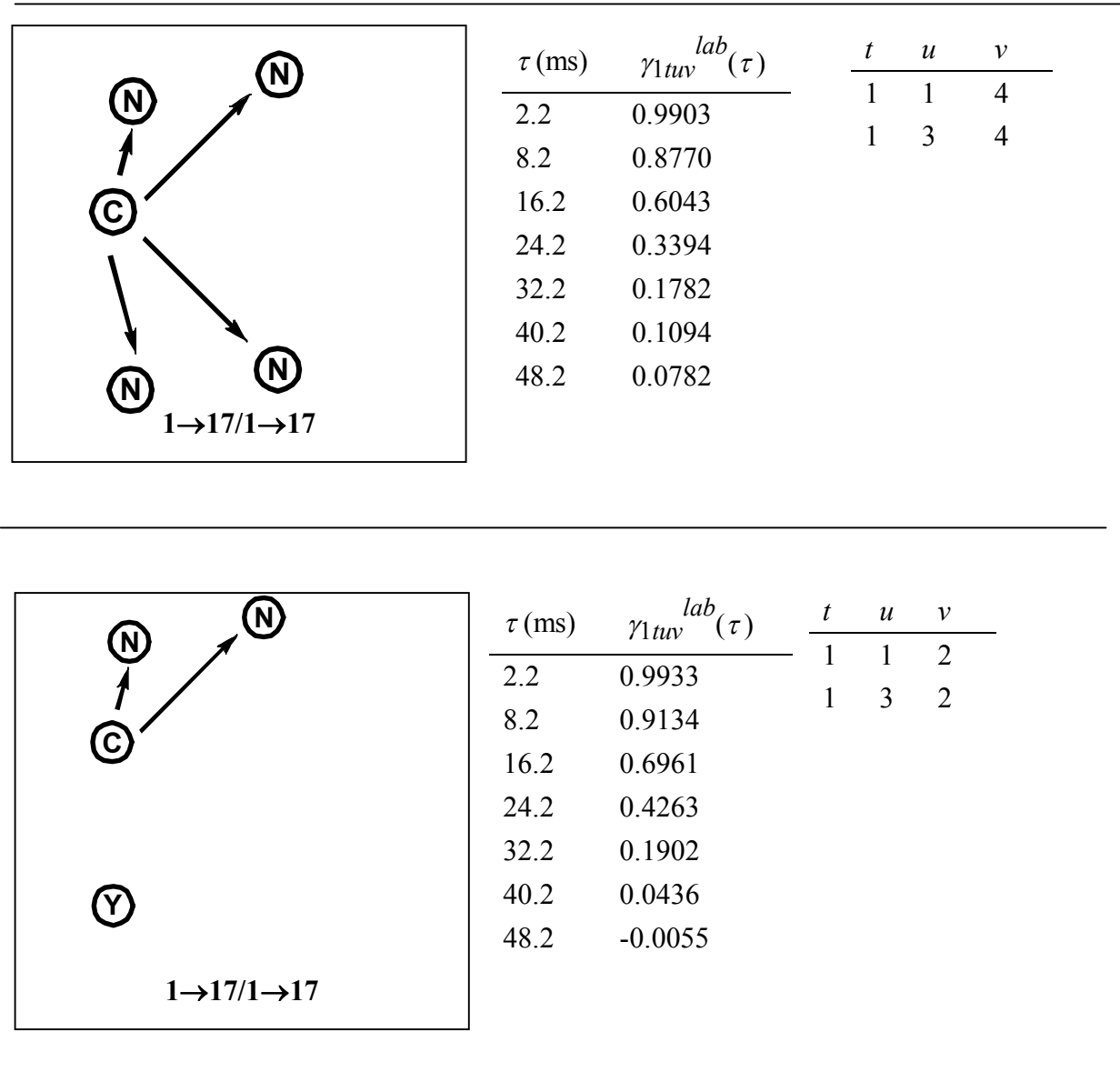
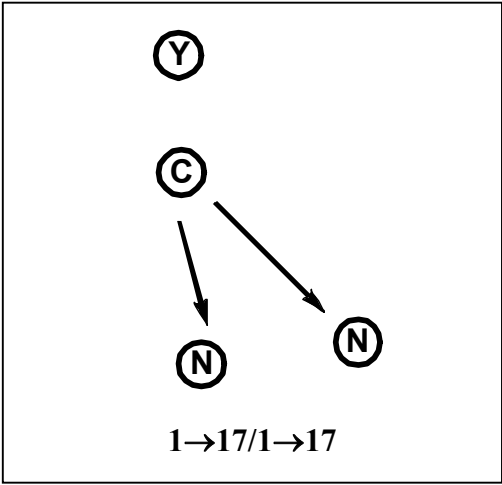
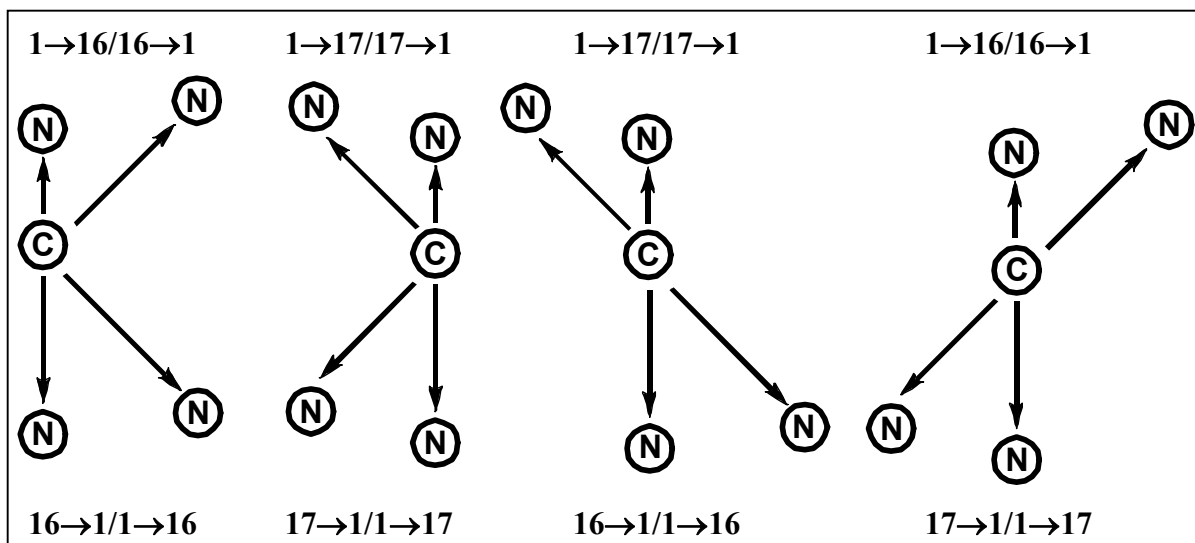


Figure 49 (cont'd).



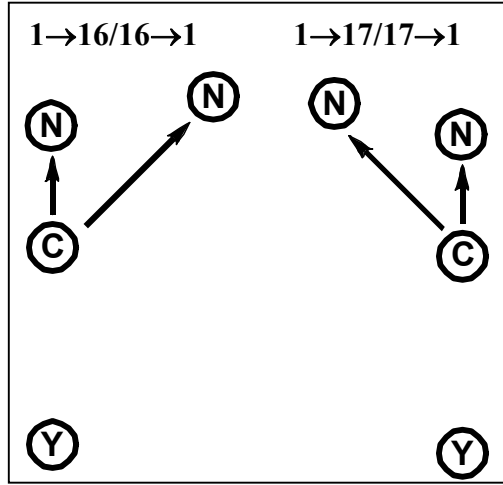
τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t	u	v
2.2	0.9979	1	1	3
8.2	0.9717	1	3	3
16.2	0.8941			
24.2	0.7773			
32.2	0.6364			
40.2	0.4883			
48.2	0.3488			

Figure 49 (cont'd).

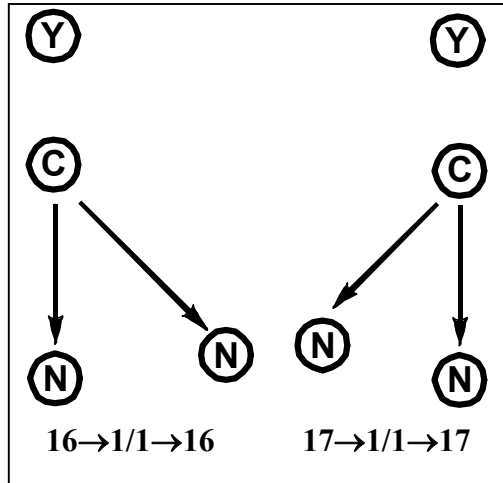


τ (ms)	$\gamma_{1tuv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9889	2	2	2	4
8.2	0.8605	2	2	3	4
16.2	0.5582				
24.2	0.2806				
32.2	0.1318				
40.2	0.0847				
48.2	0.0681				

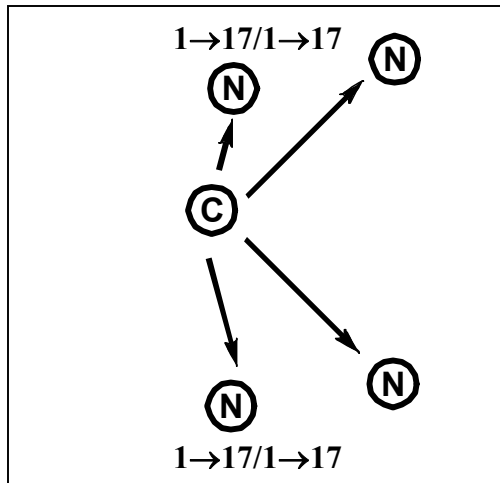
Figure 49 (cont'd).



τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9907	2	2	2	2
8.2	0.8805	2	2	3	2
16.2	0.5972	2	1	2	2
24.2	0.2859	2	1	2	4
32.2	0.0698	2	3	2	2
40.2	-0.0082	2	3	2	4
48.2	0.0112	2	3	3	2
		2	3	3	4

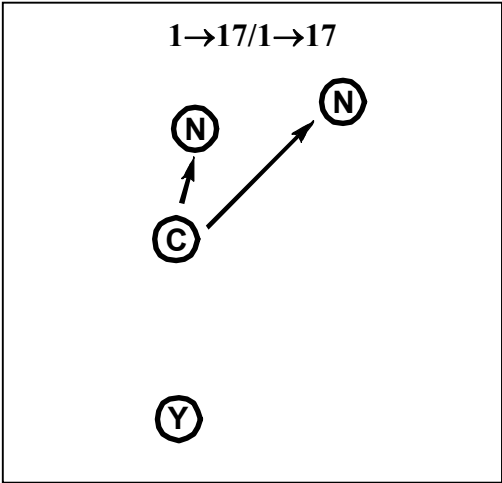


τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9982	2	2	2	3
8.2	0.9754	2	2	3	3
16.2	0.9075	1	2	2	3
24.2	0.8044	1	2	2	4
32.2	0.6779	3	2	2	3
40.2	0.5417	3	2	2	4
48.2	0.4092	3	2	3	3
		3	2	3	4

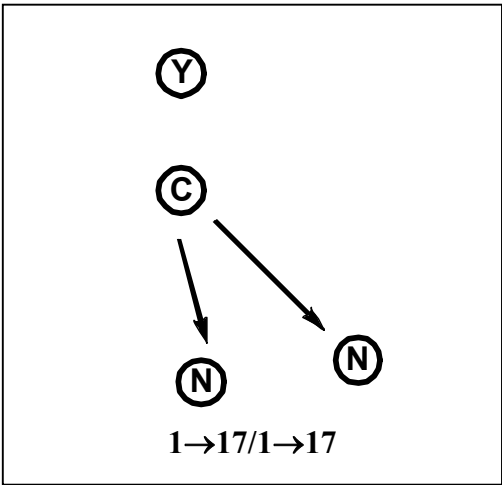


τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9903	1	1	1	4
8.2	0.8770	1	1	3	4
16.2	0.6043				
24.2	0.3394				
32.2	0.1782				
40.2	0.1094				
48.2	0.0782				

Figure 49 (cont'd).

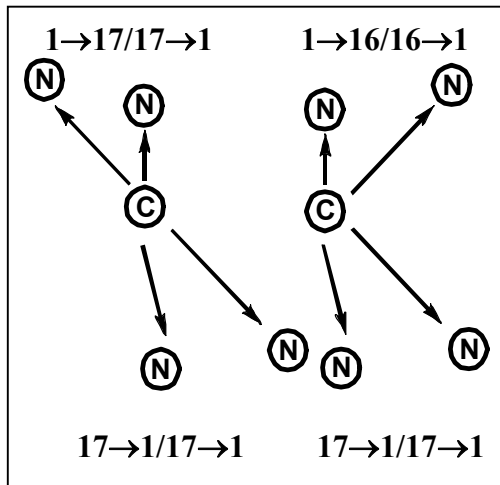


τ (ms)	$\gamma_{1tuv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9933	1	1	1	2
8.2	0.9134	1	1	3	2
16.2	0.6961	1	2	1	2
24.2	0.4263	1	2	1	4
32.2	0.1902	1	3	1	2
40.2	0.0436	1	3	1	4
48.2	-0.0055	1	3	3	2
		1	3	3	4

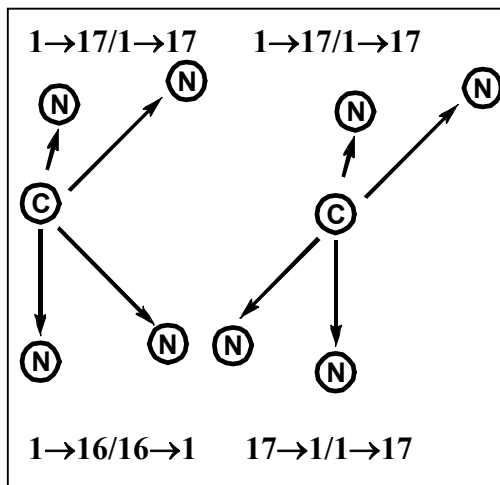


τ (ms)	$\gamma_{1tuv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9979	1	1	1	3
8.2	0.9717	1	1	3	3
16.2	0.8941	2	1	1	3
24.2	0.7773	2	1	1	4
32.2	0.6364	3	1	1	3
40.2	0.4883	3	1	1	4
48.2	0.3488	3	1	3	3
		3	1	3	4

Figure 49 (cont'd).



τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9903	2	1	3	4
8.2	0.8766				
16.2	0.5998				
24.2	0.3263				
32.2	0.1591				
40.2	0.0937				
48.2	0.0712				



τ (ms)	$\gamma_{1uv}^{lab}(\tau)$	t_1	t_2	u	v
2.2	0.9904	1	2	3	4
8.2	0.8782				
16.2	0.6061				
24.2	0.3389				
32.2	0.1753				
40.2	0.1079				
48.2	0.0805				

Appendix VIII. Five Registry Fittings

The 5 registry fittings were performed identical to the 3 registry fittings (Chapter IV). The natural abundance calculations were identical, but the number of antiparallel registries that were considered to contribute to the rdd in a given sample was increased from 3 to 5.

$$S_{lu}^{sim}(\tau) = \left\{ \left[0.99 \times J \times 0.0037 \times \gamma_{lN}^{na}(\tau) \right] + \left[0.98 \times K \times 0.011 \times \gamma_{lC}^{na}(\tau) \right] \right\} + \left\{ 0.33 - (K \times 0.011) \right\} + \left\{ \left[0.99 - (J \times 0.0037) \right] \times \left[f_X + \sum_{t=u-2}^{u+2} f_t \gamma_{ltu}^{lab}(\tau) \right] \right\} \quad (63)$$

The χ_u^2 analysis was modified from below.

$$\chi_u^2 = \sum_{m=1}^7 \left\{ \frac{\left[\frac{\Delta S(f_{t=u-2}, f_{t=u-1}, f_{t=u}, f_{t=u+1}, f_{t=u+2})}{S_0} \right]_{um}^{sim} - \left[\frac{\Delta S}{S_0} \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (64)$$

The global fitting was modified below.

$$\chi^2(f_{8-24}) = \sum_{u=8}^{24} \sum_{m=1}^7 \left\{ \frac{\left[\Delta S(f_{t=u-2}, f_{t=u-1}, f_{t=u}, f_{t=u+1}, f_{t=u+2}) / S_0 \right]_{um}^{sim} - \left[\Delta S / S_0 \right]_{um}^{exp}}{\sigma_{um}^{exp}} \right\}^2 \quad (65)$$

Appendix IX. Chapter IV Unconstrained Fitting

Similar to Chapter III, an alternate “unconstrained” fitting model was also considered for the mHFP data in Chapter IV. In Chapter IV, the unconstrained model had local mixing of registries $t = 8-24$ where different registries could be locally mixed in a single HFP oligomer. The overall goal was to approximate the best-fit f_t 's based on the $(\Delta S/S_0)^{exp}$ of many mHFP samples that were labeled differently using an unbiased data analysis method. Previous work has shown that “fully constrained” and “unconstrained” models generally yield similar fractional populations for REDOR experiments^{36,63}. Each sample was indexed by $u = \zeta + \eta - 1$ where ζ and η were the residue numbers of the ^{13}CO and ^{15}N labeled residues. For a particular “middle” HFP, the indices t_1 and t_2 describe the registries with the two adjacent strands. The adjacent strand of the t_1 registry is hydrogen bonded to the labeled ^{13}CO of the middle HFP while t_2 describes the adjacent strand that is not hydrogen bonded to the labeled ^{13}CO of the middle HFP. The simulated signals for the unconstrained model are described by **Eq (66)**.

$$S_{lu}^{sim}(\tau) = \left\{ \left[0.99 \times J \times 0.0037 \times \gamma_{lN}^{na}(\tau) \right] + \left[0.98 \times K \times 0.011 \times \gamma_{lC}^{na}(\tau) \right] \right\} + \{ 0.33 - (K \times 0.011) \} \\ + \left\{ \left[0.99 - (J \times 0.0037) \right] \times \left[\sum_{t_1=8}^{24} \sum_{t_2=8}^{24} f_{t_1} \times f_{t_2} \times \gamma_{lt_1t_2u}^{lab}(\tau) \right] \right\} \quad (66)$$

Similar to the fully constrained model, the $\gamma_{lt_1t_2u}^{lab}(\tau) = 1$ except when $t_1 = u - 1$, $t_1 = u$, $t_1 = u + 1$, $t_2 = u - 1$, $t_2 = u$, and/or $t_2 = u + 1$. In these latter cases, the $\gamma_{lt_1t_2u}^{lab}(\tau)$ were determined by SIMPSON calculations. Additionally, the spin geometry of an adjacent strand is denoted X

for t_1 when $t_1 \neq u - 1$, $t_1 \neq u$, or $t_1 \neq u + 1$ and is denoted X for t_2 when $t_2 \neq u - 1$, $t_2 \neq u$, or $t_2 \neq u + 1$. The iterative fitting for the unconstrained model is illustrated in **Figure 50** and the χ^2 analysis for each sample u data set is described by **Eq (50)**.

Table 17. Unconstrained model mHFP.

f_t	Best-fit f_t	χ^2
f_8	0.00	1.4
f_9	0.00	3.0
f_{10}	0.01	3.8
f_{11}	0.03	1.8
f_{12}	0.06	5.3
f_{13}	0.13	2.4
f_{14}	0.05	11.0
f_{15}	0.13	3.9
f_{16}	0.10	6.1
f_{17}	0.14	10.7
f_{18}	0.11	12.5
f_{19}	0.01	12.2
f_{20}	0.13	6.2
f_{21}	0.00	8.4
f_{22}	0.01	10.1
f_{23}	0.03	5.6
f_{24}	0.00	12.1
The $\sum_{t=8}^{24} f_t = 0.94$ and the $\chi^2_{\min} = 116$.		

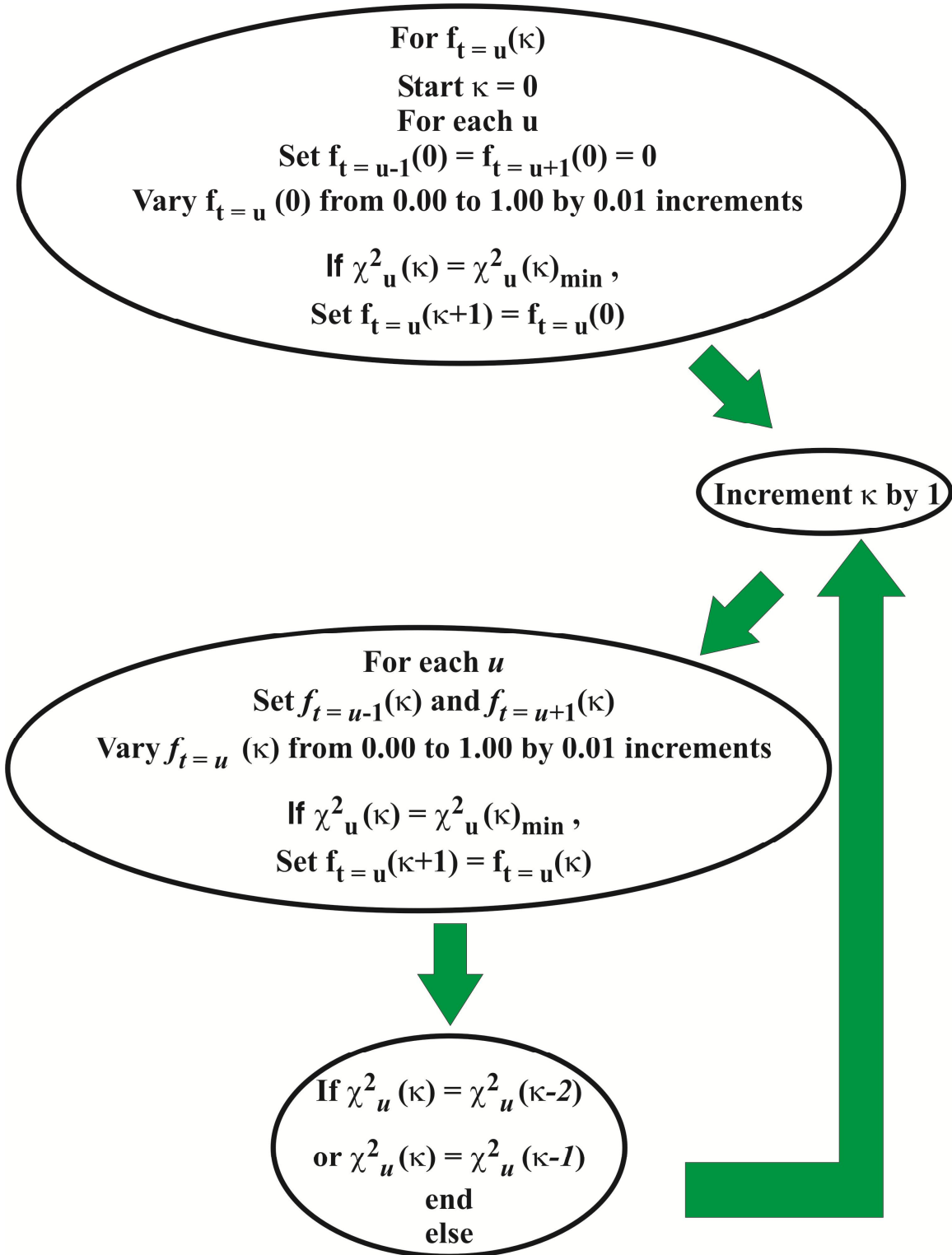


Figure 50. Flow chart for unconstrained iterative fitting. Each iteration is denoted by the variable κ , and the χ^2_u calculations are found in Chapter IV, Eq (50).

Appendix X. Freed Mutations

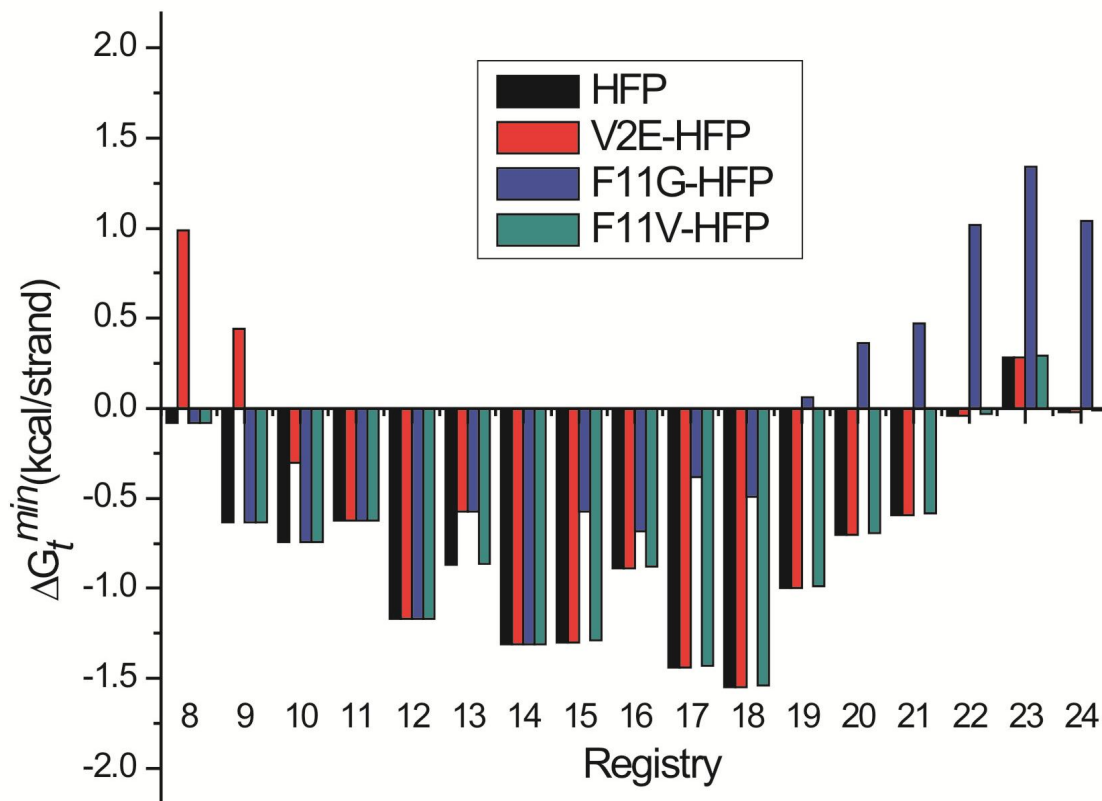


Figure 51. The ΔG_t^{\min} are plotted for registries $t = 8-24$ for HFP, V2E-HFP, F11G-HFP, and F11V-HFP. For each registry, the F11G-HFP ΔG_t^{\min} are greater than or equal to the HFP ΔG_t^{\min} which may contribute toward F11G-HFP's lower fusion activity. The F11V-HFP ΔG_t^{\min} is approximately equal to the HFP ΔG_t^{\min} relative to mHFP for each registry.

Appendix XI. Raw Data for mHFP and mV2E-HFP.

Table 18. mHFP $\Delta(S/S_0)^{\text{exp}}$.

u	HFP	$\Delta(S/S_0)^{\text{exp}}$ σ^{exp}						
		48.2 ms	40.2 ms	32.2 ms	24.2 ms	16.2 ms	8.2 ms	2.2 ms
8	A6CG3N	0.057	0.052	0.037	0.038	0.030	0.017	0.006
		0.016	0.012	0.011	0.009	0.006	0.006	0.007
9	L7CG3N	0.063	0.068	0.047	0.033	0.032	0.009	0.012
		0.018	0.012	0.011	0.009	0.009	0.006	0.006
10	F8CG3N	0.111	0.062	0.039	0.046	0.033	0.022	0.015
		0.021	0.017	0.019	0.016	0.012	0.007	0.010
11	L9CG3N	0.141	0.097	0.081	0.066	0.046	0.026	0.014
		0.022	0.017	0.011	0.012	0.009	0.006	0.005
12	L9CI4N	0.215	0.170	0.113	0.095	0.060	0.016	0.011
		0.023	0.016	0.011	0.012	0.009	0.009	0.009
13	L9CG5N	0.256	0.218	0.172	0.102	0.067	0.034	0.010
		0.025	0.016	0.014	0.015	0.012	0.008	0.005
14	L12CG3N	0.235	0.171	0.138	0.109	0.088	0.033	0.003
		0.021	0.011	0.014	0.013	0.012	0.009	0.006
15	L12CI4N	0.244	0.215	0.173	0.123	0.093	0.043	0.008
		0.019	0.019	0.014	0.011	0.012	0.008	0.008
16	L12CG5N	0.253	0.238	0.179	0.128	0.090	0.044	0.012
		0.015	0.015	0.011	0.008	0.010	0.009	0.007
17	L12CA6N	0.275	0.247	0.192	0.155	0.099	0.058	0.004
		0.021	0.016	0.011	0.013	0.007	0.010	0.009
18	L12CL7N	0.201	0.188	0.174	0.126	0.085	0.055	0.011
		0.021	0.020	0.012	0.010	0.011	0.011	0.007
19	F8CL12N	0.157	0.145	0.131	0.082	0.064	0.022	0.010
		0.013	0.013	0.011	0.008	0.006	0.005	0.007
20	F8CG13N	0.175	0.177	0.161	0.116	0.068	0.017	0.022
		0.015	0.024	0.012	0.015	0.009	0.012	0.012
21	F8CA14N	0.112	0.072	0.074	0.052	0.028	0.005	0.010
		0.016	0.016	0.013	0.013	0.011	0.010	0.009
22	F8CA15N	0.096	0.084	0.070	0.021	0.041	0.042	0.011
		0.013	0.016	0.011	0.012	0.009	0.009	0.009

Table 18 (cont'd).

23	F8CG16N	0.113	0.089	0.057	0.059	0.049	0.014	0.026
		0.019	0.019	0.016	0.018	0.011	0.014	0.010
24	L9CG16N	0.046	0.050	0.015	0.024	0.031	0.016	0.006
		0.014	0.014	0.019	0.008	0.006	0.007	0.005
28	F8CA21N	0.044	0.043	0.045	0.032	0.021	0.017	0.016
		0.021	0.017	0.013	0.013	0.011	0.010	0.009

Table 19. mV2E-HFP $\Delta(S/S_0)^{\text{exp}}$.

u	mV2E-HFP	$\Delta(S/S_0)^{\text{exp}}$						
		σ^{exp}						
		48.2 ms	40.2 ms	32.2 ms	24.2 ms	16.2 ms	8.2 ms	2.2 ms
8	A6CG3N	0.055	0.052	0.053	0.018	0.028	0.026	0.009
		0.012	0.010	0.014	0.010	0.010	0.010	0.010
9	L7CG3N	0.062	0.058	0.047	0.032	0.027	0.005	0.014
		0.017	0.012	0.010	0.010	0.010	0.010	0.010
10	F8CG3N	0.068	0.050	0.047	0.041	0.024	0.018	0.024
		0.016	0.013	0.012	0.012	0.010	0.010	0.010
11	L9CG3N	0.062	0.053	0.050	0.040	0.031	0.006	-0.010
		0.013	0.017	0.011	0.012	0.010	0.011	0.010
12	L9CI4N	0.085	0.086	0.059	0.030	0.022	0.012	0.001
		0.016	0.010	0.010	0.014	0.010	0.010	0.010
13	L9CG5N	0.129	0.078	0.063	0.047	0.038	0.044	0.002
		0.025	0.025	0.020	0.016	0.010	0.017	0.013
14	L12CG3N	0.154	0.114	0.073	0.061	0.034	0.043	0.007
		0.012	0.014	0.012	0.010	0.010	0.013	0.017
15	L12CI4N	0.195	0.175	0.143	0.095	0.046	0.039	0.004
		0.010	0.011	0.010	0.010	0.016	0.010	0.013
16	L12CG5N	0.287	0.245	0.197	0.129	0.073	0.034	0.019
		0.023	0.014	0.015	0.012	0.010	0.012	0.010
17	L12CA6N	0.310	0.271	0.238	0.175	0.078	0.035	0.019
		0.020	0.014	0.016	0.019	0.020	0.020	0.012
18	L12CL7N	0.303	0.302	0.254	0.194	0.113	0.057	0.025
		0.022	0.018	0.019	0.012	0.012	0.013	0.017

Table 19 (cont'd).

19	F8CL12N	0.346	0.280	0.213	0.144	0.069	0.046	0.013
		0.012	0.012	0.010	0.010	0.014	0.010	0.011
20	F8CG13N	0.398	0.379	0.330	0.262	0.146	0.056	0.009
		0.017	0.014	0.010	0.010	0.010	0.010	0.010
21	F8CA14N	0.257	0.198	0.179	0.130	0.075	0.043	0.007
		0.011	0.015	0.010	0.010	0.010	0.010	0.010
22	F8CA15N	0.155	0.103	0.075	0.060	0.029	0.022	0.005
		0.015	0.014	0.012	0.012	0.010	0.011	0.010
23	F8CG16N	0.113	0.103	0.087	0.077	0.041	0.016	0.011
		0.011	0.010	0.011	0.010	0.010	0.011	0.010
24	L9CG16N	0.101	0.090	0.058	0.054	0.029	0.014	0.000
		0.018	0.015	0.019	0.013	0.010	0.011	0.013

Appendix XII. Boltzmann Fraction Populations.

A Boltzmann distribution was used to calculate fractional populations of each registry by Eq. (67).

$$f_t = \frac{\exp \left(-\frac{\Delta G_t^{\min}}{kT} \right)}{\sum_{t=8}^{24} \exp \left(-\frac{\Delta G_t^{\min}}{kt} \right)} \quad (67)$$

Table 20. mHFP fully constrained model and Boltzmann distribution based f_t .

Registry (t)	mHFP	
	f_t (fully constrained)	f_t^{calc}
8	0.00	0.00
9	0.00	0.01
10	0.01	0.02
11	0.04	0.01
12	0.07	0.08
13	0.16	0.03
14	0.06	0.13
15	0.15	0.12
16	0.12	0.03
17	0.18	0.20
18	0.13	0.29
19	0.02	0.04
20	0.15	0.02
21	0.00	0.01
22	0.01	0.00
23	0.04	0.00
24	0.00	0.00

Appendix XIII. L9R Mutant Discussion

Previous experiments have demonstrated that V2E mutated gp41 expressed with wild type gp41 can dominantly inhibit membrane fusion, **Appendix III**, while L9R mutated gp41 expressed with wild type gp41 only stoichiometrically inhibits fusion²¹. Chapter IV of this dissertation demonstrated that mV2E-HFP forms a narrower distribution of registries that were shifted toward longer registries relative to mHFP. To date, there has not been any registry specific membrane insertion data collected for any HFP construct, but mV2E-HFP has a membrane surface location²⁶ and a predominant population of longer registries which is consistent with longer registries having a surface membrane location. The mV2E-HFP forms stable β sheet oligomers (Chapter IV) that are less fusion active than mHFP β sheet oligomers³¹. Since both mHFP and mV2E-HFP favorably fold into $t \sim 20$ registries, a reasonable hypothesis is that the transdominant inhibition may occur when V2E mutant gp41 forms β sheet oligomers with wild type gp41 that preferentially fold into $t \sim 20$ registries, and these registries have a membrane surface location that result in fusion inactive oligomers.

Similar to the V2E-HFP, the L9R-HFP has one hydrophobic residue mutated to a charged residue, but L9R-HFP did not dominantly inhibit membrane fusion. While this effect may have been due to amino acid type, I hypothesize that it's due to the sequential placement of the charged residue. For mV2E-HFP, the magnitude of the each ΔG_t^{\min} value was not correlated to the f_t 's (i.e. ΔG_t^{\min} does not accurately approximate each registry's free energy and other energetic contributions¹⁰⁷ must be considered), but registries with $f_t > 0.00$ generally had negative ΔG_t^{\min} values. It would seem that registries with negative ΔG_t^{\min} can stay bound to

the lipids while registries with positive ΔG_t^{\min} do not since they are typically not observed in membrane-associated HFP constructs. While mV2E-HFP has a partially inserted or membrane surface location, it stays bound to membranes and has a predominantly β sheet structure whereas lyophilized HFP without membranes forms a broad distribution of secondary structures³⁸. Both HFP and V2E-HFP had registries with negative and similar ΔG_t^{\min} values for most registries. Using similar analysis, L9R-HFP has predominantly positive ΔG_t^{\min} values and a minimum $\Delta G_t^{\min} = -0.62$ kcal/strand which is approximately equal to the $t = 11$ registry in HFP. In HFP, the $\Delta G_{11}^{\min} = -0.62$ kcal/strand and the $\Delta G_{20}^{\min} = -0.70$ kcal/strand. The magnitudes of these registries' insertion energies are similar which suggests that other factors, such as side chain packing, are more favorable for the $t = 20$ registry than the $t = 11$ registry. Since mL9R-HFP only has a couple registries with negative ΔG_t^{\min} and those registries are not formed in mHFP or mV2E-HFP, there are four hypotheses for mL9R-HFP structure: (1) mL9R-HFP will form longer registries like mV2E-HFP; (2) mL9R-HFP will form predominant $t = 11$ β sheet registries since this registry has the most negative ΔG_t^{\min} ; (3) mL9R-HFP will not preferentially fold into β sheets since the registries with favorable ΔG_t^{\min} may have unfavorable energetic contributions from other factors, such as sidechain packing; or (4) mL9R-HFP does not bind to the membrane. In mHFP, the most hydrophobic intrastrand sequence, L7-L12, likely plays a key role in stabilizing the membrane inserted structures. In mL9R-HFP, this region has a net positive contribution to $\Delta G_{n_t}^{\min}$ because $\Delta G_{Arg} = +2.58$ kcal/mol while

$2\Delta G_{Leu} + 2\Delta G_{Phe} + \Delta G_{Gly} = -1.00$ kcal/mol. This may not allow for stable binding/partial membrane insertion of β sheet oligomers. A pilot experiment was run to test whether mL9R-HFP forms longer registries like mV2E-HFP using mL9R-HFP with F8CG13N labeling. This sample was prepared identical to mHFP samples and the L9R-HFP induced aggregation among LUV's, but the LUV's stuck to the conical vial and precipitated out of solution making a clearer overall solution when compared to mHFP or mV2E-HFP, **Figure 54**. Due to difficulties acquiring signal above the noise level for REDOR experiments at dephasing times greater than 8 ms, a cross polarization spectrum was taken to observe the chemical shift distribution of the Phe-8 ^{13}CO to obtain secondary structure information, **Figure 55**. The reduced signal intensity was likely due to L9R-HFP forming a distribution of membrane bound and unbound structures where more L9R-HFP is likely unbound than observed for mHFP and mV2E-HFP samples. The NMR sample does not contain the unbound peptide. No data were collected to quantify the amount of unbound L9R-HFP. The mL9R-HFP with F8CG13N labeling had 174.6 ppm chemical shift and a 7.5 ppm line full-width at half maximum height that suggested the presence of a broad distribution of secondary structures since the chemical shifts spanned that of both α helical and β sheet secondary structures⁷⁷. This result was similar to the HFP-NC sample (Chapter III) where 175.9 ppm chemical shift was observed with lyphophilized F8C and A6NL7N labeled HFPs with a 7.0 ppm line full-width at half maximum height (Chapter III). Conversely, the HFP-F8CG13N sample had a 173.1 ppm chemical shift and a 3.5 ppm line width at half maximum height which strongly indicated β sheet secondary structure. These combined results suggest that mL9R-HFP does not favorably fold into membrane bound/partially inserted β sheet oligomers to the extent of mHFP and mV2E-HFP. These data are also consistent with the idea that L9R mutated gp41 does not form an oligomeric structure with wild type gp41 since mL9R-HFP does not preferentially

form β sheet oligomers. These results combined with the mV2E-HFP results may explain why V2E mutated gp41 transdominantly inhibits wild type gp41 while L9R mutated gp41 only stoichiometrically inhibits wild type gp41.

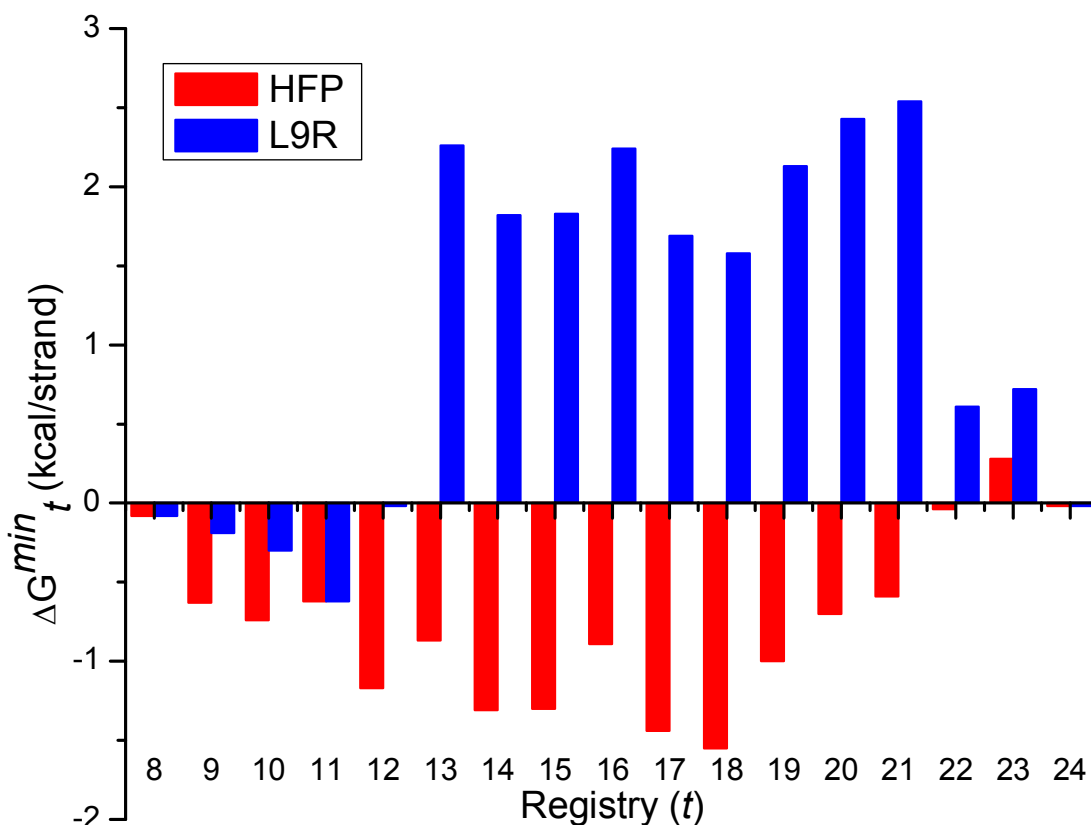


Figure 52. The membrane insertion energies were derived from the Hessa biological hydrophobicity scale for the HFP and L9R-HFP by the methods described in Chapter IV. The L9R-HFP has predominantly positive $\Delta G_{n_t}^{\min}$ whereas HFP has many registries with negative $\Delta G_{n_t}^{\min}$ which suggests that the distribution of registries should be different between constructs. Additionally, it is not obvious that mL9R-HFP should form membrane inserted β sheets since $t < 12$ registries were minimally populated in mHFP, and $t > 12$ registries have positive $\Delta G_{n_t}^{\min}$ in mL9R-HFP.



Figure 53. NMR sample of mV2E-HFP (left) compared to a water standard (right) prior to centrifugation and after mixing overnight. Aggregation of LUV's is evident in mV2E-HFP under our sample preparation conditions.

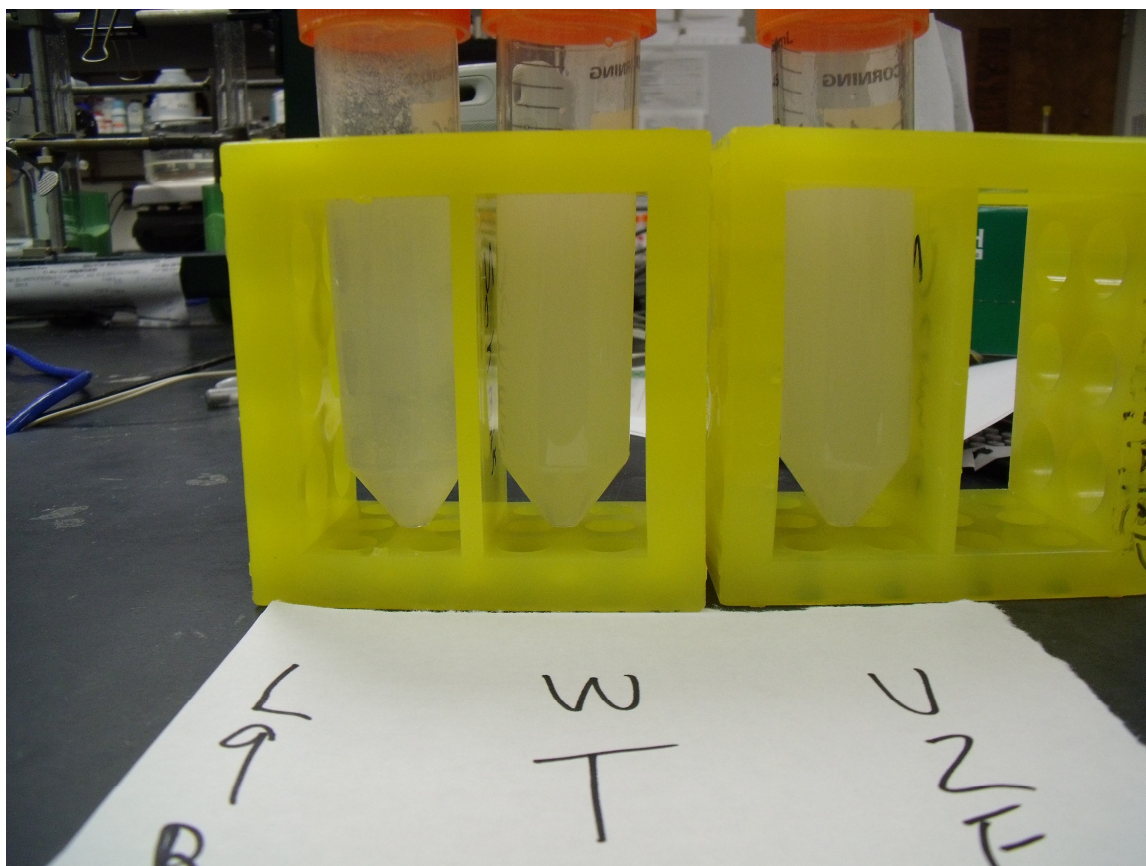


Figure 54. NMR samples of mL9R-HFP (L9R), mHFP (WT), mV2E-HFP (V2E) prior to centrifugation and after mixing overnight. The mL9R-HFP sample appeared to be more transparent than both mHFP and mV2E-HFP, but LUV aggregation was evident.

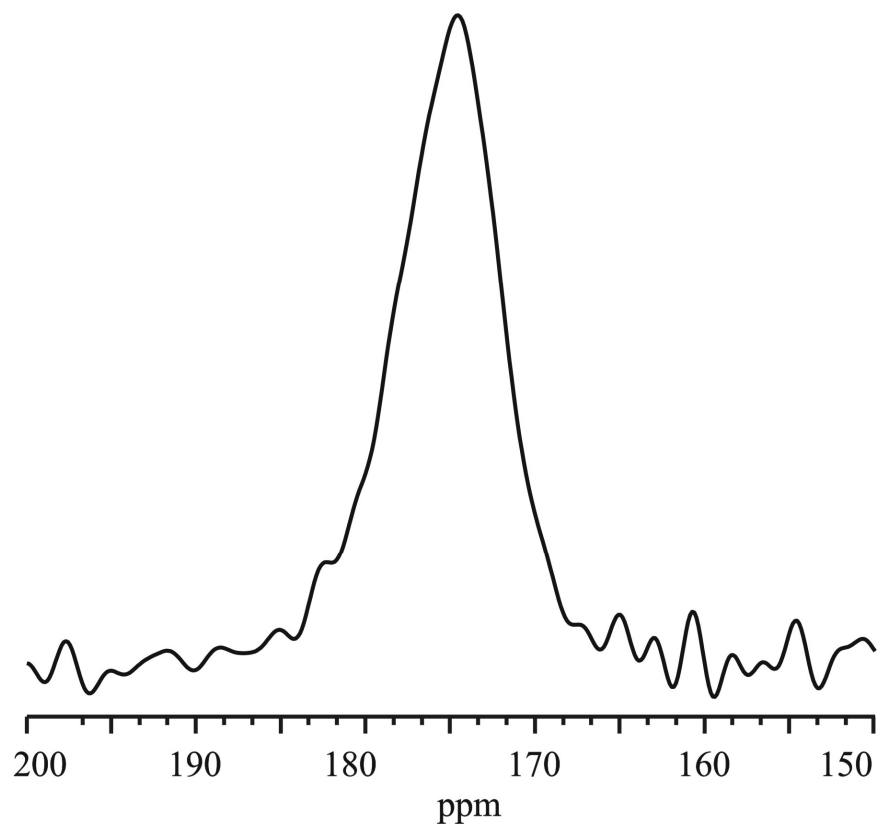


Figure 55. mL9R-HFP with F8CG13N labeling. The chemical shift of 174.6 ppm and 7.5 ppm line full-width at half maximum height indicate the presence of a distribution of secondary structures since the peak spans chemical shifts of α helical, random coil and β sheet structures.

Appendix XIV. HFP, V2E, and L9R ΔG_t^{\min} i_t , and n_t Values.

Table 21. Energy minimized membrane insertion energy parameters.

	HFP			V2E			L9R		
t	ΔG_t^{\min} (kcal/strand)	i_t	n_t	ΔG_t^{\min} (kcal/strand)	i_t	n_t	ΔG_t^{\min} (kcal/strand)	i_t	n_t
8	-0.08	1	8	0.99	3	4	-0.08	1	8
9	-0.63	1	9	0.44	3	5	-0.19	2	7
10	-0.74	2	8	-0.3	4	4	-0.3	4	4
11	-0.62	4	5	-0.62	4	5	-0.62	4	5
12	-1.17	4	6	-1.17	4	6	-0.02	5	4
13	-0.87	2	11	-0.57	5	5	2.26	2	11
14	-1.31	6	4	-1.31	6	4	1.82	6	4
15	-1.3	4	9	-1.3	4	9	1.83	4	9
16	-0.89	6	6	-0.89	6	6	2.24	6	6
17	-1.44	6	7	-1.44	6	7	1.69	6	7
18	-1.55	7	6	-1.55	7	6	1.58	7	6
19	-1	8	5	-1	8	5	2.13	8	5
20	-0.7	7	8	-0.7	7	8	2.43	7	8
21	-0.59	7	9	-0.59	7	9	2.54	7	9
22	-0.04	8	8	-0.04	8	8	0.61	10	4
23	0.28	9	7	0.28	9	7	0.72	10	5
24	-0.02	11	4	-0.02	11	4	-0.02	11	4

Appendix XV. Summary of Hessa Biological Hydrophobicity Scale

Amino acid properties, such as hydrophobicity, are thought to play a key role in determining membrane protein structure. Amino acid solvent partitioning experiments have been used to generate hydrophobicity scales for the twenty naturally occurring amino acids^{112,126}. These scales were developed in part to help with predicting membrane protein structures. Solvent partitioning experiments model hydrophobicity well, but the free energy associated with protein-membrane interactions, such as binding and folding to name a few, are not accounted for in these experiments¹⁰⁷. These concerns have led to the development of a thermodynamically based biological hydrophobicity scale where α helices with unique amino acid sequences were inserted into membranes⁹³ (more details below). I have provided a basic overview of the experiments and highlighted key points that are relevant to structure of membrane-associated HFP constructs in the remainder of this appendix.

Figure 56 illustrates the experimental setup where an α helical segment was mutated into the integral membrane protein leader peptidase. Membrane insertion of this H-segment buries the glycosylation site 2 (G2) into the cytoplasm and exposes glycosylation site 1 (G1) for glycosylation of G1 while translocation of this H-segment lead to glycosylation of both G1 and G2. SDS-PAGE gels were used to quantify the fraction of protein with only G1 glycosylated, f_{g1} , and the fraction of protein with both G1 and G2 glycosylated, f_{g2} . These fractions were expressed as the observed or “apparent” equilibrium constant for membrane insertion of the H segment where $K_{app} = f_{g1}/f_{g2}$. The apparent change in free energy between the membrane inserted and translocated H-segments was calculated by $\Delta G_{app}^{ins,seg} = -RT \ln K_{app}$. Therefore,

determination of the change in free energy for membrane insertion was determined by mutating a single amino acid within a sequence and comparing the $\Delta G_{app}^{ins,seg}$ between the two sequences

$$\text{by } \Delta G_a^{ins} = \Delta G_{app}^{ins,mutant} - \Delta G_{app}^{ins,seg}$$

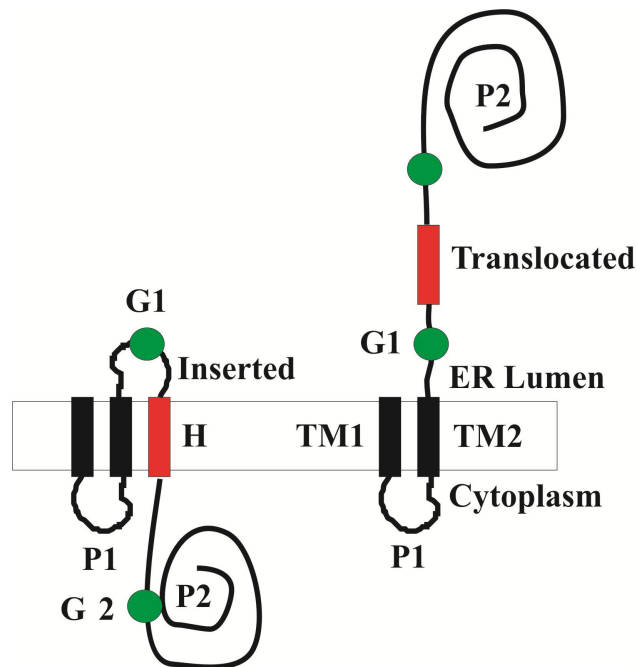


Figure 56. The model systems were composed of two transmembrane domains (TM1 and TM2), two luminal domains (P1 and P2), and two glycosylation acceptor sites (G1 and G2). A third helical transmembrane domain (H) is illustrated in red. Translocation of the H segment from the membrane allows for glycosylation of both G1 and G2 while membrane insertion of the H segment only allows for glycosylation of G1. This figure was modified from literature⁹³.

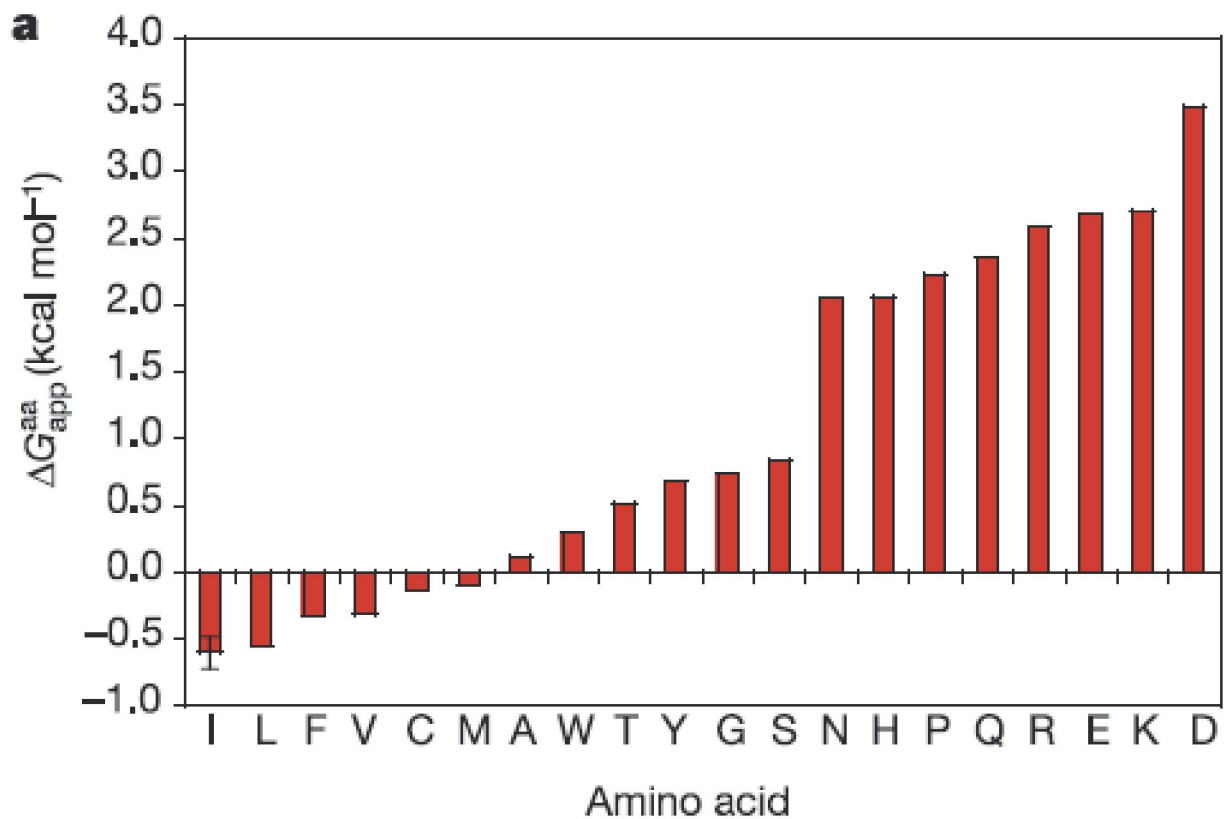


Figure 57. The Hessa biological hydrophobicity scale. This figure was taken from literature⁹³.

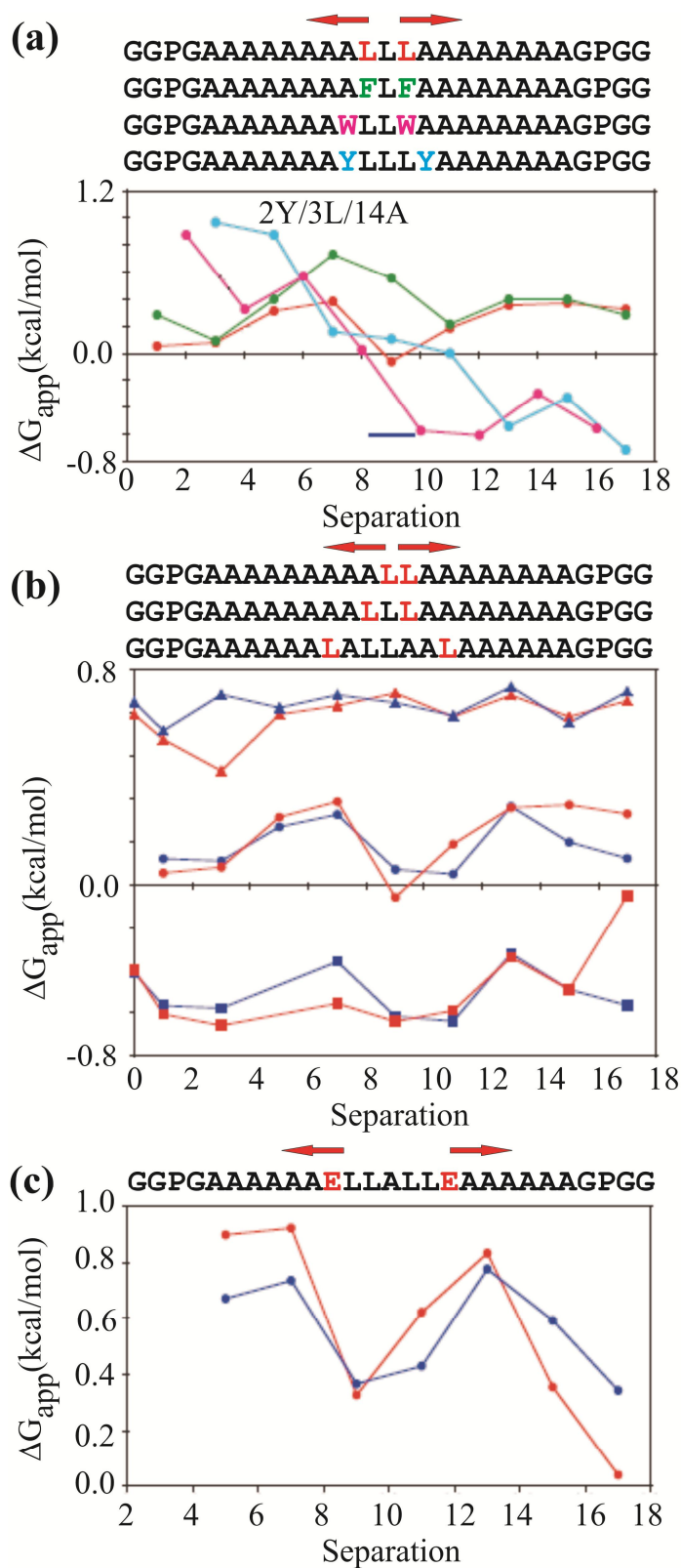
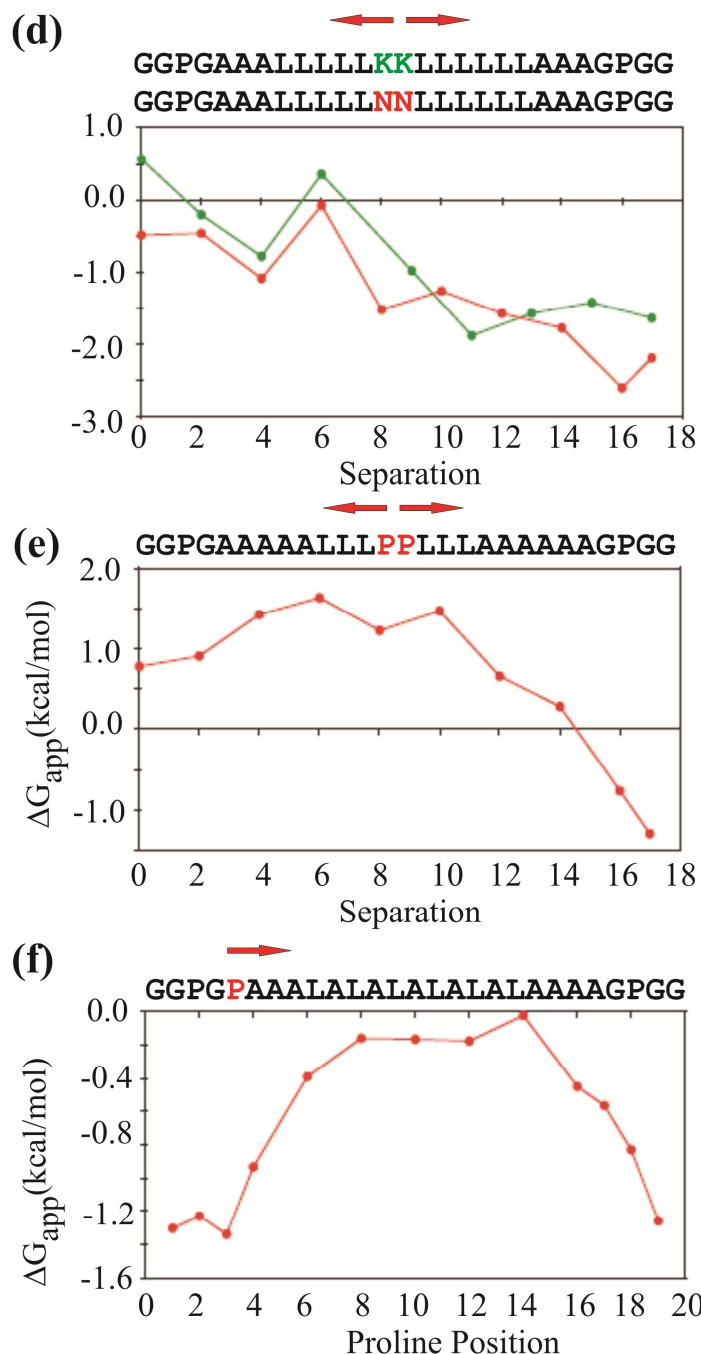


Figure 58. The positional dependence of amino acids. Key points from these figures are discussed below. These figures were taken from literature⁹³.

Figure 58 (cont'd).



The H-segments were generally composed of GGPG-X₁₉-GP GG where X is the H-segment primarily composed of Ala and Leu residues, and the Pro and Gly residues were added adjacent to the α helical segment so that flanking residues had a low probability of secondary structure formation. Sequences composed of three or four Leu and sixteen or fifteen Ala,

respectively, were used to evaluate the flanking residues contribution to $\Delta G_{\text{app}}^{\text{ins,seg}}$ using the sequence $G_z\text{PG-X}_{19}\text{-GPG}_z$ where z was varied from 2 to 6. The resultant $\Delta G_{\text{app}}^{\text{ins,seg}}$ values varied by ± 0.2 kcal/mol from the $z = 2$ construct which suggested that residues flanking the H-segments had minimal contributions to the $\Delta G_{\text{app}}^{\text{ins,seg}}$. Additionally, different sequences of the H-segment were developed with 3 and 4 Leu residues where the six flanking Gly residues were mutated to Asn residues, $\text{NNPN-X}_{19}\text{-NPNN}$, and the $\Delta G_{\text{app}}^{\text{ins,seg}}$ increased by an average of only +0.5 kcal/mol whereas six Gln residues placed in the central region of the H-segment should increase $\Delta G_{\text{app}}^{\text{ins,seg}}$ by $\sim +14.2$ kcal/mol. These combined results suggest that flanking residues have a minimal contribution to the $\Delta G_{\text{app}}^{\text{ins,seg}}$ values, and support that the K_6 tag in our HFP sequence minimally affects the insertion energy of specific β sheet registries since the K_6 tag is C-terminal of the membrane inserted region. Additionally, there are residues without regular secondary structure between the membrane inserted region and the K_6 tag.

In Chapter IV, the insertion energy of a structure was determined by summing ΔG_a^{ins} values, but the “apparent” ΔG_a^{ins} , where the subscript “a” denotes amino acid type, are only approximately additive since the insertion energies have different positional dependences within a structure. For example, the hydrophobic residues Leu and Phe demonstrated a small positional dependence, **Figure 58a,b**, while more polar residues, such as Tyr, Trp, Asn, Gln, Lys, and Ser, demonstrated a large positional dependence where $\Delta G_{\text{app}}^{\text{ins,seg}}$ were lowered as these residues were placed away from the central region and toward the edges of the H-segment, **Figure 58a,c,d**. Additionally, Pro residues are known to disrupt secondary structure. All secondary

structure types have hydrogen bonding between amide and carbonyl groups that result from dipole-dipole interactions. Thus, a sequence with secondary structure is generally considered to be less polar and more energetically favorable for membrane insertion than its unordered counterpart. In this study, residues within the H-segment were mutated into Pro residues at different positions, **Figure 58e,f**. Pro mutations in the central region of the H-segment drastically increased $\Delta G_{app}^{ins,seg}$. As Pro was moved away from the central region of the H-segments, the $\Delta G_{app}^{ins,seg}$ became more negative which was consistent with the observed position dependence of the more polar residues. Polar residues are frequently found in secondary structures and do not inherently disrupt secondary structure; however, relative to an aqueous environment, it is energetically unfavorable to insert polar residues into the hydrophobic environment of the membrane interior.

Appendix XVI. Chapter IV Tables for $\gamma_{1tuv}^{lab}(\tau)$ or $\gamma_{1t_1t_2uv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON for the three registry fittings.

Similar to **Appendix VII**, the spin geometries are illustrated for the associated calculated $\gamma_{1tuv}^{lab}(\tau)$ for specific t , u , and v in the fully constrained model or calculated $\gamma_{1t_1t_2uv}^{lab}(\tau)$ for specific t_1 , t_2 , u , and v in the unconstrained model. In each spin geometry schematic, N or C respectively denote a ^{15}N nucleus or a ^{13}CO nucleus included in the simulation, and X denotes X registry. Each arrow denotes ^{13}CO - ^{15}N dipolar coupling considered in the simulation. For unconstrained model geometries, the top/middle strand registry is at the top of the schematic and the middle/bottom strand registry is at the bottom of the schematic. When more than one spin geometries are shown, each $\gamma(\tau)$ is the average for the displayed geometries.

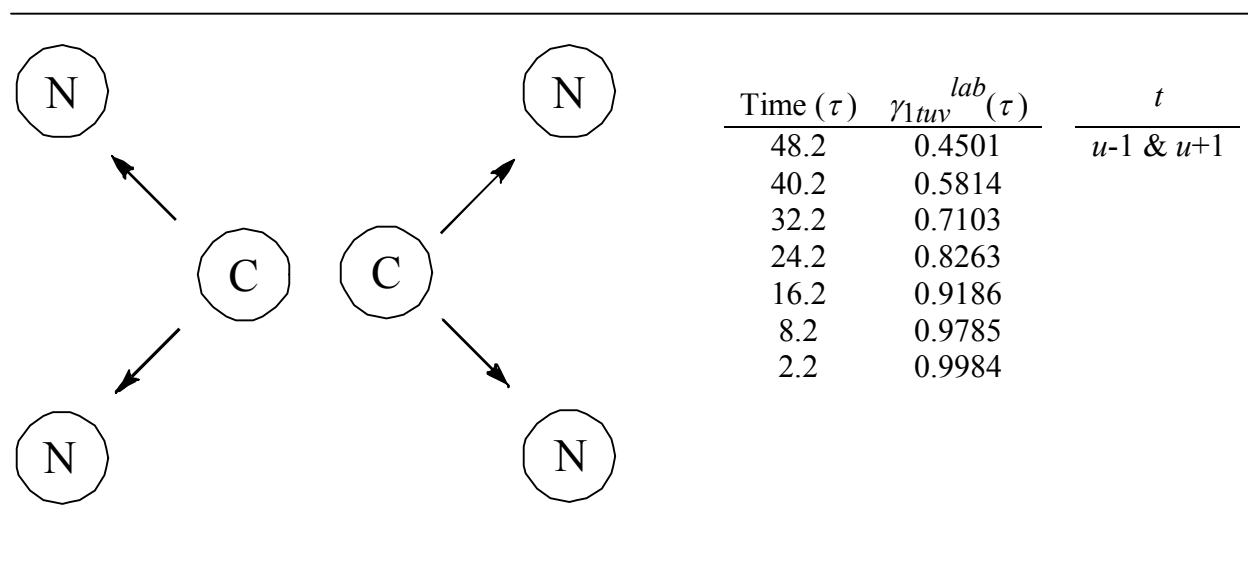
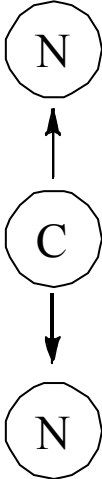


Figure 59. Spin geometries and simulated data for Chapter IV three registry fittings.

Figure 59 (cont'd).

	Time (τ)	$\gamma_{1uv}^{lab}(\tau)$	$\frac{t}{u}$
	48.2	0.0939	
	40.2	0.1186	
	32.2	0.1964	
	24.2	0.3786	
	16.2	0.6453	
	8.2	0.8938	
	2.2	0.9917	

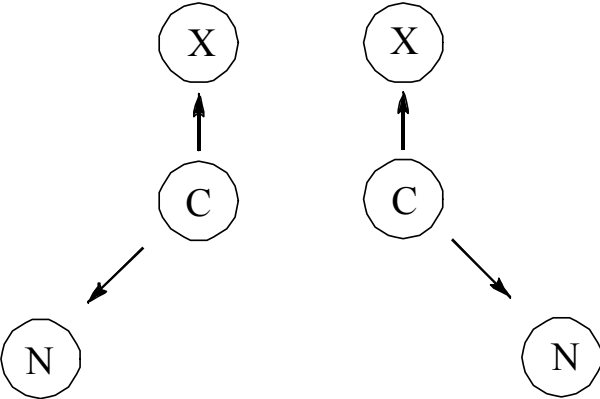
	Time (τ)	$\gamma_{l_1 t_2 uv}^{lab}(\tau)$	$\frac{t_1}{x}$
	48.2	0.6917	
	40.2	0.7769	
	32.2	0.8521	$\frac{t_2}{u-1 \text{ \& } u+1}$
	24.2	0.9143	
	16.2	0.9608	
	8.2	0.9898	
	2.2	0.9992	

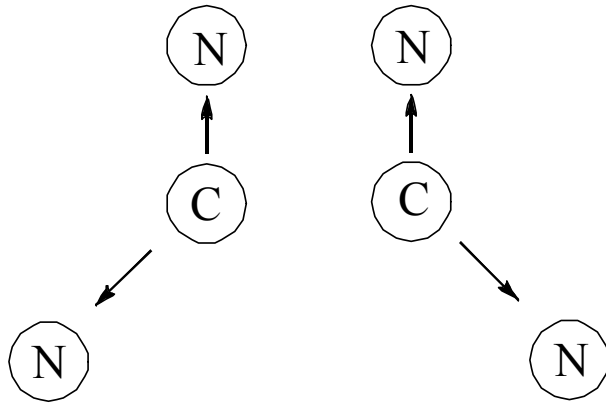
Figure 59 (cont'd).

Time (τ)	$\gamma_{lt_1t_2uv}^{lab}(\tau)$	t_l
48.2	0.6592	$u-1 \text{ \& } u+1$
40.2	0.7520	t_2
32.2	0.8348	x
24.2	0.9039	
16.2	0.9560	
8.2	0.9885	
2.2	0.9991	

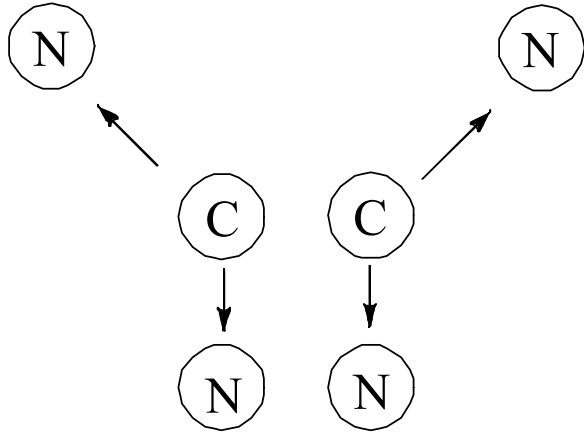
Time (τ)	$\gamma_{lt_1t_2uv}^{lab}(\tau)$	t_l
48.2	-0.0470	u
40.2	-0.0230	t_2
32.2	0.1236	x
24.2	0.3778	
16.2	0.6710	
8.2	0.9064	
2.2	0.9928	

Time (τ)	$\gamma_{lt_1t_2uv}^{lab}(\tau)$	t_l
48.2	0.5675	x
40.2	0.6824	t_2
32.2	0.7870	u
24.2	0.8754	
16.2	0.9427	
8.2	0.9850	
2.2	0.9989	

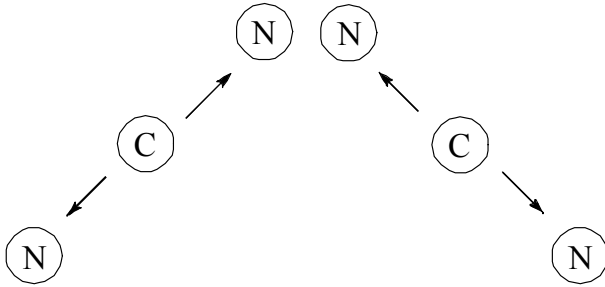
Figure 59 (cont'd).



Time (τ)	$\gamma_{l_1 t_2 uv}^{lab}(\tau)$	t_l
48.2	0.0011	u
40.2	0.0156	t_2
32.2	0.1288	$u-1 \text{ \& } u+1$
24.2	0.3565	
16.2	0.6476	
8.2	0.8974	
2.2	0.9921	



Time (τ)	$\gamma_{l_1 t_2 uv}^{lab}(\tau)$	t_l
48.2	0.3739	$u-1 \text{ \& } u+1$
40.2	0.5132	t_2
32.2	0.6571	u
24.2	0.7914	
16.2	0.9012	
8.2	0.9737	
2.2	0.9980	



Time (τ)	$\gamma_{l_1 t_2 uv}^{lab}(\tau)$	t_l
48.2	0.4926	$u+1 \text{ or } u-1$
40.2	0.6037	t_2
32.2	0.7200	$u-1 \text{ or } u+1$
24.2	0.8293	
16.2	0.9191	
8.2	0.9785	
2.2	0.9984	

Appendix XVII. Chapter IV Tables for the unique $\gamma_{1tuv}^{lab}(\tau)$ spin geometries with calculated values from SIMPSON for the five registry fittings.

Similar to **Appendix VII**, the spin geometries are illustrated for the associated calculated $\gamma_{1tuv}^{lab}(\tau)$ for specific t , u , and v in the fully constrained model. In each spin geometry schematic, N or C respectively denote a ^{15}N nucleus or a ^{13}CO nucleus included in the simulation. Each arrow denotes ^{13}CO - ^{15}N dipolar coupling considered in the simulation. Two spin geometries are shown and each $\gamma(\tau)$ is the average for the displayed geometries.

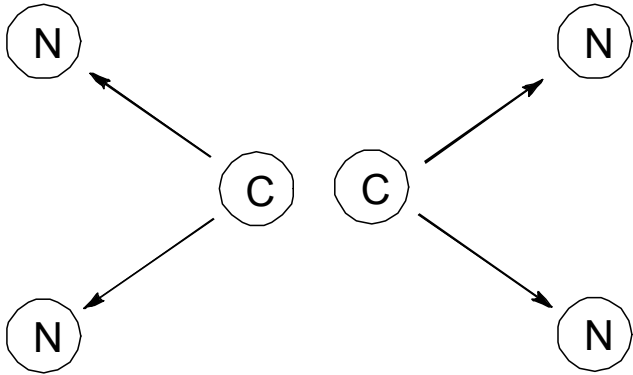
	Time (τ)	$\gamma_{1tuv}^{lab}(\tau)$	t
	48.2	0.9052	$u-2$ & $u+2$
	40.2	0.9334	
	32.2	0.9569	
	24.2	0.9755	
	16.2	0.9890	
	8.2	0.9971	
	2.2	0.9998	

Figure 60. Additional spin geometries and simulated data for Chapter IV five registry fittings.

REFERENCES

REFERENCES

1. White, J.M., Delos, S.E., Brecher, M. & Schornberg, K. Structures and mechanisms of viral membrane fusion proteins: Multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.* 43, 189-219 (2008).
2. Zepp, F. Principles of vaccine design-Lessons from nature. *Vaccine* 28, C14-C24.
3. Korber, B. et al. Evolutionary and immunological implications of contemporary HIV-1 variation. *British Medical Bulletin* 58, 19-42 (2001).
4. Tavassoli, A. Targeting the protein-protein interactions of the HIV lifecycle. *Chemical Society Reviews* 40, 1337-1346 (2011).
5. Levy, J.A. Pathogenesis Of Human-Immunodeficiency-Virus Infection. *Microbiological Reviews* 57, 183-289 (1993).
6. Austin, C.P., Brady, L.S., Insel, T.R. & Collins, F.S. NIH Molecular Libraries Initiative. *Science* 306, 1138-1139 (2004).
7. Stein, R.L. High-throughput screening in academia: The Harvard experience. *Journal Of Biomolecular Screening* 8, 615-619 (2003). 8. Grewe, C., Beck, A. & Gelderblom, H.R. HIV: early virus-cell interactions. *J Acquir Immune Defic Syndr* 3, 965-74 (1990).
9. Wyatt, R. & Sodroski, J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280, 1884-8 (1998).
10. Kwong, P.D. et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648-59 (1998).
11. Kibler, K.V. et al. Polyarginine inhibits gp160 processing by furin and suppresses productive human immunodeficiency virus type 1 infection. *Journal Of Biological Chemistry* 279, 49055-49063 (2004).
12. Moore, J.P., McKeating, J.A., Weiss, R.A. & Sattentau, Q.J. Dissociation Of Gp120 From Hiv-1 Virions Induced By Soluble Cd4. *Science* 250, 1139-1142 (1990).
13. Chan, D.C. & Kim, P.S. HIV entry and its inhibition. *Cell* 93, 681-684 (1998).
14. Agnihotri, K.D., Tripathy, S.P., Jere, A.J., Kale, S.M. & Paranjape, R.S. Molecular analysis of gp41 sequences of HIV type 1 subtype C from India. *J AIDS-Journal Of Acquired Immune Deficiency Syndromes* 41, 345-351 (2006).
15. Buzon, V. et al. Crystal Structure of HIV-1 gp41 Including Both Fusion Peptide and Membrane Proximal External Regions. *Plos Pathogens* 6(2010).
16. 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.* 94, 12303-12308 (1997).

17. Caffrey, M. et al. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J.* 17, 4572-4584 (1998).
18. Yang, Z.N. et al. The crystal structure of the SIV gp41 ectodomain at 1.47 Å resolution. *J. Struct. Biol.* 126, 131-144 (1999).
19. Buzon, V. et al. Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *Plos Pathogens* 6, e1000880 (2010).
20. Eckert, D.M. & Kim, P.S. Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.* 70, 777-810 (2001).
21. Freed, E.O., Delwart, E.L., Buchsacher, G.L., Jr. & Panganiban, A.T. A mutation in the human immunodeficiency virus type 1 transmembrane glycoprotein gp41 dominantly interferes with fusion and infectivity. *Proc. Natl. Acad. Sci. U.S.A.* 89, 70-74 (1992).
22. Teissier, E., Penin, F. & Pecheur, E.I. Targeting Cell Entry of Enveloped Viruses as an Antiviral Strategy. *Molecules* 16, 221-250 (2011).
23. Bai, X.F. et al. Impact of the enfuvirtide resistance mutation N43D and the associated baseline polymorphism E137K on peptide sensitivity and six-helix bundle structure. *Biochemistry* 47, 6662-6670 (2008).
24. Hill, A.M., Cho, M. & Mrus, J.M. The Costs of Full Suppression of Plasma HIV RNA in Highly Antiretroviral-Experienced Patients. *Aids Reviews* 13, 41-48.
25. Qiang, W. & Weliky, D.P. HIV fusion peptide and its cross-linked oligomers: Efficient syntheses, significance of the trimer in fusion activity, correlation of beta strand conformation with membrane cholesterol, and proximity to lipid headgroups. *Biochemistry* 48, 289-301 (2009).
26. Qiang, W., Sun, Y. & Weliky, D.P. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 106, 15314-15319 (2009).
27. Markosyan, R.M., Leung, M.Y. & Cohen, F.S. The Six-Helix Bundle of Human Immunodeficiency Virus Env Controls Pore Formation and Enlargement and Is Initiated at Residues Proximal to the Hairpin Turn. *Journal Of Virology* 83, 10048-10057 (2009).
28. Delahunty, M.D., Rhee, I., Freed, E.O. & Bonifacino, J.S. Mutational analysis of the fusion peptide of the human immunodeficiency virus type 1: identification of critical glycine residues. *Virology* 218, 94-102 (1996).
29. Yang, J. & Weliky, D.P. Solid state nuclear magnetic resonance evidence for parallel and antiparallel strand arrangements in the membrane-associated HIV-1 fusion peptide. *Biochemistry* 42, 11879-11890 (2003).

30. Yang, R., Prorok, M., Castellino, F.J. & Weliky, D.P. A trimeric HIV-1 fusion peptide construct which does not self-associate in aqueous solution and which has 15-fold higher membrane fusion rate. *J. Am. Chem. Soc.* 126, 14722-14723 (2004).
31. Kliger, Y. et al. Fusion peptides derived from the HIV type 1 glycoprotein 41 associate within phospholipid membranes and inhibit cell-cell fusion. Structure- function study. *J. Biol. Chem.* 272, 13496-13505 (1997).
32. Chang, D.K., Chien, W.J. & Cheng, S.F. 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.* 247, 896-905 (1997).
33. Yang, J., Prorok, M., Castellino, F.J. & Weliky, D.P. Oligomeric β -structure of the membrane-bound HIV-1 fusion peptide formed from soluble monomers. *Biophys. J.* 87, 1951-1963 (2004).
34. 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* 44, 13354-13364 (2005).
35. Yang, J. et al. Solid state NMR measurements of conformation and conformational distributions in the membrane-bound HIV-1 fusion peptide. *J. Mol. Graph. Model.* 19, 129-135 (2001).
36. Qiang, W., Bodner, M.L. & Weliky, D.P. 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. *Journal Of The American Chemical Society* 130, 5459-5471 (2008).
37. Yang, J., Gabrys, C.M. & Weliky, D.P. Solid-state nuclear magnetic resonance evidence for an extended beta strand conformation of the membrane-bound HIV-1 fusion peptide. *Biochemistry* 40, 8126-8137. (2001).
38. Yang, J., Parkanzky, P.D., Bodner, M.L., Duskin, C.G. & Weliky, D.P. Application of REDOR subtraction for filtered MAS observation of labeled backbone carbons of membrane-bound fusion peptides. *J. Magn. Reson.* 159, 101-110 (2002).
39. Yang, J., Prorok, M., Castellino, F.J. & Weliky, D.P. Oligomeric beta structure of the membrane-bound HIV-1 fusion peptide formed from soluble monomers. *Biophys. J.* 87, 1951-1963 (2004).
40. Qiang, W., Sun, Y. & Weliky, D.P. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15314-15319 (2009).

41. Yang, X.Z., Kurteva, S., Ren, X.P., Lee, S. & Sodroski, J. Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1. *J. Virol.* 79, 12132-12147 (2005).
42. Magnus, C., Rusert, P., Bonhoeffer, S., Trkola, A. & Regoes, R.R. Estimating the stoichiometry of Human Immunodeficiency Virus entry. *J. Virol.* 83, 1523-1531 (2009).
43. Sougrat, R. et al. Electron tomography of the contact between T cells and SIV/HIV-1: Implications for viral entry. *PLOS Pathogens* 3, 571-581 (2007).
44. Lau, W.L., Ege, D.S., Lear, J.D., Hammer, D.A. & DeGrado, W.F. Oligomerization of fusogenic peptides promotes membrane fusion by enhancing membrane destabilization. *Biophys. J.* 86, 272-284 (2004).
45. Pan, J.H., Lai, C.B., Scott, W.R.P. & Straus, S.K. Synthetic fusion peptides of tick-borne Encephalitis virus as models for membrane fusion. *Biochemistry* 49, 287-296 (2010).
46. Sougrat, R. et al. Electron tomography of the contact between T cells and SIV/HIV-1: Implications for viral entry. *Plos Pathogens* 3, 571-581 (2007).
47. Sackett, K. & Shai, Y. The HIV fusion peptide adopts intermolecular parallel β -sheet structure in membranes when stabilized by the adjacent N-terminal heptad repeat: A ^{13}C FTIR study. *J. Mol. Biol.* 350, 790-805 (2005).
48. Haque, M.E., Koppaka, V., Axelsen, P.H. & Lentz, B.R. Properties and structures of the influenza and HIV fusion peptides on lipid membranes: Implications for a role in fusion. *Biophys. J.* 89, 3183-3194 (2005).
49. Han, X. & Tamm, L.K. A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13097-13102 (2000).
50. Qiang, W., Bodner, M.L. & Weliky, D.P. Solid-state NMR spectroscopy of human immunodeficiency virus fusion peptides associated with host-cell-like membranes: 2D correlation spectra and distance measurements support a fully extended conformation and models for specific antiparallel strand registries. *J. Am. Chem. Soc.* 130, 5459-5471 (2008).
51. Bodner, M.L. et al. Temperature dependence and resonance assignment of ^{13}C NMR spectra of selectively and uniformly labeled fusion peptides associated with membranes. *Magn. Reson. Chem.* 42, 187-194 (2004).
52. Benzinger, T.L. et al. Propagating structure of Alzheimer's β -amyloid(10-35) is parallel β -sheet with residues in exact register. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13407-13412 (1998).
53. Tycko, R. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Quart. Rev. Biophys.* 39, 1-55 (2006).

54. Zheng, Z., Yang, R., Bodner, M.L., and Weliky, D.P. Conformational flexibility and strand arrangements of the membrane-associated HIV fusion peptide trimer probed by solid-state NMR spectroscopy. *Biochemistry* 45, 12960-12975 (2006).
55. Zheng, Z., Qiang, W. & Weliky, D.P. 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.* 245, S247-S260 (2007).
56. 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 succinimidoxymethyl anion as a leaving group in amino-acid protection. *Synthesis-Stuttgart* 8, 671-673 (1983).
57. Chan, W.C., and White, P.D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, 94-109 (2000).
58. Brugger, B. et al. The HIV lipidome: A raft with an unusual composition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2641-2646 (2006).
59. Siminovitch, D.J., Ruocco, M.J., Makriyannis, A. & Griffin, R.G. The effect of cholesterol on lipid dynamics and packing in diether phosphatidylcholine bilayers. X-ray diffraction and ^2H -NMR study. *Biochim. Biophys. Acta* 901, 191-200 (1987).
60. Komatsu, H. & Rowe, E.S. Effect of cholesterol on the ethanol-induced interdigitated gel phase in phosphatidylcholine: Use of fluorophore pyrene-labeled phosphatidylcholine. *Biochemistry* 30, 2463-2470 (1991).
61. Gabrys, C.M. et al. 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* 1798, 194-201 (2010).
62. Huster, D., Xiao, L.S. & Hong, M. Solid-state NMR investigation of the dynamics of the soluble and membrane-bound colicin Ia channel-forming domain. *Biochemistry* 40, 7662-7674 (2001).
63. Schmick, S.D. & Weliky, D.P. Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide. *Biochemistry* 49, 10623 (2010).
64. Koy, C. et al. Matrix-assisted laser desorption/ionization-quadrupole ion trap-time of flight mass spectrometry sequencing resolves structures of unidentified peptides obtained by in-gel tryptic digestion of haptoglobin derivatives from human plasma proteomes. *Proteomics* 3, 851-858 (2003).
65. Gehrig, P.M., Roschitzki, B., Rutishauser, D., Reiland, S. & Schlapbach, R. Phosphorylated serine and threonine residues promote site-specific fragmentation of

- singly charged, arginine-containing peptide ions. *Rapid Communications In Mass Spectrometry* 23, 1435-1445 (2009).
66. Zhang, X., Jai-Nhuknan, J. & Cassady, C.J. Collision-induced dissociation and post-source decay of model dodecapeptide ions containing lysine and glycine. *International Journal Of Mass Spectrometry* 171, 135-145 (1997).
 67. Patterson, S.D. & Katta, V. Prompt Fragmentation Of Disulfide-Linked Peptides During Matrix-Assisted Laser-Desorption Ionization Mass-Spectrometry. *Analytical Chemistry* 66, 3727-3732 (1994).
 68. Harris, R.K. *Nuclear magnetic resonance spectroscopy*, Medium: X; Size: Pages: 272 (1986).
 69. Wylie, B.J. et al. Chemical-shift anisotropy measurements of amide and carbonyl resonances in a microcrystalline protein with slow magic-angle spinning NMR spectroscopy. *Journal Of The American Chemical Society* 129, 5318-+ (2007).
 70. Hartzell, C.J., Whitfield, M., Oas, T.G. & Drobny, G.P. Determination of the N-15 and C-13 Chemical-Shift Tensors of L- [C-13]Alanyl-L-[N-15]Alanine From the Dipole-Coupled Powder Patterns. *Journal of the American Chemical Society* 109, 5966-5969 (1987).
 71. Long, H.W. & Tycko, R. Biopolymer conformational distributions from solid-state NMR: alpha-helix and 3(10)-helix contents of a helical peptide. *J. Am. Chem. Soc.* 120, 7039-7048 (1998).
 72. Frye, J.S. & Maciel, G.E. Setting The Magic Angle Using A Quadrupolar Nuclide. *Journal Of Magnetic Resonance* 48, 125-131 (1982).
 73. Bennett, A.E., Rienstra, C.M., Auger, M., Lakshmi, K.V. & Griffin, R.G. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 103, 6951-6958 (1995).
 74. Kricheldorf, H.R. & Muller, D. Secondary structure of peptides. 3. ¹³C NMR cross polarization/magic angle spinning spectroscopic characterization of solid polypeptides. *Macromolecules* 16, 615-623 (1983).
 75. Yannoni, C.S. High-Resolution Nmr In Solids - The Cpmas Experiment. *Accounts Of Chemical Research* 15, 201-208 (1982).
 76. Chevelkov, V., Fink, U. & Reif, B. Accurate Determination of Order Parameters from (1)H,(15)N Dipolar Couplings in MAS Solid-State NMR Experiments. *Journal Of The American Chemical Society* 131, 14018-14022 (2009).
 77. Zhang, H.Y., Neal, S. & Wishart, D.S. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* 25, 173-195 (2003).

78. Gullion, T. & Schaefer, J. Rotational-echo double-resonance NMR. *J. Magn. Reson.* 81, 196-200 (1989).
79. Gullion, T., Baker, D.B. & Conradi, M.S. New, Compensated Carr-Purcell Sequences. *J. Magn. Reson.* 89, 479-484 (1990).
80. Fowler, D.J., Weis, R.M. & Thompson, L.K. Kinase-active signaling complexes of bacterial chemoreceptors do not contain proposed receptor-receptor contacts observed in crystal structures. *Biochemistry* 49, 1425-1434 (2010).
81. Bennett, A.E. et al. Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* 108, 9463-9479 (1998).
82. Gullion, T. Introduction to rotational-echo, double-resonance NMR. *Concepts Magn. Reson.* 10, 277-289 (1998).
83. Bevington, P.R. & Robinson, D.K. *Data Reduction and Error Analysis for the Physical Sciences*, 38-52 (McGraw-Hill, Boston, 1992).
84. Bak, M., Rasmussen, J.T. & Nielsen, N.C. SIMPSON: A general simulation program for solid-state NMR spectroscopy. *J. Magn. Reson.* 147, 296-330 (2000).
85. Farrar, T.C. Density Matrices in NMR Spectroscopy: Part I. *Concepts in Magnetic Resonance* 2, 1-12 (1990).
86. Farrar, T.C. Density Matrices in NMR Spectroscopy: Part II. *Concepts in Magnetic Resonance* 2, 55-61 (1990).
87. Oas, T.G., Hartzell, C.J., McMahon, T.J., Drobny, G.P. & Dahlquist, F.W. The carbonyl ^{13}C chemical-shift tensors of 5 peptides determined from ^{15}N dipole-coupled chemical shift powder patterns. *J. Am. Chem. Soc.* 109, 5956-5962 (1987).
88. Bak, M., Schultz, R., Vosegaard, T. & Nielsen, N.C. Specification and visualization of anisotropic interaction tensors in polypeptides and numerical simulations in biological solid-state NMR. *J. Magn. Reson.* 154, 28-45 (2002).
89. Gabrys, C.M., Yang, J. & Weliky, D.P. Analysis of local conformation of membrane-bound and polycrystalline peptides by two-dimensional slow-spinning rotor-synchronized MAS exchange spectroscopy. *J. Biomol. NMR* 26, 49-68 (2003).
90. Pereira, F.B., Goni, F.M., Muga, A. & Nieva, J.L. Permeabilization and fusion of uncharged lipid vesicles induced by the HIV-1 fusion peptide adopting an extended conformation: dose and sequence effects. *Biophys. J.* 73, 1977-1986 (1997).
91. Jaroniec, C.P. et al. Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry* 44, 16167-16180 (2005).

92. Tristram-Nagle, S. et al. HIV fusion peptide penetrates, disorders, and softens T-cell membrane mimics. *J. Mol. Biol.* 402, 139-153 (2010).
93. Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H., and Heijne, G. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433, 377-381 (2005).
94. Petkova, A.T. et al. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* 307, 262-265 (2005).
95. Yang, R., Yang, J. & Weliky, D.P. Synthesis, enhanced fusogenicity, and solid state NMR measurements of cross-linked HIV-1 fusion peptides. *Biochemistry* 42, 3527-3535 (2003).
96. Yang, X.Z., Kurteva, S., Ren, X.P., Lee, S. & Sodroski, J. Stoichiometry of envelope glycoprotein trimers in the entry of human immunodeficiency virus type 1. *Journal Of Virology* 79, 12132-12147 (2005).
97. Magnus, C., Rusert, P., Bonhoeffer, S., Trkola, A. & Regoes, R.R. Estimating the stoichiometry of Human Immunodeficiency Virus entry. *Journal Of Virology* 83, 1523-1531 (2009).
98. Nelson, R. et al. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435, 773-778 (2005).
99. Jehle, S. et al. Solid-state NMR and SAXS studies provide a structural basis for the activation of alpha B-crystallin oligomers. *Nature Structural & Molecular Biology* 17, 1037-U1.
100. Antzutkin, O.N., Leapman, R.D., Balbach, J.J. & Tycko, R. Supramolecular structural constraints on Alzheimer's beta-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. *Biochemistry* 41, 15436-15450 (2002).
101. Galzitskaya, O.V., Garbuzynskiy, S.O. & Lobanov, M.Y. Expected packing density allows prediction of both amyloidogenic and disordered regions in protein chains. *Journal Of Physics-Condensed Matter* 19(2007).
102. Antzutkin, O.N. et al. Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. *Proc Natl Acad Sci U S A* 97, 13045-50. (2000).
103. Balbach, J.J. et al. Amyloid fibril formation by A β ₁₆₋₂₂, a seven-residue fragment of the Alzheimer's β -amyloid peptide, and structural characterization by solid state NMR. *Biochemistry* 39, 13748-13759. (2000).
104. Gordon, D.J., Balbach, J.J., Tycko, R. & Meredith, S.C. Increasing the amphiphilicity of an amyloidogenic peptide changes the beta-sheet structure in the fibrils from antiparallel to parallel. *Biophys. J.* 86, 428-434 (2004).

105. Trovato, A., Chiti, F., Maritan, A. & Seno, F. Insight into the structure of amyloid fibrils from the analysis of globular proteins. *Plos Computational Biology* 2, 1608-1618 (2006).
106. Qiang, W., Yang, J. & Weliky, D.P. Solid-state nuclear magnetic resonance measurements of HIV fusion peptide to lipid distances reveal the intimate contact of beta strand peptide with membranes and the proximity of the Ala-14-Gly-16 region with lipid headgroups. *Biochemistry* 46, 4997-5008 (2007).
107. White, S.H. & Wimley, W.C. Membrane protein folding and stability: Physical principles. *Annual Review Of Biophysics And Biomolecular Structure* 28, 319-365 (1999).
108. Chou, P.Y. & Fasman, G.D. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* 47, 45-148 (1978).
109. Giesler, M., Thorgerson, M., Masterson, L. & Loh, A.P. Effect of peptide chain length and sidechain steric hinderance on helix structure and dynamics. *Biophysical Journal* 80, 398A-398A (2001).
110. Pritsker, M., Rucker, J., Hoffman, T.L., Doms, R.W. & Shai, Y. Effect of nonpolar substitutions of the conserved Phe11 in the fusion peptide of HIV-1 gp41 on its function, structure, and organization in membranes. *Biochemistry* 38, 11359-11371 (1999).
111. Merkel, J.S. & Regan, L. Aromatic rescue of glycine in beta sheets. *Folding & Design* 3, 449-455 (1998).
112. Wimley, W.C., and White, S.H. Experimental determined hydrophobicity scales for proteins at membrane interfaces. *Nature Struc. Biol.* 8, 842-848 (1996).
113. Wimley, W.C. et al. Folding of beta-sheet membrane proteins: A hydrophobic hexapeptide model. *Journal Of Molecular Biology* 277, 1091-1110 (1998).
114. Ahmed, M. et al. Structural conversion of neurotoxic amyloid-[beta]1-42 oligomers to fibrils. *Nat Struct Mol Biol* 17, 561.
115. Joh, N.H. et al. Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453, 1266-U73 (2008).
116. Riess, J.G. Fluorinated Vesicles. *Journal Of Drug Targeting* 2, 455-468 (1994).
117. Riess, J.G. & Krafft, M.P. Fluorinated Phosphocholine-Based Amphiphiles As Components Of Fluorocarbon Emulsions And Fluorinated Vesicles. *Chemistry And Physics Of Lipids* 75, 1-14 (1995).
118. Adachi, T., Takahashi, H., Ohki, K. & Hatta, I. Interdigitated Structure Of Phospholipid-Alcohol Systems Studied By X-Ray-Diffraction. *Biophysical Journal* 68, 1850-1855 (1995).

119. Ozdirekcan, S., Rijkers, D.T.S., Liskamp, R.M.J. & Killian, J.A. Influence of flanking residues on tilt and rotation angles of transmembrane peptides in lipid bilayers. A solid-state H-2 NMR study. *Biochemistry* 44, 1004-1012 (2005).
120. Sack, I., Goldbourt, A., Vega, S. & Buntkowsky, G. Deuterium REDOR: Principles and applications for distance measurements. *Journal Of Magnetic Resonance* 138, 54-65 (1999).
121. Gullion, T. Measuring C-13-D-2 dipolar couplings with a universal REDOR dephasing curve. *Journal Of Magnetic Resonance* 146, 220-222 (2000).
122. Villalain, J. Location of cholesterol in model membranes by magic-angle-sample-spinning NMR. *European Journal Of Biochemistry* 241, 586-593 (1996).
123. Harroun, T.A., Katsaras, J. & Wassall, S.R. Cholesterol hydroxyl group is found to reside in the center of a polyunsaturated lipid membrane. *Biochemistry* 45, 1227-1233 (2006).
124. Harroun, T.A., Katsaras, J. & Wassall, S.R. Cholesterol is found to reside in the center of a polyunsaturated lipid membrane. *Biochemistry* 47, 7090-7096 (2008).
125. Ingallinella, P. et al. Addition of a cholesterol group to an HIV-1 peptide fusion inhibitor dramatically increases its antiviral potency. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 106, 5801-5806 (2009).
126. Wimley, W.C., Geamer, T.P., and White, S.H. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* 35, 5109-5124 (1996).