

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/56857>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

**Development of Nuclear
Magnetic Resonance Methods
for Determination of Membrane
Protein Structure**

Dharmesh K. Patel

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy



Department of Chemistry

September 2012

TABLE OF CONTENTS

i.	Table of contents.....	i
ii.	List of figures.....	vii
iii.	List of tables	xi
iv.	Acknowledgments.....	xii
v.	Declaration.....	xiii
vi.	Abbreviations.....	xiv
vii.	Summary	xx
1	INTRODUCTION	1
1.1	Membrane proteins	1
1.1.1	Challenges when studying membrane proteins.....	2
1.1.2	Expression.....	2
1.1.3	Solubilisation.....	4
1.1.4	Purification.....	5
1.2	Membrane mimetic systems.....	6
1.2.1	Detergent micelles	7
1.2.2	Liposomes.....	7
1.2.3	Amphipols and nanodiscs.....	9
1.2.4	Bicelles	10
1.3	Methods for structure determination	12
1.3.1	X-ray crystallography.....	12
1.3.2	Solution NMR.....	13
1.3.3	Solid state NMR (ssNMR).....	14

Table of contents

1.3.4	Chemical shift anisotropy (CSA).....	15
1.3.5	Dipolar coupling.....	15
1.3.6	Quadrupolar coupling.....	16
1.4	NMR Theory.....	17
1.4.1	Spin.....	17
1.4.2	Magnetisation.....	19
1.4.3	R.F pulses and the Rotating Frame.....	21
1.4.4	Chemical Shift.....	23
1.4.5	Two dimensional (2D) NMR.....	24
1.5	Solid state NMR.....	26
1.5.1	Magic angle spinning (MAS).....	27
1.5.2	Cross polarisation (CP).....	28
1.5.3	Proton decoupling.....	29
1.5.4	Recoupling.....	30
1.5.5	Dipolar Assisted Rotational Recoupling (DARR).....	30
1.6	Glycophorin A (GpA) as a model TM.....	32
1.7	Bovine Papillomavirus E5 (BPV E5).....	33
1.8	Aims and objectives of this study.....	35
1.9	References.....	36
2	MATERIALS AND METHODS.....	41
2.1	Reagents and Chemicals.....	41
2.2	Buffers.....	41
2.3	Peptide design.....	42
2.3.1	Glycophorin A (GpA) peptides.....	42
2.3.2	Bovine Papillomavirus E5 peptides.....	42

2.4	Peptide purification using reverse-phase high performance liquid chromatography	43
2.5	Mass Spectrometry.....	45
2.5.1	Electrospray Ionisation (ESI) Mass spectrometry	45
2.5.2	Matrix Assisted Laser desorption Ionisation (MALDI) Mass spectrometry	45
2.6	Peptide preparation and characterisation	46
2.6.1	Lyophilisation.....	46
2.6.2	Protein concentration determination	46
2.6.3	SDS-PAGE analysis of purified peptides	47
2.6.4	Coomassie staining	47
2.6.5	Silver staining.....	47
2.7	Peptide reconstitution into lipid vesicles	48
2.7.1	Detergent solubilisation of peptides	48
2.7.2	Detergent removal	49
2.7.3	Co-solubilisation	50
2.7.4	Transfer of samples to solid state NMR rotor.....	51
2.8	Characterisation of proteoliposomes.....	51
2.8.1	Circular Dichroism (CD) spectroscopy.....	51
2.8.2	Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy	52
2.8.3	Electron Microscopy (EM)	53
2.9	Magic Angle Spinning (MAS) Solid state NMR spectroscopy.....	54
2.9.1	Solid state NMR experimental procedure.....	54
2.9.2	1D ³¹ P/ ¹ H lipid NMR experiments.....	55
2.9.3	1D ¹³ C NMR experiments	55

2.9.4	2D ^{13}C - ^{13}C DARR NMR experiments.....	56
2.9.5	2D ^{13}C - ^{15}N NMR experiments.....	56
2.10	Solution NMR experiments.....	57
2.10.1	Bicelle preparation.....	57
2.10.2	Solution NMR experimental procedure.....	58
2.10.3	1D ^1H NMR spectroscopy.....	58
2.10.4	2D heteronuclear NMR spectroscopy.....	58
2.10.5	3D heteronuclear NMR spectroscopy.....	59
2.11	References.....	59
3	PEPTIDE DESIGN, PREPARATION AND CHARACTERISATION.....	62
3.1	Introduction.....	62
3.2	Peptide design and purification for ssNMR analyses.....	63
3.3	Reconstitution of GpA peptide into lipid vesicles.....	75
3.3.1	CD for screening of detergents.....	76
3.3.2	Electron Microscopy.....	81
3.3.3	ATR-FTIR analysis of reconstituted peptides.....	83
3.3.4	Co-solubilisation of E5 peptides.....	87
3.4	Summary.....	90
3.5	References.....	92
4	SOLID STATE NMR ANALYSIS OF RECONSTITUTED GPA.....	94
4.1	Introduction.....	94
4.2	^{13}C NMR of doubly labelled GpA _{VG} peptide reconstituted using the detergent removal method.....	95
4.2.1	1D ^{13}C ssNMR spectra of doubly labelled (GpA _{VG}) peptide.....	96
4.2.2	Secondary shift analysis.....	101

4.2.3	1D ^{13}C ssNMR of GpA _{VG} prepared using the co-solubilisation method.....	104
4.2.4	2D ^{13}C - ^{13}C DARR ssNMR.....	106
4.2.5	Relating cross peaks observed to existing GpA structure	111
4.3	Alternative peptide labelling scheme.....	113
4.3.1	ssNMR of singly labelled GpA peptides.....	115
4.3.2	Singly labelled GpA build-up curves.....	118
4.4	^{15}N GpA ssNMR.....	121
4.4.1	1D ^{15}N ssNMR of singly labelled GpA sample	121
4.4.2	2D ^{15}N - ^{13}C ssNMR of singly labelled GpA sample.....	123
4.5	Summary.....	129
4.6	References.....	131
5	SOLID STATE NMR ANALYSIS OF BPV E5.....	133
5.1	Introduction.....	133
5.2	1D ^{13}C ssNMR of BPV E5 labelled at leucine and phenylalanine (BPV E5 _{LF})	134
5.2.1	2D ^{13}C - ^{13}C ssNMR of BPV E5 _{LF}	139
5.3	1D ^{13}C ssNMR of BPV E5 labelled at tyrosine and phenylalanine (BPV E5 _{FY})	144
5.3.1	Optimisation of experimental parameters in 1D CP experiments	149
5.3.2	Optimisation of experimental parameters: temperature	150
5.3.3	Optimisation of experimental parameters: contact time.....	153
5.3.4	2D ^{13}C - ^{13}C ssNMR of BPV E5 labelled at phenylalanine & tyrosine (BPV E5 _{FY}) ...	156
5.4	ssNMR characterisation of DMPC/cholesterol lipid membranes.....	160
5.4.1	Characterisation of lipid membranes by static ^{31}P ssNMR	160
5.4.2	Characterisation of lipid membranes by ^1H MAS ssNMR	165
5.5	Summary.....	170
5.6	References.....	173

6	SOLUTION NMR ANALYSIS OF BPV E5	175
6.1	Introduction	175
6.2	Solution NMR analysis of BPV E5 _{V2}	178
6.2.1	Solution NMR analysis of BPV E5 _{V2} in TFE	179
6.3	Solution NMR analysis of BPV E5 _{V2} in bicelles	182
6.3.1	Solution NMR of BPV E5 _{V2} in q=0.33 bicelles	186
6.3.2	3D HSQC TOCSY and NOESY analysis of BPV E5 _{V2}	188
6.4	Solution NMR analysis of BPV E5 _{LF} in bicelles	195
6.5	Characterisation of DMPC/DHPC bicelles	197
6.5.1	Analysis of bicelles by ³¹ P solution NMR	198
6.5.2	Analysis of bicelles by DLS	199
6.6	Summary	201
6.7	References	204
7	DISCUSSION AND FUTURE WORK	206
7.1	Discussion	206
7.2	References	210

LIST OF FIGURES

Figure 1.1 Cartoon representation of a detergent micelle	4
Figure 1.2 Schematic representation of a nanodisc	10
Figure 1.3 Cartoon representation of a bicelle.....	12
Figure 1.4 Single spin angular magnetic moment	17
Figure 1.5 Zeeman splitting of energy levels.....	19
Figure 1.6 Orientation and precession of nuclear spins	20
Figure 1.7 Magnetisation and rotating frame of reference	21
Figure 1.8 Free Induction Decay (FID).....	23
Figure 1.9 2D NMR experiment.	25
Figure 1.10 Illustration of rotor assembly in a magic angle spinning experiment.....	27
Figure 1.11 Dipolar Assisted Rotational Resonance (DARR) pulse sequence.....	31
Figure 1.12 Schematic of the magnetisation transfer within a protein in a DARR experiment.	31
Figure 1.13 Structural model of the Glycophorin A TM domain dimer	33
Figure 1.14 Peptide sequence encoded for by Type I BPV E5 gene	33
Figure 1.15 Schematic of the symmetrical model of the BPV E5 dimer	35
Figure 2.1 Schematic of ATR-FTIR setup	53
Figure 3.1 3D structure of the alpha-helical homodimer transmembrane Glycophorin A (GpA)	63
Figure 3.2 Ball and stick representation of the GpA homodimer interface.....	64
Figure 3.3. Isotopically labelled GpA TM domain peptide sequences.....	65
Figure 3.4 Molecular model of BPV E5 transmembrane protein generated from CHI	66
Figure 3.5 Molecular model of selected E5 dimer interfacial regions generated using CHI...68	
Figure 3.6 Isotopically labelled E5 TM domain peptide sequences.....	68
Figure 3.7 Representative RP-HPLC chromatogram of crude GpA purification	70
Figure 3.8 Representative RP-HPLC chromatogram of crude BPV E5 purification	70
Figure 3.9 RP-HPLC chromatogram of purified BPV E5 peptide	71
Figure 3.10 GpA Deconvoluted ESI-MicroTOF mass spectra	73
Figure 3.11 BPV E5 Deconvoluted ESI-MicroTOF mass spectra.....	74
Figure 3.12 CD spectra of GpA peptide dissolved in TFE	78
Figure 3.13 CD spectra of GpA peptide dissolved in varying detergents.....	78
Figure 3.14 Secondary structure analysis of CD data	79
Figure 3.15 Overview of reconstitution protocols	80
Figure 3.16 Images obtained by TEM using negative staining.....	81
Figure 3.17 CD spectrum of GpA proteoliposome solution obtained by Capillary CD	82
Figure 3.18 CD spectrum of GpA proteoliposome solution prepared using the co- solubilisation method.....	83

Figure 3.19 FTIR spectrum of the GpA peptide in DMPC liposomes containing 5% cholesterol prepared using the detergent removal method	85
Figure 3.20 FTIR spectrum of the GpA peptide in DMPC liposomes containing 5% cholesterol prepared using the co-solubilisation method	86
Figure 3.21 1D ^{13}C spectra obtained from samples prepared by detergent removal (blue) and co-solubilisation (red)	87
Figure 3.22 FTIR spectrum of the E5 _{TM} peptide in DMPC liposomes prepared using the co-solubilisation method.....	88
Figure 3.23 CD spectrum of BPV E5 proteoliposome solution prepared for solid state NMR using the co-solubilisation method	89
Figure 4.1 Chemical structure of valine and glycine.....	96
Figure 4.2 1D proton-decoupled ^{13}C CP-MAS spectrum of doubly labelled GpA _{VG} in DMPC liposomes prepared using the detergent removal method.....	97
Figure 4.3 1D proton-decoupled ^{13}C CP-MAS spectrum of DMPC with 5% cholesterol	98
Figure 4.4 Experimentally derived secondary chemical shifts for GpA.....	104
Figure 4.5 1D proton-decoupled ^{13}C CP-MAS spectrum of doubly labelled GpA _{VG} in DMPC liposomes prepared using the co-solubilisation method.....	105
Figure 4.6 20 ms 2D ^{13}C - ^{13}C DARR spectrum of doubly labelled GpA _{VG} in DMPC liposomes.	107
Figure 4.7 Overlay of 20 & 400 ms 2D ^{13}C - ^{13}C DARR spectra of doubly labelled GpA _{VG} in DMPC liposomes.	108
Figure 4.8 Aliphatic regions of 2D ^{13}C - ^{13}C DARR spectra of doubly labelled GpA _{VG} at short and long mixing times.	110
Figure 4.9 Molecular model of human GpA homodimer	113
Figure 4.10 Diagrammatic form of alternative peptide labelling scheme	115
Figure 4.11 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled GpA _V + GpA _G mixture in DMPC liposomes prepared using the co-solubilisation method.....	116
Figure 4.12 Overlaid 400 ms 2D ^{13}C - ^{13}C DARR correlation spectrum of doubly labelled GpA (GpA _{VG}) vs. singly labelled GpA _V + GpA _G	118
Figure 4.13 Build-up curves for intra-residue and inter-residue GpA cross peaks.....	120
Figure 4.14 1D proton-decoupled ^{15}N CP-MAS spectrum of singly labelled GpA in DMPC liposomes prepared using the co-solubilisation method.....	122
Figure 4.15 2D ^{15}N - ^{13}C TEDOR spectra of singly labelled GpA _V + GpA _G peptides in DMPC/cholesterol liposomes	125
Figure 4.16 1D extracted rows from 2D ^{15}N - ^{13}C TEDOR spectra of valine resonances in singly labelled GpA _V + GpA _G peptides	126
Figure 4.17 1D extracted row from 2D ^{15}N - ^{13}C TEDOR spectra of glycine resonances in singly labelled GpA _V + GpA _G peptides	127
Figure 5.1 Chemical structure of leucine and phenylalanine	135

List of figures

Figure 5.2 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled BPV E5 _{LF} in DMPC liposomes.....	137
Figure 5.3 20 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5 _{LF} in DMPC liposomes.	140
Figure 5.4 400 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5 _{LF} in DMPC liposomes.	141
Figure 5.5 Overlay of 20 and 400 ms 2D ^{13}C - ^{13}C DARR spectra of singly labelled BPV E5 _{LF} in DMPC liposomes.	142
Figure 5.6 Chemical structure of phenylalanine and tyrosine.....	144
Figure 5.7 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled BPV E5 _{FY} in DMPC liposomes.....	147
Figure 5.8 1D proton-decoupled ^{13}C CP-MAS spectra of singly labelled BPV E5 _{FY} at decreasing temperature	151
Figure 5.9 Average peak width of ^{13}C BPV E5 _{FY} resonances as a function of temperature....	152
Figure 5.10 1D proton-decoupled ^{13}C CP-MAS spectra of singly labelled BPV E5 _{FY} at increasing CP contact time	154
Figure 5.11 Graph of 1D ^{13}C resonance signal intensity at increasing cross polarisation (CP) contact times	155
Figure 5.12 50 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5 _{YF} in DMPC liposomes	157
Figure 5.15 400 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5 _{YF} in DMPC liposomes.....	158
Figure 5.16 50 vs 400 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5 _{YF} in DMPC liposomes.....	159
Figure 5.15 Lipid polymorph and phase behaviour	161
Figure 5.16 1D proton-decoupled static ^{31}P spectra of DMPC liposomes with increasing cholesterol concentration at varying temperature.....	162
Figure 5.17 Variation of ^{31}P chemical shift anisotropy of DMPC/cholesterol liposomes with temperature	164
Figure 5.18 1D ^1H spectra of DMPC liposomes with increasing cholesterol concentration at varying temperature.....	165
Figure 5.19 Cartoon of cholesterol intercalating between lipid molecules at high and low temperatures.....	169
Figure 6.1 Peptide sequence of ^{15}N labelled BPV E5 _{V2} peptide	178
Figure 6.2 Molecular model of BPV E5 _{V2} homodimer with ^{15}N labelled amino acids indicated	179
Figure 6.3 1D ^1H spectrum of BPV E5 _{V2} in deuterated TFE.....	180
Figure 6.4 2D ^{15}N - ^1H HSQC spectrum of BPV E5 _{V2} in deuterated TFE	181
Figure 6.5 2D ^{15}N - ^1H HSQC spectra of BPV E5 _{V2} in DMPC/DHPC bicelles at increasing temperature	184

List of figures

Figure 6.6 Extracted planes from 2D ^{15}N - ^1H HSQC spectra at increasing temperatures.....	185
Figure 6.7 Overlay of 2D ^{15}N - ^1H HSQC of BPV E5 _{V2} reconstituted into q=0.33 and q=0.5 bicelles.....	187
Figure 6.8 ^{15}N edited ^1H - ^1H HSQC TOCSY of BPV E5 _{V2} reconstituted into q=0.33 bicelles ..	190
Figure 6.9 ^{15}N edited ^1H - ^1H NOESY-HSQC spectrum of BPV E5 _{V2} reconstituted into q=0.33 bicelles.....	192
Figure 6.10 2D ^{15}N - ^1H HSQC of BPV E5V2 reconstituted into q=0.33 bicelles with tentative amino acid assignments.....	193
Figure 6.11 2D ^{15}N - ^1H HSQC of BPV E5 _{LF} reconstituted into q=0.25 bicelles.....	196
Figure 6.12 ^{31}P solution NMR spectra of q=0.33 bicelles at increasing temperature	198

LIST OF TABLES

Table 2.1 List of isotopically labelled GpA peptides used in this study	42
Table 2.2 List of isotopically labelled BPV E5 peptides used in this study.....	43
Table 2.3 Gradient used for the purification of GpA peptides by RP-HPLC.....	44
Table 2.4 Gradient used for the purification of BPV E5 peptides by RP-HPLC.....	44
Table 2.5 List and properties of detergents used in this study	49
Table 3.1 Table of the shortest inter-helical distance between amino acids at the E5 dimer interface.....	67
Table 3.2 Summary of observed charge states for peptides used in this study analysed by ESI-MS.....	72
Table 4.1 List of previously assigned dimyristoyl- <i>sn</i> -phosphatidylcholine (DMPC) resonances	99
Table 4.2 ¹³ C chemical shift data for labelled GpA Val 80 and Gly 83.....	100
Table 4.3 ¹³ C chemical shift data for secondary species in labelled GpA Val 80 and Gly 83..	101
Table 4.4 Comparison of experimental GpA chemical shifts compared to random coil values	103
Table 4.5 Average distances between valine 80 and glycine 83 carbon atoms in the published GpA homodimer structure.....	111
Table 4.6 ¹⁵ N chemical shift data for labelled GpA Val 80 and Gly 83.....	122
Table 4.7 ¹³ C line widths for GpA Val 80 and Gly 83 extracted from rows of 2D ¹⁵ N- ¹³ C z-filtered TEDOR experiments	128
Table 5.1 ¹³ C chemical shift data for labelled BPV E5 Leu 24 and Phe 28	138
Table 5.2 ¹³ C chemical shift data for labelled BPV E5 Phe 28 and Tyr 31	149
Table 5.3 ¹ H chemical shift data for DMPC natural abundance lipid resonances.....	166
Table 6.1 Chemical shift assignments for ¹⁵ N labelled amino acids in BPV E5 _{v2} reconstituted in DMPC/DHPC bicelles.....	194
Table 6.2 Tentative chemical shift assignments for labelled BPV E5 Leu 24 and Phe 28 in bicelles.....	197
Table 6.3 Hydrodynamic radii of empty and bicelles with inserted BPV E5 _{v2} peptides.....	200

ACKNOWLEDGMENTS

ॐ गं गणपतये नमः

First and foremost I would like to thank my supervisor Ann Dixon, for giving me the opportunity to work on a project in a tough and exciting field of science, her excellent supervision and guidance throughout this project and her input in the preparation of this manuscript have been invaluable. Not only has she been a brilliant supervisor, but also a great friend, someone who I could always go to talk to or share a laugh with and as such will I will always be indebted to her.

I would also like to thank Steven Brown for his input and suggestions in the direction of this study, his knowledge of the field has helped immensely. I would also like to thank Johanna-Becker Baldus, who taught me the many joys of setting up solid-state NMR experiments, Fredrik Romer and all other members of Solid state NMR group who have helped when things have gone wrong. My thanks also to my academic panel Peter Sadler and Pat Unwin for their input into my project and to Janet Crawford for synthesising the peptides used in this study.

I would also like to thank all members of the Dixon group, past and present for their helpful discussion (including Friday Pictionary), friendship and support throughout the four years that I have been here at Warwick, in particular Gemma Warren, Michael Chow, Fay Probert, Esther Martin for making researching towards a thesis enjoyable and especially Maria Tareen for all her help in proofreading and to print this manuscript. Many thanks to all my friends in Chemical Biology and the whole Chemistry department whom I have not named here. I would also like to thank Alison Rodger and the members of her group for the use of equipment and helpful discussion throughout my research.

Finally I would like to thank my family, without whom I would not be here today, their constant love, support and guidance has and always will be immeasurable and for always keeping my freezer stocked with food!

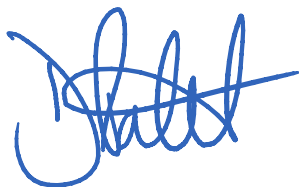
DECLARATION

The work in this thesis is original, and was conducted by the author, unless otherwise stated, under the supervision of Dr Ann M. Dixon (Department of Chemistry) and in collaboration with Professor Steven Brown (Department of Physics).

It has not previously been presented for another degree.

Funding was provided by an EPSRC studentship.

All sources of information have been acknowledged by means of reference.



Dharmesh Kantilal Patel

ABBREVIATIONS

A

aa	Amino acids
ACN	Acetonitrile
AU	Absorbance units
ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy

B

β-OG	β -Octyl glucoside
BMRB	Biological magnetic resonance databank
BPV	Bovine papilloma virus

C

CD	Circular dichroism
CHI	CNS searching of helix interactions
CMC	Critical micelle concentration
CP	Cross polarisation
CP MAS	Cross polarisation magic angle spinning
CSA	Chemical shift anisotropy

D

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
DARR	Dipolar assisted rotational recoupling
dH₂O	Distilled water
DDM	Dodecylmaltoside
DHPC	1,2-Dihexanoyl- <i>sn</i> -glycero-3-phosphocholine
DLS	Dynamic light scattering
DMPC	1,2-Dimyristoyl- <i>sn</i> -glycero-3-phosphocholine
DPC	Dodcylphosphocholine

E

EM	Electron microscopy
ESI-MS	Electrospray ionisation mass spectrometry

F

FT	Fourier transform
-----------	-------------------

G

g	Grams
GpA	Glycophorin A

H

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High performance liquid chromatography
hr	Hours
HSQC	Hetronuclear single quantum coherence
Hz	Hertz

I

IPA	Isopropanol
------------	-------------

K

kDa	kilo Dalton
kHz	Kilo Hertz

L

L_α	Lamellar liquid crystalline phase
L_β	Lamellar gel phase

Abbreviations

L_o Lamellar liquid ordered phase

LD Linear dichroism

LPR Lipid to protein ratio

M

MALDI-TOF Matrix-assisted laser desorption ionisation time of flight

MAS Magic angle spinning

mg Milligram

MHz Mega Hertz

mL Millilitre

mM Millimolar

ms Millisecond

N

NMR Nuclear magnetic resonance

nm Nanometre

NOE Nuclear Overhauser enhancement

NOESY Nuclear Overhauser spectroscopy

O

OCD Oriented circular dichroism

OG octyl-glucoside

P

ppm Parts per million

R

REDOR Rotational echo double resonance

RF Radio frequency

RPM Revolutions per minute

RT Room temperature

S

SDS Sodium dodecyl sulphate

ssNMR Solid state NMR

T

T₁ Longitudinal relaxation time

T₂ Transversal relaxation time

TEDOR Transferred echo double resonance

Abbreviations

TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluoroethanol
TOCSY	Total correlation spectroscopy
TM	Transmembrane
TRIS	tris (hydroxymethyl) aminomethane



UV	Ultraviolet
-----------	-------------



w/v	Weight per volume
------------	-------------------

Greek symbols

ϵ	Extinction coefficient
λ	Wavelength
μg	Micro gram
μL	Micro litre
μM	Micro molar

SUMMARY

Membrane proteins represent over a third of all proteins encoded for by the human genome and play a vital role in the functionality of the cell, by controlling a vast number of cellular processes. With over half of pharmacological drugs targeting membrane proteins, their importance is not to be underestimated. Yet the number of three-dimensional membrane protein structures reported to date falls well short of that of their water soluble counterparts. This discrepancy can directly be attributed to the difficulties involved in studying membrane protein structure due to their hydrophobic nature, resulting in a number of challenges in the production and purification of protein, whilst requiring the use of a suitable membrane mimetic upon extraction from their native membrane.

Solid state NMR (ssNMR) as a technique for studying membrane protein structure is well placed in being able to obtain structural information for membrane proteins in “native-like” lamellar bilayer environments but there are challenges involved in preparing suitable samples for analysis. As there is no “one suit fits all” method for preparing membrane protein samples for ssNMR analysis, conditions that result in fully reconstituted protein, that also allow for high resolution structural analysis have to be trialled.

This study presents work on sample preparation methods for the reconstitution of the small alpha helical transmembrane (TM) proteins, using the well characterised TM protein Glycophorin A (GpA) as a model peptide. Established biophysical and NMR techniques were used to characterise DMPC lipid embedded peptides prepared using two reconstitution techniques. The limited site specific labelling at key positions of the GpA homodimer was used to evaluate the feasibility of using similar sample preparation and labelling schemes when applied to that of the Bovine Papillomavirus E5 (BPV E5) TM protein, for which no solved three-dimensional structure exists. Characterisation of the DMPC membranes into which membrane proteins were reconstituted was also conducted. To compliment ssNMR analysis of BPV E5, preliminary work on the use of fast tumbling isotropic bicelles to study membrane protein structure by solution NMR is also presented.



INTRODUCTION

1.1 Membrane proteins

An integral part of a biological cell is the lipid membrane. In addition to sequestering the contents of the cell, the lipid membrane provides an interface through which the cell can interact with its external environment. The properties of the lipid membrane are influenced by membrane proteins which span across the lipid bilayer. Membrane proteins account for over 30% of the proteins expressed by the human genome (Wallin and von Heijne 1998) and play a vital role in controlling a vast number of cellular functions such as; cell signalling, signal transduction and the trafficking of molecules across the cell membrane. Membrane proteins therefore play a key role in the functionality of the cell and the organism as a whole, as defects in their function can be associated to many diseases and causes of tumorigenesis in eukaryotes (Sanders and Nagy 2000; Sanders and Myers 2004; Aperia 2007). It is for these reasons in particular that membrane proteins are of such significant interest due to their huge potential as pharmaceutical targets in the development of novel therapies against a range of diseases. Currently it is estimated that up to 50% of current drug pharmacological drugs target membrane proteins for their action (Russell and Eggleston 2000), with G-protein coupled receptors (GPCRs) representing the most popular membrane protein target (Russell and Eggleston 2000). Therefore obtaining structural information for membrane proteins is highly beneficial in characterising their functionality and towards designing more effective drugs that are more specific towards their membrane protein targets.

Despite their importance, the number of three-dimensional structures determined for membrane proteins to date remains relatively small when compared to that of soluble proteins, comprising less than one in a hundred structures that have been deposited in the Protein Data Bank (PDB) (www.pdb.org, (Berman, Westbrook et al. 2000)). The membrane protein data bank (MPDB) (Raman, Cherezov et al. 2006) lists only 407 unique membrane protein structures that have been solved and deposited to date and as such membrane proteins are vastly under-represented. Although the rate at which structures are determined increases year by year exponentially, the number of membrane protein structures falls well below the tens of thousands of soluble structures that are available.

1.1.1 Challenges when studying membrane proteins

The relatively small number of membrane protein structures determined to date can be directly attributed to the experimental difficulties involved in working with membrane proteins due to their hydrophobic nature. With their domains inserted in lipid bilayers, the production of membrane proteins for characterisation presents a challenge in comparison to soluble proteins that can be readily solubilised and expressed using typical over expression techniques and purification techniques for studying protein *in vitro*. As membrane proteins natively exist in a non-polar environment of the lipid bilayer, upon extraction from the membrane the three-dimensional structure is typically lost upon solubilisation, as membrane proteins will aggregate in solution unlike water soluble proteins, making techniques for characterisation of the three-dimensional structure a challenge unless a suitable membrane mimetic is used (Warschawski, Arnold et al. 2011). The areas of particular challenge when preparing hydrophobic membrane proteins for structural analysis can be identified as explained in the following section with regards to; the expression, purification, solubilisation and eventual reconstitution of the protein into a suitable mimetic system.

1.1.2 Expression

Membrane proteins are typically expressed at low levels in the native host system and this is therefore a problem for structural studies as high resolution structure determination methods typically require high concentrations (milligram quantities) of

protein. Therefore overexpression of the membrane protein being studied is typically employed. Overexpression of functional eukaryotic membrane proteins is a challenge (Tate 2001) as the prokaryotic host systems commonly employed used for expression of soluble proteins, such as *Escherichia coli* (*E. coli*) typically do not result in high expression yields due to a number of contributing factors. Of the factors that can result in poor yield, most are attributed to the differences between prokaryotic and eukaryotic cellular systems. Unlike soluble proteins which are accumulated in the cytoplasm when expressed in bacteria, membrane proteins are targeted for insertion into membranes due to their hydrophobic nature and therefore the lipid composition of bacterial membranes, which differ considerably from that of eukaryotic membranes, can result in an inhospitable environment for the expressed protein to be inserted into leading to aggregation or toxicity resulting in cell death. Prokaryotic expression system also typically lack the necessary cellular host cell machinery for post translation modification of expressed proteins and therefore modifications such as glycosylation of the expressed protein are not possible and can result in lack of functionality. Prokaryotic expression systems also lack the appropriate chaperones for correct folding and therefore can lead to incorrectly folded or aggregated protein being produced. In some cases where expression of the target membrane protein is high, the expressed protein may be targeted to inclusion bodies containing aggregated protein (Wagner, Baars et al. 2007). Whilst it common place to solubilise and refold soluble proteins isolated from inclusion bodies, the hydrophobic nature of membrane proteins this is generally much harder in practice.

Other expression systems such as yeast (Sreekrishna, Brankamp et al. 1997) and insect cells, that offer eukaryotic expression systems which allow for post translational modification whilst not as popular as bacterial expression systems are now more common place for membrane protein expression. Cell Free Expression systems have also recently been developed (Jermutus, Ryabova et al. 1998; Sawasaki, Ogasawara et al. 2002) that circumvent complications arising from membrane protein toxicity, refolding or reconstitution into lipids. Using cell lysates and a mixture of tRNA, enzymes, nucleotides and amino acids, specifically labelled proteins can be prepared on an mg scale. Additionally the

use of synthesised peptides, using solid phase chemistry is also a viable option for the preparation of short hydrophobic membrane proteins.

1.1.3 Solubilisation

Following successful production for the target membrane protein the next challenge is the solubilisation of the expressed protein from the host system followed by purification. Solubilisation of the target membrane protein requires extraction from the host membrane, this step is usually conducted with the use of detergents. Detergents are amphipathic (containing both polar (water-soluble) and nonpolar (not water-soluble)) molecules that consist of a polar head group and a hydrophobic tail. When placed in an aqueous solution at a concentration above that of their corresponding critical micelle concentration (CMC), they spontaneously form spherical micellar structures (**Figure 1.1**).

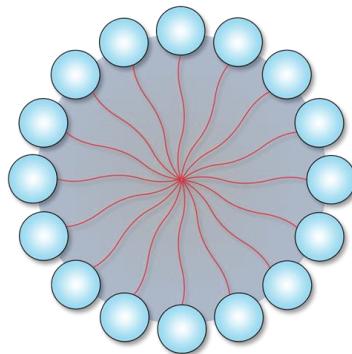


Figure 1.1 Cartoon representation of a detergent micelle

Schematic representation of a detergent micelle, the polar head group (blue balls) generally occupy a larger area than the hydrophobic tails (red lines) and therefore have a cone like shape, these individual detergent molecules then pack together above the critical micelle concentration (CMC) to form detergent micelles.

These micellar structures can be used to shield the hydrophobic membrane protein from the aqueous environment and mimic the native lipid bilayer. The choice of detergent used for solubilisation is crucial in maintaining the structure of the membrane protein but also its function. Whilst detergents can be exchanged through processes such as dialysis,

thought also has to be given to the compatibility with the purification process and subsequent structure determination method with which the membrane protein is to be studied. Proteins are typically solubilised with detergent, whilst maintaining some of the original host membrane, in which the protein was expressed in order to surround the protein and prevent it from aggregating. No one detergent is universal in for successful solubilisation and therefore a detergent screen is generally required to identify the most suitable detergent that will solubilised the protein in an unaggregated state.

1.1.4 Purification

Protein purity and sample homogeneity are important factors for structural analysis studies, therefore it is important to purify over expressed samples and even those obtained from natural membranes, in order to remove unwanted proteins that may result in sample contamination such as bacterial proteins in the case of over expression. It is important that during the purification process the membrane proteins structural integrity is maintained so as to avoid aggregation, here again the choice of solubilising detergent plays an important role as certain detergents are incompatible with certain purification methods. For example, charged detergents such as Sodium Dodecyl Sulfate (SDS) are incompatible with ion-exchange methods. The high detergent concentrations typically used in the solubilisation stage can also have a negative effect on the purification of membrane proteins and therefore excess detergent is often required to be removed prior to purification. Detergent exchange (or removal) can be achieved either by dialysis methods or by using on column chromatographic techniques such as affinity/ion-exchange chromatography, or size exclusion. These techniques can also be used in the purification of detergent solubilised over expressed protein. A common method, that has been successfully used for many years now is the use of affinity tags such as poly-histidine (His tag) that have been incorporated into the expressed target membrane protein, so as to aid purification using affinity chromatography. Using affinity tags the target protein can be bound to a column, such as a Nickel chelating column (in the case of His-tagged proteins). Another challenge when purifying membrane proteins is to also ensure the functionality of the purified protein is maintained following successful expression and solubilisation, in the case of enzymes, the enzymatic activity can be assayed in order to confirm the functionality of the purified

protein. Once purified, the target membrane protein then needs to be reconstituted into a suitable membrane mimetic system for further structural studies.

1.2 Membrane mimetic systems

The lipid membrane is a complex heterogeneous environment composed not only of phospholipids but also numerous other components such as cholesterol, carbohydrates, glycolipids and membrane proteins that exist in the fluid state, as first proposed by Singer and Nicholson (Singer and Nicolson 1972). Therefore to study the three-dimensional structure of a membrane protein a suitable mimetic, representative of the native membrane in which it is found is generally required. This is not always a possibility in particular when using detergents as they do not form bilayers but rather single layered micelles and as such a chosen membrane mimetic system must serve to relate as closely as possible to that of the native environment of which the membrane protein is found, whilst also allowing for high-resolution structural information to be acquired on the embedded protein. The challenge of finding a suitable mimetic for the solubilisation of membrane proteins is made even harder due to lack of a single universally applicable mimetic, with proteins behaving differently, in different mimetics based upon their biological properties, therefore a typical prerequisite for obtaining structural information for solubilised membrane proteins is the screening of different mimetics in order to find the system most suited for not only the protein being studied but also the characterisation technique being used. Properties of the membrane that need to be considered when identifying a suitable mimetic system include; charge, the acyl chain length and saturation of the acyl chains as these factors dictate the properties of the membrane such; as the propensity to form bilayers, the fluidity, curvature and thickness. Thickness of the membrane mimetic is important as the orientation and tilt of transmembrane helices is affected by the thickness, as having a thinner bilayer due to shorter acyl chains for example, can cause an hydrophobic mismatch, as a result an excessive tilt angle can occur as the transmembrane helices try to accommodate their hydrophobic regions within the membrane of the mimetic system.

1.2.1 Detergent micelles

Of all the methods for solubilising membrane proteins for solution NMR, detergent micelles are the most commonly used membrane mimetic for the study of membrane protein structure and function (Kang and Li 2011). Detergent micelles are commonly used to solubilise hydrophobic membrane proteins and can be screened in order to identify suitable detergents that can also preserve native structure and biophysical function. Since detergent micelles have a much smaller overall diameter than that of liposomes, typically ~3 - 5 nm (30 - 50 Å) (Warschawski, Arnold et al. 2011) this allows the protein embedded micellar complex to tumble faster, making detergent micelles suitable for studying membrane proteins by solution NMR and also X-ray crystallography, allowing for high resolution structural information to be obtained. Although detergent micelles are much smaller and tumble faster on the NMR time scale than large lipid complexes, the smaller diameter of micelles in comparison to that of liposomes can often result in curvature stress (i.e. lateral pressure on the surface of the micelle) and increased lateral pressure on the embedded membrane protein, causing minor or in some cases major alterations to the protein fold as it tries to accommodate itself within the micellar structure (Cross, Sharma et al. 2011). In addition to causing curvature stress and lateral pressure to embedded proteins, detergent molecules are not representative of the native lipid bilayer environment in which membrane proteins are found, making them less than ideal membrane mimetic systems (Poget and Girvin 2007).

1.2.2 Liposomes

Of all membrane mimetic systems available, the system that most resembles the native environment in which membrane proteins are found are liposomes, also referred to as lipid vesicles. Liposomes are membranes made up of phospholipids, molecules that are made up of a phosphate head group, a diglyceride and an organic molecule such as a choline. Like detergents, they are mostly amphiphatic with polar and non-polar ends and therefore when placed in an aqueous solution will spontaneously form liposomes. These lipid molecules form bilayers as found in native membranes and can be made up of either a single lipid type or a more complex composition, through the use of membrane extracts. In both plant and animal membranes the most commonly occurring lipid is

phosphatidylcholine (PC) which has a polar choline head group and can make up to 60% of the membrane of cellular organelles such as the Golgi apparatus and Endoplasmic reticulum (van Meer 1998). Phosphatidylethanolamine (PE) is the second most abundant phospholipid found in cell membranes (~30 %) with an ethanolamine head group is a neutral (zwitterionic) lipid and is the major component of microbial membranes. Phosphatidylinositol (PI), with an inositol head group is another common lipid found in plant and animal cell membranes (~25 %) and also plays a role in cell signalling. Other lipids commonly found in the cell membrane include phosphatidylserine (PS) (~ 10%), an acidic lipid due to its serine head group and also Sphingomyelin (SM) a Sphingolipid derived from sphingosine (van Meer, Voelker et al. 2008). The type of lipid used can influence the properties of the lipid bilayer and thereby the liposomes that they makeup. As mentioned earlier in the chapter the length and saturation of the lipid acyl chains together with the type of head group dictate properties of the membrane such the fluidity, curvature and thickness. Membranes formed of lipids with phosphatidylethanolamine (PE) head groups, will tend to form curved membranes, as the smaller head group results in a more conical shaped molecule and therefore forms bilayers with negative curvature. Membranes formed of lipids with more cylindrical shaped molecules such as with phosphatidylcholine (PC) head groups will form flat, planer bilayers as the molecules stack together laterally (Frolov, Shnyrova et al. 2011). Lipid preparations can be used in order to reconstitute membrane proteins in a range of vesicle sizes dependent upon the nature of the lipids and sample preparation methods used. Liposomes can vary in size from 1-2 nm to over 300 nm in diameter. Various preparation methods can be used such as ultra-sonication in order to form small unilamellar vesicles (SUV) or freeze-thaw methods to form large multilamellar vesicles (LMV).

Although liposomes are more biologically relevant membrane mimetics, their large size results in slow tumbling making them unsuitable for analysis of membrane proteins by solution NMR but ideal for solid state NMR methods. Although static NMR spectra are generally broad with poor resolution, interactions between the embedded protein and the lipid membrane can be monitored through ^{31}P and ^2H NMR to gain information about dynamics and lipid/protein interactions (Marius, de Planque et al. 2012). Using magic angle

spinning the resolution of spectra obtained can be greatly improved. MAS-NMR has been used to solve or provide information about the three-dimensional structures of a number of membrane proteins including the viral Influenza M2 channel protein (Wang, Kim et al. 2001), Glycophorin A (GpA) (Smith, Jonas et al. 1994) in DMPC bilayers. One of the issues with using liposomes for the solubilisation of membrane proteins is the oversimplification of the membrane environment, as using a single lipid component in artificial membranes is not as representative of the native bilayer as using more complex mixtures that more closely resemble the native membrane. The use of more complex mixtures to more closely resemble the native membrane, whilst appears more idea, can result in complications due to the complex phase diagrams for mixed lipid membrane systems that can cause issue when working at certain temperatures whilst studying the structure of the embedded membrane protein and due to the increase in unwanted background signals that may interfere and complicate the signals of interest when using techniques such as NMR.

1.2.3 Amphipols and nanodiscs

Methods for the solubilisation of hydrophobic membrane proteins, that allow for analysis by solution NMR whilst also more closely resembling the native membrane environment have recently been developed (Warschawski, Arnold et al. 2011). These membrane mimetic systems include amphipols, a family of synthetic amphipathic polymers developed to replace detergent molecules for membrane protein solubilisation (Tribet, Audebert et al. 1996). Amphipols contain a number of hydrophobic chains that surround solubilised membrane proteins with a greatly reduced number of molecules in non-micellar form, which reduces the viscosity of the solution that can lead to slow tumbling and can be tuned to specific length and charge. Nanodiscs (Bayburt and Sligar 2003) are another recently developed membrane mimetic system that more closely resembles the native environment of a membrane bilayer (Denisov, Grinkova et al. 2004). Nanodiscs are composed of small portions of phospholipid bilayers that have been stabilised by the addition of stabilising amphipathic helical membrane scaffold proteins (MSPs) (**Figure 1.2**). These scaffold proteins result in the formation of disc-shaped soluble portions of

biomembranes, with a typical diameter of ~10 nm and a thickness of ~ 4nm (Warschawski, Arnold et al. 2011) suitable for the solubilisation of membrane proteins. With an increased lateral diameter this makes nanodiscs more suitable than detergent micelles for studying membrane protein structure due to reduced curvature (Lyukmanova, Shenkarev et al. 2008).

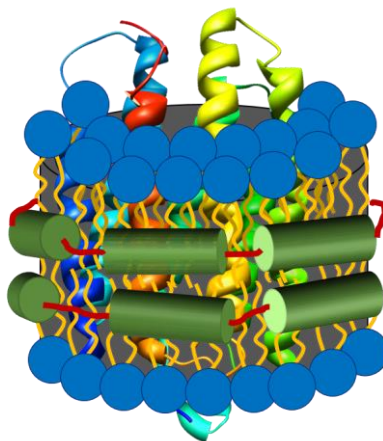


Figure 1.2 Schematic representation of a nanodisc

Representation of a nanodisc with an embedded membrane protein. A small section of a lipid bilayer (blue) surrounding the embedded membrane protein forms a disc like structure which is held in place by helical membrane scaffold proteins (MSP) shown in green.

1.2.4 Bicelles

Although newer membrane mimetic systems such as nanodiscs have been developed, in the last ten years the most popular lipid based membrane mimetic for studying small and multi-spanning membrane proteins by solution NMR are lipid bicelles (Marcotte and Auger 2005) whilst larger alignable bicelles have also been used with ssNMR. Bicelles (**bilayered micelles**) are disc shaped lipid aggregates prepared from a mixture of long chain phospholipids and detergents or short chain phospholipids. The morphology of bicelles depends upon the ratio of long to short chain lipids, total phospholipid concentration and temperature (Gabriel and Roberts 1984; Sanders and Prosser 1998). The long chain phospholipids form a central planar bilayer, with the short chain detergents/lipids forming a rim around the circumference that shields the long chain hydrophobic lipid tails from water

(Figure 1.3). The planar structure with an increased lateral diameter and thickness of ~ 4 nm (40 Å) resembling that of a native bilayer (Luchette, Vetman et al. 2001) and the presence of natural lipids, which more closely mimic *in vivo* membranous structures, make bicelles a more attractive membrane mimetic system for the solubilisation of membrane proteins. Bicelles can typically be produced in two different sizes, with small isotropic bicelles being more suitable for high resolution solution NMR studies due to their rapid tumbling in solution (Vold, Prosser et al. 1997) and larger bicelles that are magnetically alignable with the magnetic field of the spectrometer and can be used for studying orientation and crossing angles of embedded membrane proteins (Sanders and Schwonek 1992; Sanders and Prosser 1998). Larger magnetically alignable bicelles are also suitable for solid state NMR analyses, providing high resolution data for structural assignment (De Angelis, Nevzorov et al. 2004). Whilst the classical description of isotropic bicelles is that of disc like shape with a DMPC bilayer closed by DHPC molecules at the rim, is the accepted morphology for smaller bicelles below the phase transition temperature T_m of DMPC, the structure of larger bicelles formed with higher amounts of long chain to short chain ratios, above the T_m of DMPC is debated. Evidence suggests that at higher temperatures, these larger magnetically alignable bicelles fuse together to form large lamellar bilayer sheets of DMPC with perforated holes that are lined with short chain DHPC molecules (Nieh, Raghunathan et al. 2004), this bicelle morphology is referred to as the Swiss cheese model. At even higher temperatures bicelles have been observed to form 'worm-like' micelles in the liquid crystalline phase (Harroun, Koslowsky et al. 2005).

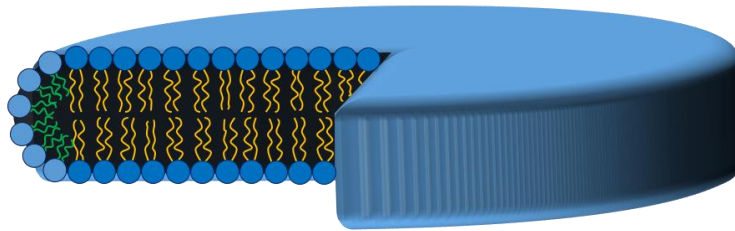


Figure 1.3 Cartoon representation of a bicelle.

Representation of a discoid like lipid bicelle, made up of long chain lipid molecules (typically DMPC) represented with orange acyl chains at the centre and short chain detergent molecules (typically DHPC) represented with green hydrophobic tails at the outer rim.

1.3 Methods for structure determination

Although the number of membrane protein structures determined to date is relatively low, recent advances in technology such as Synchrotron sources for X-ray crystallography, high field NMR and high resolution electron microscopy have led to increased knowledge in the area of membrane protein biochemistry. Each technique presents its own set of advantages and disadvantages.

1.3.1 X-ray crystallography

X-ray crystallography is currently the most popular method for obtaining the three-dimensional structure of membrane proteins, accounting for over 80% of all membrane protein structures deposited at the PDB (Berman, Westbrook et al. 2000; Raman, Cherezov et al. 2006). Structures obtained by X-ray crystallography are typically of high resolution, much more so than of those obtained using other structure determination methods, although a prerequisite is often the requirement of diffraction quality crystals that can be hard to grow and often require screening of a large number conditions in order to find those in which crystal growth occurs. Analysis of membrane proteins using X-ray crystallography also presents difficulties as preparation of crystals requires the solubilisation of membrane proteins prior to crystallisation, whilst maintaining the structural integrity of the solubilised protein. Typically membrane proteins are solubilised in detergent, but the

presence of detergent when crystallising can prevent the formation of crystal contacts and can also result in distorted structures due to curvature stress induced by the small diameter of detergent micelles. X-ray crystallography is also more suited to larger multi-spanning membrane proteins as larger proteins crystallise more readily than smaller proteins, as the larger the protein the greater the surface area for which crystal contacts can form that are required for crystal growth. For smaller proteins the surface area is greatly reduced thereby reducing the possibility of forming electrostatic contacts between unit cells in a crystal thereby reducing the possibility of forming diffraction quality crystals for analysis, although in recent years the crystallography of membrane proteins in lipid membranes has become viable by the growth of crystals in lipidic mesophases (also referred to as Lipid cubic phase (LCP) crystallisation (Landau and Rosenbusch 1996; Cherezov 2011; Caffrey, Li et al. 2012). Lipids in the LCP form highly curved bilayers that form cubic lattice structures. First used to obtain high resolution structural data for bacteriorhodopsin (Landau and Rosenbusch 1996) this method has now been used to crystallise a variety of bitopic membrane proteins including GPCRs and helical proteins (Cherezov, Rosenbaum et al. 2007; Jaakola, Griffith et al. 2008; Wu, Chien et al. 2010). Therefore whilst X-ray crystallography is positioned to provide high resolution structures at atomic resolution, the strategies involved in producing viable crystallisation conditions can often result in structures that differ vastly from their native form (Cross, Arseniev et al. 1999).

1.3.2 Solution NMR

Next to X-ray crystallography, solution NMR is the second most popular method for the determination of membrane protein three-dimensional structure. Solution NMR techniques allow the possibility of exploring protein-protein and protein-ligand interactions in a dynamic environment without the need for protein crystals. Unlike X-ray crystallography solution NMR can be used in order to study the dynamics of a protein rather than in a fixed state within a crystal. Typically, membrane proteins are studied in detergent micelles, but this is not ideal as detergent micelles have been identified to cause curvature stress to embedded membrane proteins altering the structure of proteins when compared to bilayer

bound samples (Chou, Kaufman et al. 2002). In addition the resolution and sensitivity of the spectra obtained by solution NMR is strongly affected by how fast a molecule tumbles in solution. Due to rapid random tumbling rates of small molecules on the NMR timescale (of typically nano/picoseconds), orientation dependant anisotropic interactions are averaged out to zero resulting in sharp resonances. As the size of the molecule in solution increases as does the rate of tumbling, as such larger molecules therefore have much slower tumbling rates and correspondingly shorter spin-spin (transverse) T_2 relaxation times due to enhanced spin-spin interactions. Shorter T_2 relaxation times result in line broadening and intensity loss as a result in of the reduction in the sensitivity of complicated multi-pulse NMR experiments that often use long delays for the necessary coherence transfer steps between nuclei. Therefore whilst the protein of interest to be studied by NMR may be small, once reconstituted into a detergent micelle or lipid embedded environment, the size of the complex typically becomes much larger, thereby tumbling much more slowly (Watts and Spooner 1991; Marcotte and Auger 2005), and in the case of proteins embedded in lipid vesicles can exceed the size limit (100 kDa) for this technique resulting in severe line broadening and signal intensity loss.

1.3.3 Solid state NMR (ssNMR)

In contrast to solution NMR, solid state NMR (ssNMR) is not restricted by an upper molecular weight size limit, thereby making it a powerful technique for the study of higher molecular weight proteins and of those embedded in lipid environments. Therefore, unlike solution NMR, this allows for the study of membrane proteins in hydrated lipid bilayers, thereby representing a more “native-like” environment and therefore resulting in high resolution structures in more biologically relevant confirmations. Typically high resolution ssNMR spectra have been obtained for microcrystalline or amyloid fibril samples with high structural homogeneity (Bockmann and Meier 2010), whereas spectra obtained in hydrated bilayers are often much broader in comparison. As such ssNMR has become an invaluable tool for obtaining structural information of membrane proteins under physiological conditions (Watts, Burnett et al. 1999) such as Gramicidin (Ketchem, Hu et al. 1993), Influenza

M2 (Cady, Mishanina et al. 2009; Luo, Cady et al. 2009) and human Phospholamban (Verardi, Shi et al. 2011).

In spite of the numerous advantages of studying membrane protein structure by ssNMR, a number of disadvantages are also associated with the technique, in particular the resolution of ssNMR spectra recorded in comparison to solution NMR spectra is greatly reduced. Inherently, ssNMR spectra are much harder to interpret and assign as they are much more complicated in their nature when compared to solution NMR spectra as the full effect of orientation-dependant (anisotropic) interactions are still present and observed in the spectra obtained. These anisotropic interactions which are normally averaged out in solution for small rapid tumbling molecules are still present in solid samples, as molecular motions are restricted with rotational correlation times much longer than in solution i.e. nanosecond to seconds. In addition, in solid samples, molecules are simultaneously present in a large number of orientations. The presence of anisotropic interactions in solid samples results in considerable broadening of resonances (typically 0.5 – 2 ppm) in comparison to those recorded in solution NMR spectra, often leading to complicated, difficult to resolve. These anisotropic interactions that are still present in ssNMR experiments are listed below.

1.3.4 Chemical shift anisotropy (CSA)

In solid (powder) samples all molecular orientations are present in random orientations, with random distribution, which gives rise to powder patterns in recorded spectra. These powder patterns arise as a result of each different molecular orientation (with respect to the applied magnetic field B_0) having its own chemical shift, with each orientation giving rise to its own (sharp) resonance. The overlapping of these individual resonances gives rise to the broad axially symmetrical unresolvable powder patterns typically observed in solid samples.

1.3.5 Dipolar coupling

When spins from individual nuclei come into close contact with each other in a sample, the magnetic field generated by each nucleus can act through-space to have an influence on the spin energy of neighbouring nuclei. Dipolar coupling can occur between

nuclei of the same type i.e. homonuclear dipolar coupling between ^{13}C and ^{13}C , or between nuclei of different atoms i.e. heteronuclear dipolar coupling between ^{13}C and ^1H . In ssNMR dipolar coupling can lead to a detrimental effect on the spectra recorded due to signal broadening and decay of magnetisation through effect such as dipolar truncation (Hodgkinson and Emsley 1999). Dipolar interactions can also be useful for probing distance measurements between nuclei due to its r^3 dependence, where r is the inter nuclear distance, therefore making the strength of the dipolar coupling between nuclei a good measure for the distance between them. Dipolar coupling can and also for signal enhancement through transfer of magnetisation such as through cross polarisation, as described further in **Section 1.5.2**.

1.3.6 Quadrupolar coupling

For nuclei which have spin greater than $\frac{1}{2}$, these nuclei are referred to a quadrupolar, i.e. they possess a nuclear electronic quadrupole moment. The quadrupole moment in the nucleus arises from the non-spherical distribution of charge within. This quadrupole moment is, in addition to the magnetic dipolar moment as possessed by nuclei with spin $\frac{1}{2}$. Electric quadrupoles interact with electric field gradients therefore such nuclei not only interact with the applied and local magnetic fields, but also with any electric field gradients within the nucleus, thereby affecting the nuclear spin energy levels. The strength of the interaction depends upon the magnitude of the quadrupole moment. The effect of the quadrupolar interaction is observed as a substantial broadening of the observed ssNMR spectra.

Therefore, whilst the broad lines in ssNMR spectra contain a wealth of information regarding structure and dynamics of the protein (Warschawski, Traikia et al. 1998), they typically have detrimental effects on spectra recorded, obscuring peaks and leading to the poor resolution of resonances observed, making it difficult to resolve individual resonances due to spectral overcrowding. Therefore in order to avoid spectral crowding, a number of elaborate labelling schemes can employed, such as the use of 1,3- ^{13}C labelled glycerol when expressing membrane proteins, that give rise to specific cross peak patterns that make the assignment process easier.

1.4 NMR Theory

1.4.1 Spin

Nuclear magnetic resonance (NMR) spectroscopy can be used to exploit the intrinsic property possessed by certain atomic nuclei of "spin" **Figure 1.4** and the fact that those nuclei that possess spin can undergo transitions between nuclear spin energy levels defined by the Zeeman quantised spin angular momentum, in order to determine the magnetic environment of the nuclei.

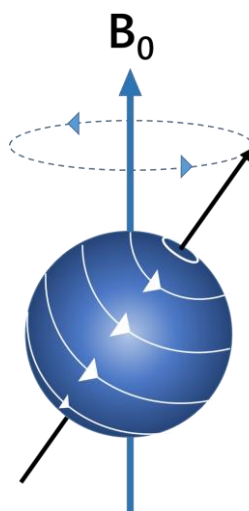


Figure 1.4 Single spin angular magnetic moment

An isolated nucleus has spin angular momentum and a precession frequency dependent only upon the type of nucleus (gyromagnetic ratio) and upon the strength of the applied magnetic field B_0

Quantum mechanically sub atomic particles have intrinsic angular momentum, which is characterised by its spin quantum number I , where I is an integer or half integer. In a number of atoms e.g. ^{12}C , ^{16}O , spins are paired, cancelling each other out and result in an atom with no overall spin ($I=0$) and are therefore NMR-inactive. If the number of neutrons and the number of protons in an atom are even then the nucleus has no spin, if the number of neutrons in addition to protons is odd then the nucleus will have half-integer spin (e.g. $I=1/2, 3/2, 5/2$). If both protons and neutrons are odd in number, then the nucleus has integer

spin (e.g. $I = 1, 2, 3$). The spin quantum number dictates magnitude of the angular magnetic moment (μ) for the nucleus as related by **Equation 1**.

$$\mu = \gamma I \quad (1)$$

Where γ is the gyromagnetic ratio, this ratio is characteristic unique to each nuclear isotope. When an external magnetic field of strength B_0 is applied, a spinning nucleus will align its nuclear magnetic moment in a quantised number of orientations, given by $2I+1$, either with or against the magnetic field. For example in the case of a nucleus with spin $I = \frac{1}{2}$, only one transition is possible between two energy levels, the energetically favourable, aligned with the applied magnetic field (spin $m = +\frac{1}{2}$) also referred to as α and a higher energy orientation aligned against the applied magnetic field (spin $m = -\frac{1}{2}$) or β orientation (**Figure 1.5**). The splitting of these nuclear spin energy levels by the applied magnetic field is due to the Zeeman Effect, responsible for an identical splitting between each of the nuclear spin energy levels and is the dominant interaction in NMR. The distribution of these nuclear spin energy levels is governed by the Boltzmann distribution (**Equation 2**), where N values are the number of nuclei in each respective spin state, ΔE the energy level difference between each spin state ($= \hbar\omega_0$) the strength of the external magnetic field, k the Boltzmann constant and T the absolute temperature.

$$\frac{N_{upper}}{N_{lower}} = e^{-\left(\frac{\Delta E}{kT}\right)} \quad (2)$$

NMR spectroscopy is used to probe the transitions between these two energy levels, since the difference between the two energy states and the number of nuclei in each state is minimal in comparison to techniques such as UV/Vis spectroscopy, where the energy separation between each state is comparatively larger, this renders NMR an insensitive technique. This means that NMR experiments give rise to relatively weak signals, requiring a

large number of spectra to be collected in order to obtain adequate signal to noise. To improve sensitivity of NMR experiments, higher field strength magnets, that result in an increase in the size of the applied magnetic field, as well as decreasing the temperature at which experiments are conducted are two of the most common methods used in order to increase the Boltzmann distribution between the two energy states, thereby resulting in higher sensitivity.

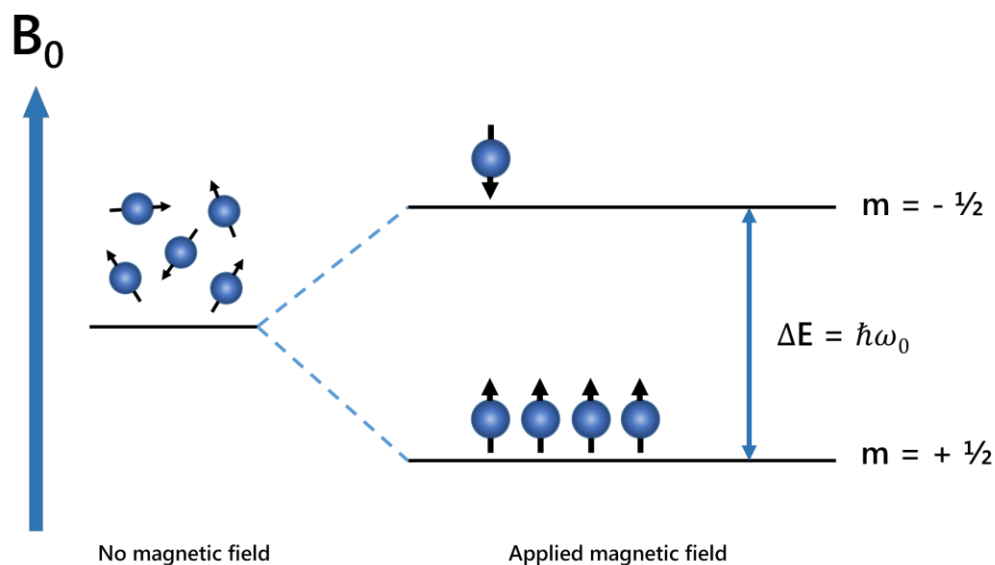


Figure 1.5 Zeeman splitting of energy levels

Illustration of Zeeman splitting of energy levels for a nucleus with spin $I = \frac{1}{2}$. There are two possible Eigenstates for such nuclei, a low energy $+ \frac{1}{2}$ and a high energy $- \frac{1}{2}$ state, The transition energy between the two states is related to the strength of the applied magnetic field and is described by the equation $\Delta E = \hbar\omega_0$. When a magnetic field is applied there is a Boltzmann distribution of the spins in two states

1.4.2 Magnetisation

When a nucleus with angular momentum is placed in a magnetic field of strength B_0 , the nucleus will precess about the external magnetic field z-axis, due to the torque generated by the interaction of the nuclear angular momentum with the magnetic field, at the nuclide specific Larmor frequency defined by **Equation 3**.

$$\omega_0 = -\gamma B_0 \quad (3)$$

Where ω is the angular velocity of precession, γ is the gyromagnetic ratio of the nuclide, B_0 is the external, static magnetic field and the sign denotes the direction of motion of the magnetic moment about the static field. In a real sample there are a large number of nuclear spins in the system, all precessing about the z-axis (**Figure 1.6**), this ensemble of nuclei of the same kind, would precess around the applied magnetic field with a common angular frequency, giving rise to a net resultant bulk magnetisation, i.e. the sum of all individual magnetic moments, M_0 .

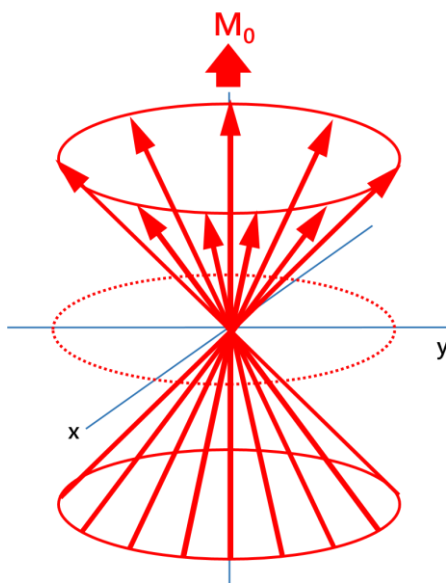


Figure 1.6 Orientation and precession of nuclear spins

An ensemble of nuclear magnetic moments (red arrows) distributed across the two spin energy levels, will precess around the magnetic field B_0 without any phase coherence. The excess population of spins in the lower energy level gives rise to the bulk nuclear spin magnetisation M_0 .

1.4.3 R.F pulses and the Rotating Frame

Once a sample is placed in the NMR spectrometer, the nuclear spins will align with the applied magnetic field B_0 where they will reach equilibrium with a bulk magnetisation. Nuclear magnetic resonance occurs when electromagnetic radiation with a frequency matching with the Larmor frequency of the nuclei of interest is applied in order to perturb the magnetic moments (as shown in **Figure 1.7**), causing the nuclei to change its spin state. Using the rotating frame coordinate system (x' , y' and z), where the external magnetic field is considered to be along the z -axis, a radio frequency (RF) pulse of radiation is applied along the x' -axis (B_1), this will impose a torque on the bulk magnetisation vector (M_0) resulting from the precessing nuclei (B_1 and M_0 are stationary and at right angles in the rotating frame of reference).

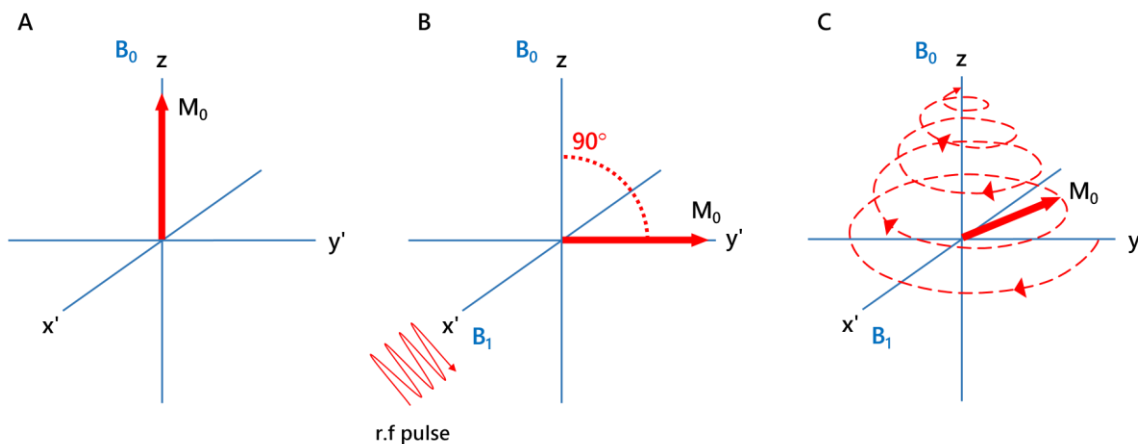


Figure 1.7 Magnetisation and rotating frame of reference

(A) Bulk magnetisation M_0 at thermal equilibrium precessing in the z -axis in the magnetic field B_0 . (B) Application of a 90° ($\pi/2$) RF pulse in the B_1 x' -axis causes perturbation of the bulk magnetisation vector, rotating it into the y' -axis. (C) Once the RF pulse is switched off, the bulk magnetisation vector relaxes back gradually to thermal equilibrium, precessing around the y,z plane giving rise to the FID signal recorded in the detector.

This torque will be perpendicular to the external field vector and rotates the bulk magnetisation vector M_0 away from its equilibrium position along B_0 and around into the y' -axis with the appropriate pulse length. The spins will then begin to precess about the $x'y'$

plane (in which the detector is placed), causing a very weak oscillating voltage to be induced in the coil surrounding the sample, which is responsible for the observed NMR signal. The angle of rotation (θ), is dependent upon the gyromagnetic ratio of the nucleus γ , the amplitude of the B_1 RF pulse and upon the length time (t) that the pulse is applied, (as given in **Equation 4**), the example of a 90° or $\pi/2$ pulse is given in **Figure 1.7 B**. Once the applied RF pulse is switched off, the system undergoes relaxation, with the bulk magnetisation vector gradually returning back to its thermal equilibrium state along the z-axis, this is referred to as longitudinal (T_1) relaxation. Whilst T_1 relaxation describes the decay of signal back into the z-axis, transverse (T_2) relaxation characterises the relaxation in the transverse (xy) plane as a result of excited nuclei exchanging spins or losing coherence with each other and is therefore also referred to as 'spin-spin' relaxation. This relaxation back to thermal equilibrium causes the signal observed in the receiver coil to decay with time.

$$\theta = \gamma B_1 t \quad (4)$$

This decaying signal consists of contributions from all the different target nuclei in the sample. This signal cannot be digitised directly due to its high frequency and is therefore mixed with a lower frequency signal in order to produce an interferogram of low frequency. This interferogram is digitised and is responsible for the detected NMR response called a Free Induction Decay (FID) (**Figure 1.8**). Fourier transformation of this FID yields a frequency domain spectrum.

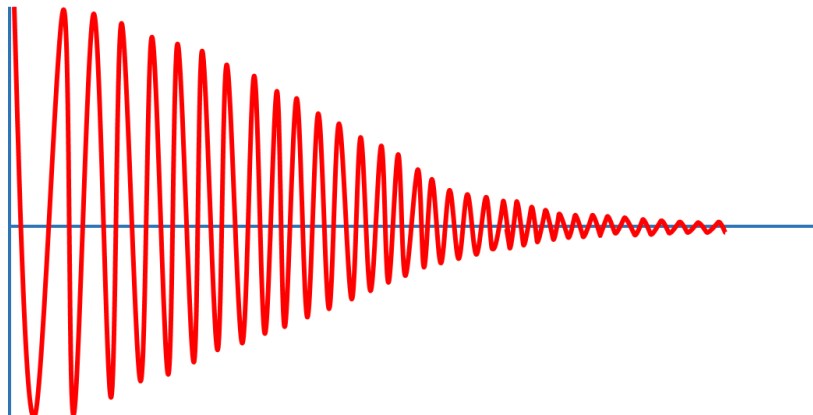


Figure 1.8 Free Induction Decay (FID)

A Free Induction Decay (FID) signal produced as a result of relaxation of excited spins in an NMR experiment, Fourier transformation of which gives rise to a frequency domain spectrum.

1.4.4 Chemical Shift

NMR frequencies are set as chemical shifts in units of parts per million (ppm). Nuclei in the sample are eminent by differences in their chemical shift, as nuclei of different elements have different gyromagnetic ratios, which yield signals at different frequencies dependant on magnetic field strength. However nuclei of the same element can also give rise to signals of different frequencies. This is due to the Larmor frequency of a nucleus being dependent upon its local electronic environment. The applied external magnetic field induces currents within the electron cloud surrounding the nucleus, thereby inducing a small local magnetic field. This induced local magnetic field is opposed to the applied magnetic field and in effect shields the nucleus from the applied field. By this effect nuclei of the same type in different chemical environments experience slightly different local magnetic fields and therefore have different Larmor frequencies. Although this induced local field is considerably small in comparison to the applied external field, the difference in Larmor frequency between nuclei in different environments is still measurable as relative signals in the NMR spectrum. A reference compound is typically chosen and the difference

between the position of the signal of interest and that of the reference is termed the chemical shift.

NMR chemical shift values are typically expressed in ppm rather than in Hz, so as to remove the dependency of the magnetic field strength (operating frequency) at which the sample was recorded using **Equation 5**.

$$\delta_{ppm} = \frac{\omega_0 - \omega_{ref}}{\omega_{ref}} \times 10^6 \quad (5)$$

Where ω_{ref} is the Larmor frequency for a reference compound, e.g. tetramethylsilane (TMS) for ^1H or 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) for ^{13}C . This results in a scale that is independent of the applied external magnetic field B_0 used for the experiment. As nuclei in a complex sample such as a protein will experience different chemical environments resulting in differing Larmor frequencies, the resultant dispersion one dimensional (1D) NMR spectrum obtained can be much more complicated to interpret in comparison to simple single molecule samples due to the crowding and overlapping of signals, therefore multidimensional (two or three dimensional 2D/3D) experiments are typically used in order to simplify assignment of the chemical shifts recorded.

1.4.5 Two dimensional (2D) NMR

A number of homo-nuclear 2D experiments exist, all of which share the same basic principle. There are four steps to a 2D NMR experiment; in the first step (called the preparation time), all nuclei in the sample are excited simultaneously using one or more pulses, creating magnetisation in the xy plane. The resulting magnetisation is allowed to evolve during the evolution period (t_1) during which time encoding is carried out in the indirect dimension (F_1) and the chemical shift of the first nucleus (e.g. A) is recorded. This is

followed by a mixing time (t_{mix}) in which magnetisation is allowed to transfer to the second nucleus (e.g. B) using either a combination of pulses or delay periods. The final stage is detection where the chemical shift of the second nucleus is recoded in the direct dimension (F_2). Raw data from a 2D NMR experiment consists of a series of FIDs, each one acquired with a slightly longer t_1 duration than the previous; the 2D data can then be Fourier transformed in order to produce a 2D NMR spectrum. By selectively labelling specific residues of interest, the presence of cross-peaks off the diagonal in the 2D spectrum (**Figure 1.9**) obtained can provide information about which amino acids are close together in space. Such experiments make use of the structural information contained within through-space dipolar couplings by applying a recoupling sequence during the mixing time.

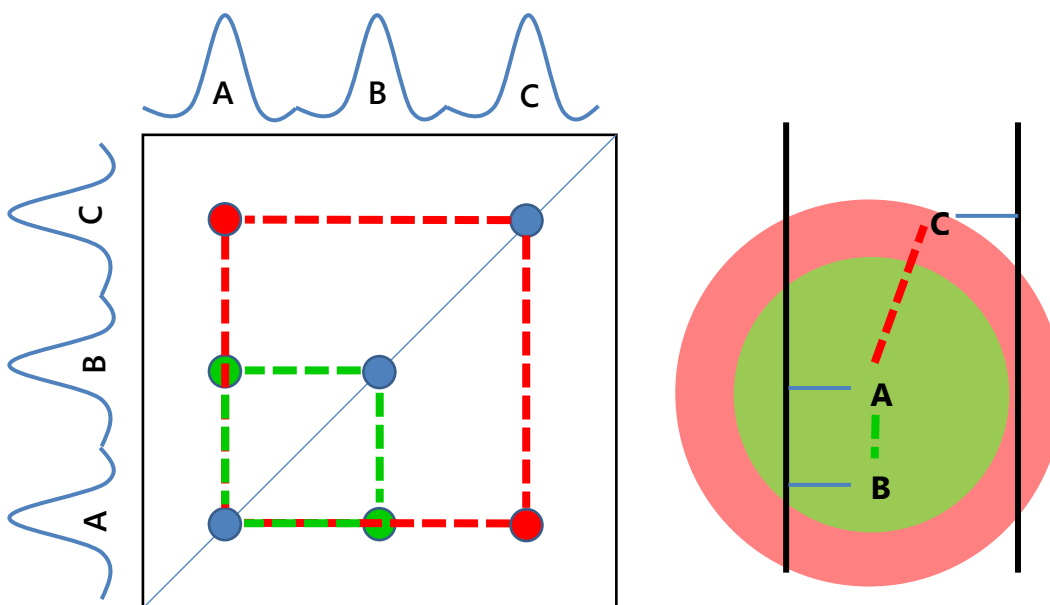


Figure 1.9 2D NMR experiment.

Representation of a 2D NMR spectrum shown on left, and a representation of peptide dimer with isotopic labels indicated as A, B and C. At short mixing times (green circle) magnetisation travels only far enough to see short range correlations (green cross peaks). At longer mixing times (red circle) the magnetisation is allowed to travel further and as such longer range correlations are observed (red cross peaks).

1.5 Solid state NMR

In solid state NMR (ssNMR) the samples being analysed are typically *powder samples* i.e. samples consisting of many crystallites in random orientations. The anisotropic nuclear spin interactions that affect ssNMR spectra such as; CSA, dipolar and quadrupolar coupling (as detailed in **Section 1.3.4-6**) are all dependent upon the orientation of crystallite orientations. In solution NMR this is not a problem as anisotropic interactions are averaged to zero, as molecules in solution exhibit Brownian motion, tumbling faster than the frequency of the interactions and reorientation of molecules occurs on the NMR timescale i.e. pico/nanoseconds (dependent upon the size of the molecule). By contrast, ssNMR experiments concentrate on solid samples with restricted molecular motion (i.e. milliseconds to seconds on the NMR timescale) where fast molecular tumbling does not exist, therefore the effect of anisotropic interactions, that are averaged out in solution NMR are still present in spectra obtained by ssNMR. Additionally, in solid samples molecules exist simultaneously in a number of orientations and as a result ssNMR spectra exhibit broad features, powder patterns composed of a superposition of signals from a number of different orientations. As a result ssNMR spectra contain significant structural information that is typically lost in solution NMR, although due to the lack of resolution in ssNMR spectra obscures any information that the spectrum may contain. Additionally due to the strong dipolar coupled network of spins, protons are generally not the preferred nuclei for observation in solid state NMR as these interactions typically result in broadened spectra. For ssNMR of protein samples structural details are primarily obtained from low- γ and dilute $I=3/2$ spins i.e. ^{13}C and ^{15}N . The detection of low γ nuclei typically requires isotope enrichment for sensitivity enhancement. Therefore in order to obtain high resolution NMR spectra of solid samples, techniques such magic angle spinning (MAS) have been developed, whilst methods such as cross-polarisation (CP) and high power proton decoupling, in order to remove ^1H - ^{13}C and ^1H - ^{15}N couplings that are too strong to remove by sample rotation alone, are commonly used to improve sensitivity.

1.5.1 Magic angle spinning (MAS)

Magic angle spinning (MAS), first introduced by Andrew and Lowe (Andrew, Bradbury et al. 1959) is routinely used in solid state NMR to mimic the rapid isotropic tumbling that occurs in solution that does not occur in solid samples via mechanical rotation of the sample. MAS is essential for obtaining high resolution ssNMR spectra and is used to remove the effects of chemical shift anisotropy and to assist the removal of heteronuclear dipolar coupling. As shown in **Figure 1.1** typically the sample filled rotor is spun about its axis at $\beta = 54.74^\circ$, "the magic angle", with respect to the magnetic field B_0 , and is rotated at a rate ω_r . Spinning at the magic angle simulates the rapid isotropic tumbling that occurs in solutions, which averages the molecular orientation dependence of the transition frequencies to zero on the NMR timescale. Chemical shielding and dipolar coupling both contain a molecular orientation dependence term of the form $3\cos^2\theta - 1$ with respect to the magnetic field B_0 . In solution the rapid tumbling averages this component to zero thereby removing their effects, whereas in ssNMR this angular component can be averaged to zero by mechanically rotating and spinning at the magic angle.

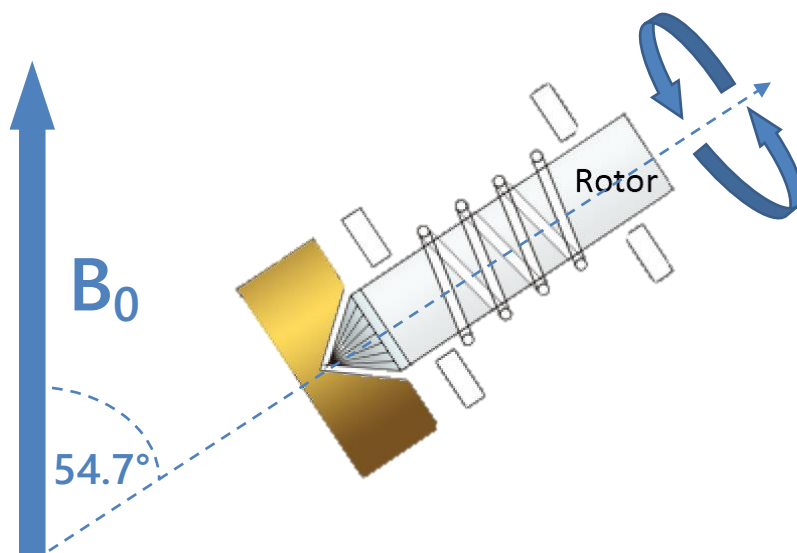


Figure 1.10 Illustration of rotor assembly in a magic angle spinning experiment

The rotor containing sample is spun at 54.7° (magic angle) with respect to the magnetic field B_0 , the sample can be cooled whilst spinning in order to reduce the mobility of the packed sample which can lead to improved line width.

In order for MAS to reduce a powder pattern to a single line at the isotropic chemical shift, the rate of sample spinning ω_r must be around 3 to 4 times greater than the frequency of the interaction to be able to average it out to zero. With current spinning speeds reaching approximately 70 kHz, it is possible to suppress CSA and dipolar interactions but not quadrupolar couplings. When spinning slower and line broadening is not completely suppressed, this results in the appearance of *spinning sidebands*, set at the spinning rate apart, in addition to the peak at the isotropic chemical shift. This limits MAS when applied to high gamma nuclei such as protons, which have dipolar coupling in excess of 100 kHz, as such spinning speeds are much more difficult to achieve.

1.5.2 Cross polarisation (CP)

In addition to MAS, cross polarisation (CP) (Urbina and Waugh 1973) is another commonly used technique central to ssNMR. CP is used in order to improve the signal intensity of dilute spins (S-spins) by transferring magnetisation from abundant nuclei (I-spins), with high gyromagnetic ratios such as ^1H , and ^{31}P to dilute nuclei with low gyromagnetic ratios, for example ^{13}C or ^{15}N . The use of CP also overcomes the problem of long T_1 (longitudinal) relaxation times, when using direct polarisation of dilute spin nuclei with low abundance and low- γ , which typically result in longer experimental times, as using CP, the repetition time of the experiment is dictated by the shorter relaxation rate of the more abundant, high- γ nuclei (i.e. ^1H) allowing for shorter experimental times. The CP transfer is mediated by the dipolar interaction between ^1H and X spins and uses the principle that when two nuclei are brought close together, magnetisation tends to travel from more highly polarised nuclei to nuclei with lower polarisation. In the CP experiment, the field strengths of both nuclei (e.g. ^1H and ^{13}C) are set to the Hartmann-Hahn Matching condition (**Equation 6**) the dipolar spin pair system is considered in a doubly rotating frame, in which ^1H and ^{13}C magnetisation precess about B_0 . CP begins with an initial 90° RF pulse creating transverse ^1H magnetisation along the $-y$ axis, which is maintained along the rotating frame (*spin-locked*) using a *contact pulse* over a contact time period. Spin locking involves applying a continuous pulse along this axis which is known as the *spin-lock field* and is labelled $B_1(^1\text{H})$. Simultaneously a 90° contact pulse is also applied to the ^{13}C channel to maintain magnetisation in the transverse frame $B_1(^{13}\text{C})$. The amplitude of the two contact

pulses is then carefully set in order to achieve the Hartmann-Hahn Matching condition. Magnetisation transfer then occurs between the spins (i.e. from ^1H to ^{13}C nuclei) provided that there is dipolar coupling between them.

$$\gamma_H B_1(1H) = \gamma_X B_1(1X) \quad (6)$$

1.5.3 Proton decoupling

MAS is sufficient to remove the effect of dipolar coupling for the majority of nuclei, but for nuclei with large gyromagnetic ratios such as ^1H , which are also in high abundance, the dipolar coupling is too strong to be removed by MAS alone, resulting in broad peaks and low resolution and over complicated spectra. This is a problem particularly with protein samples where the large number of high γ protons interfere with the low abundance ^{13}C , ^{15}N low γ nuclei typically being detected. It is therefore necessary to remove these residual ^1H - ^{13}C dipolar couplings in order to improve the resolution of the spectra being collected, therefore high power decoupling is typically used. The simplest method of decoupling involves the continuous application of an high power RF pulse (usually 80 – 100 KHz) during acquisition of the FID. This high power RF pulse causes the ^1H spins to flip between the α - β , parallel-antiparallel state. As the influence of ^1H dipolar coupling on ^{13}C is determined by the z-component of the ^1H magnetization, if the α - β state transition is faster than the ^1H - ^{13}C dipolar coupling frequency, then the ^1H - ^{13}C dipolar coupling will be averaged to zero. The continuous application of high power RF can result in sample heating and also damage to the NMR probe and therefore more sophisticated decoupling methods such as Two-pulse phase modulation (TPPM) (McGeorge, Alderman et al. 1999) and SPINAL-64 have been developed which significantly improve the efficiency of decoupling and in which the decoupling pulses are divided up into a sequence of pulses with varying phase, which result in narrower line widths than continuous wave (CW) decoupling.

1.5.4 Recoupling

Whilst MAS is used to average away the anisotropic interactions that lead to broad unresolvable spectra, the wealth of structural information regarding orientation and inter nuclear distance is also removed from the resulting spectra obtained. Dipolar coupling for example can provide distance constraint information regarding nuclei that are close together in space due to its r^3 dependence as a factor of inter nuclear distance between dipolar coupled nuclei. Therefore it can be beneficial to run experiments that simultaneously provide high resolution whilst also selectively reintroducing (recoupling) anisotropic interactions. A number of different techniques have been developed for recoupling dipolar interactions that typically incorporate the application of RF pulse trains to disrupt the averaging of these interactions due to sample rotation.

1.5.5 Dipolar Assisted Rotational Recoupling (DARR)

Dipolar Assisted Rotational Recoupling (DARR) (Takegoshi, Nakamura et al. 2001) is a recoupling mechanism that uses a combination of mechanical rotation of the sample and the ^{13}C - ^{13}C dipolar interaction in order to reintroduce homo-nuclear dipolar couplings that are averaged out by MAS (**Figure 1.11**). This is achieved by the application of continuous low power ^1H 'recoupling' pulses with the frequency ω_1 , satisfying the rotary resonance condition $\omega_1 = \omega_R$ during the mixing period t_{mix} , where ω_R is the MAS spinning frequency. DARR has been previously applied to a number of membrane protein samples and has been used to observe long range, through-space correlations (Crocker, Patel et al. 2004; Abdine, Verhoeven et al. 2010).

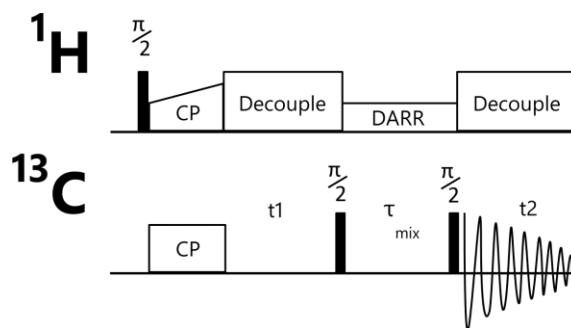


Figure 1.11 Dipolar Assisted Rotational Resonance (DARR) pulse sequence.

Pulse sequence for 2D Dipolar Assisted Rotational Resonance (DARR) experiment. A 90° pulse on the proton channel flips ^1H magnetisation into the X-Y plane which is transferred to ^{13}C via a ramped CP pulse and ^1H decoupling applied during the t_1 evolution period. Following the evolution period, magnetisation is placed along the z-axis with a 90° pulse and mixing occurs longitudinally, with a low power ^1H DARR recoupling pulse equal to that of the MAS sample spinning speed is applied. A final 90° pulse is applied and high power ^1H decoupling applied during detection.

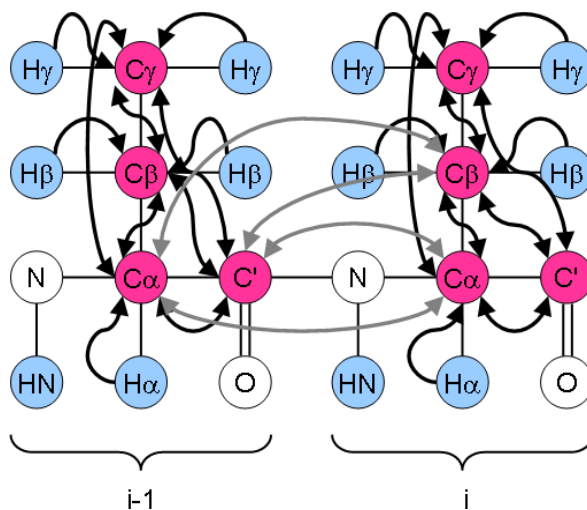


Figure 1.12 Schematic of the magnetisation transfer within a protein in a DARR experiment.

Pink atoms indicate those observed and blue atoms indicate those through which magnetisation flows between residues i and $i-1$ in a protein chain. Black arrows show magnetisation transfer and grey arrows indicate magnetisation that is normally only observed through longer mixing times (image taken from http://www.protein-nmr.org.uk/spectra_ssnmr.html).

1.6 Glycophorin A (GpA) as a model TM

The helical TM erythrocyte protein Glycophorin A is an example of a membrane protein that has been very well-characterised using both solution (MacKenzie, Prestegard et al. 1997) and ssNMR (Smith, Jonas et al. 1994; Liu, Crocker et al. 2003). The incorporation of both ^{13}C and ^{15}N isotopic labels into the protein has been used to obtain structural information at atomic resolution for the GpA transmembrane (TM) domain dimer (Smith, Jonas et al. 1994) and has provided information such as inter-helical interactions in the lipid bilayer (Smith, Song et al. 2001). The structural model obtained in the bilayer by ssNMR was shown to differ from that obtained by solution NMR with a decrease in crossing angle (Smith, Eilers et al. 2002), where the crossing angle is the angle between the two helix axes when projected onto their plane of contact (Chothia, Levitt et al. 1981). This demonstrates, as with other comparatively studied membrane proteins (Cross, Tian et al. 1999; Cady, Mishanina et al. 2009) that protein conformation is altered in a micelle when compared to its more native-like lipid bilayer environment. ssNMR was used to provide the first direct measurements of helix to helix contacts in the GpA TM domain by demonstrating close packing of side chain methyl groups of valine and glycine residues across the dimer interface (see **Figure 1.13**) (Smith and Bormann 1995).

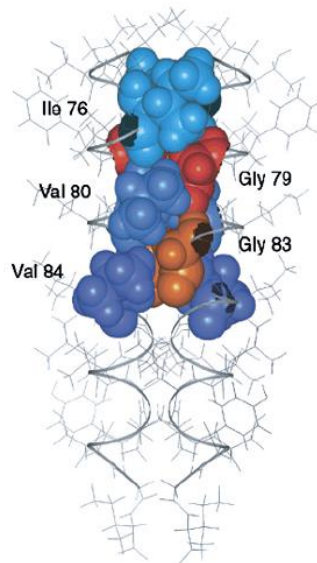


Figure 1.13 Structural model of the Glycophorin A TM domain dimer

Structural model of GpA homodimer obtained from ^{13}C - ^{13}C distances, showing close packing of valine and glycine residues at the dimer interface. Reproduced from Smith *et al* 2001

1.7 Bovine Papillomavirus E5 (BPV E5)

Viruses have evolved a number of mechanisms in order to transform host cells so as to ensure continual replication of infected host cells. Bovine Papillomavirus BPV is a non-enveloped double stranded DNA virus belonging to the family *Papillomaviridae* which causes fibropapillomas in cattle (Lancaster and Olson 1982). BPV efficiently transforms healthy cells via the action of the product of the E5 gene (**Figure 1.14**) producing what is currently the smallest known viral oncoprotein: the E5 protein. E5 is a highly hydrophobic 44 amino acid type II transmembrane oncoprotein which forms a 7 kDa membrane-spanning homodimer (Schlegel, Wade-Glass et al. 1986).

NH_2 -MPNLWFLFLGLVAAMQLLLLLLFLFFLVYWDHFECSTGLPF-COOH

Figure 1.14 Peptide sequence encoded for by Type I BPV E5 gene

The BPV E5 protein is composed of an α -helical amino terminal transmembrane domain followed by a 14 amino acid hydrophilic carboxyl terminal region (Burkhardt, DiMaio et al. 1987). As suggested by its hydrophobic character, BPV E5 has been shown by immunoelectron microscopy to localise primarily to the Golgi and plasma membranes as a type II membrane protein, with its carboxyl-terminus facing the Golgi lumen (Burkhardt, Willingham et al. 1989). Expression of E5 alone, in the absence of other viral gene products has been shown to be sufficient to transform immortalised murine cells. Mutational studies (DiMaio, Guralski et al. 1986; Windisch, Hoffmann et al. 2010) have demonstrated E5 to be fairly insensitive to single point residue mutations, but have identified a small number of key residues (i.e. a glutamine at position 17 (Klein, Polack et al. 1998), aspartate at position 33, and two carboxyl-terminal cysteine residues) that are thought to be critical to maintaining its function.

The oncogenic transmembrane E5 protein has been shown to exhibit its cell transforming ability by targeting the platelet derived growth factor beta receptor (PDGF β R) (Petti, Nilson et al. 1991), a receptor tyrosine kinase. In virally transformed cells, E5 binding to PDGF β R (Drummond-Barbosa, Vaillancourt et al. 1995) causes sustained activation of the receptor leading to uncontrolled cell growth and tumour formation. Although other E5 targets such as epidermal growth factor receptor (EGFR) (Martin, Vass et al. 1989) and the 16K subunit of vacuolar H⁺ ATPase (Goldstein, Finbow et al. 1991) have been suggested, E5 has been identified as binding preferentially to PDGF β R (Petti and DiMaio 1994; Freeman-Cook, Edwards et al. 2005).

No three-dimensional structure for E5 exists to date; however, molecular dynamics simulation, polarised IR and mutagenesis studies (Surti, Klein et al. 1998; Adduci and Schlegel 1999) have been used to produce a simplified model of the E5 dimer and its complex with PDGF β R. E5 is believed to interact with the transmembrane domain of PDGF β R (**Figure 1.15**), bringing two PDGF β R molecules together and resulting in trans-phosphorylation and activation of the tyrosine kinase domain (Mattoon, Gupta et al. 2001). Although this model exists, alternative models have also previously been suggested (Adduci and Schlegel 1999) proposing that the E5 dimer is asymmetrical.

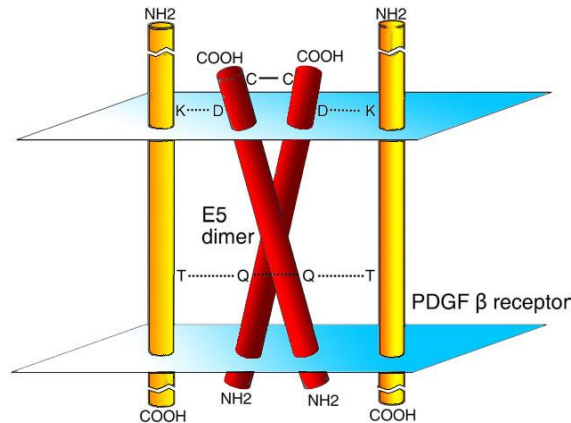


Figure 1.15 Schematic of the symmetrical model of the BPV E5 dimer

Schematic of the BPV E5 homodimer (red) and its interaction with the PDGFβR (yellow), adapted from Talbert-Slagle and Dimaio 2008.

1.8 Aims and objectives of this study

This project aims to evaluate and optimise sample preparation methods for reproducible and reliable reconstitution of small TM proteins for ssNMR and solution NMR analyses. Using a range of biophysical techniques in order to assess sample quality and reconstitution methods, we aimed to identify the ideal conditions to prepare samples for the development of protein solid state and solution NMR methods. In collaboration with the solid state NMR group in the Department of Physics here at Warwick University, we hoped to gain structural information for the Glycophorin A (GpA) homodimer in a lipid bilayer environment using fewer selective uniformly labelled amino acids at the homodimer interface. Once established using GpA, we then hoped to apply these techniques to the Bovine Papillomavirus (BPV) E5 protein, for which no structure exists and which is of biological interest due to its oncogenic properties. In conjunction with ssNMR methods, Using BPV E5, we aimed to evaluate sample preparation conditions for the reconstitution of small TM proteins into bicelles for solution NMR studies. It is hoped that by developing such methods for studying membrane proteins, critical information will be obtained that will

facilitate the development of rationally-designed drugs based upon target protein structure for a wide variety of diseases in which membrane proteins play a key role.

1.9 References

- Abdine, A., M. A. Verhoeven, K. H. Park, A. Ghazi, E. Guittet, C. Berrier, C. Van Heijenoort and D. E. Warschawski (2010). "Structural study of the membrane protein MscL using cell-free expression and solid-state NMR." *J Magn Reson* **204**(1): 155-159.
- Adduci, A. J. and R. Schlegel (1999). "The Transmembrane Domain of the E5 Oncoprotein Contains Functionally Discrete Helical Faces." *J Biol Chem* **274**(15): 10249-10258.
- Andrew, E. R., A. Bradbury and R. G. Eades (1959). "Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation." *Nature* **183**(4678): 1802-1803.
- Aperia, A. (2007). "Membrane transport proteins in health and disease." *J Intern Med* **261**(1): 2-4.
- Bayburt, T. H. and S. G. Sligar (2003). "Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers." *Protein Sci* **12**(11): 2476-2481.
- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne (2000). "The Protein Data Bank." *Nucleic Acids Res* **28**(1): 235-242.
- Bockmann, A. and B. H. Meier (2010). "Prions: En route from structural models to structures." *Prion* **4**(2): 72-79.
- Burkhardt, A., D. DiMaio and R. Schlegel (1987). "Genetic and biochemical definition of the bovine papillomavirus E5 transforming protein." *EMBO J* **6**(8): 2381-2385.
- Burkhardt, A., M. Willingham, C. Gay, K. Jeang and R. Schlegel (1989). "The E5 oncoprotein of bovine papillomavirus is oriented asymmetrically in Golgi and plasma membranes." *Virology* **170**(1): 334-339.
- Cady, S., T. Mishanina and M. Hong (2009). "Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding." *J Mol Biol* **385**(4): 1127-1141.
- Cady, S. D., T. V. Mishanina and M. Hong (2009). "Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding." *J Mol Biol* **385**(4): 1127-1141.
- Caffrey, M., D. Li and A. Dukkupati (2012). "Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes." *Biochemistry* **51**(32): 6266-6288.
- Cherezov, V. (2011). "Lipidic cubic phase technologies for membrane protein structural studies." *Curr Opin Struct Biol* **21**(4): 559-566.
- Cherezov, V., D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." *Science* **318**(5854): 1258-1265.
- Chothia, C., M. Levitt and D. Richardson (1981). "Helix to helix packing in proteins." *J Mol Biol* **145**(1): 215-250.

- Chou, J., J. Kaufman, S. Stahl, P. Wingfield and A. Bax (2002). "Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel." *J Am Chem Soc* **124**(11): 2450-2451.
- Crocker, E., A. B. Patel, M. Eilers, S. Jayaraman, E. Getmanova, P. J. Reeves, M. Ziliox, H. G. Khorana, M. Sheves and S. O. Smith (2004). "Dipolar assisted rotational resonance NMR of tryptophan and tyrosine in rhodopsin." *J Biomol NMR* **29**(1): 11-20.
- Cross, T., A. Arseniev, B. Cornell, J. Davis, J. Killian, R. n. Koeppe, L. Nicholson, F. Separovic and B. Wallace (1999). "Gramicidin channel controversy--revisited." *Nat Struct Biol* **6**(7): 610-611; discussion 611-612.
- Cross, T., F. Tian, M. Cotten, J. Wang, F. Kovacs and R. Fu (1999). "Correlations of structure, dynamics and function in the gramicidin channel by solid-state NMR spectroscopy." *Novartis Found Symp* **225**: 4-16; discussion 16-22.
- Cross, T. A., M. Sharma, M. Yi and H. X. Zhou (2011). "Influence of solubilizing environments on membrane protein structures." *Trends Biochem Sci* **36**(2): 117-125.
- De Angelis, A. A., A. A. Nevzorov, S. H. Park, S. C. Howell, A. A. Mrse and S. J. Opella (2004). "High-resolution NMR spectroscopy of membrane proteins in aligned bicelles." *J Am Chem Soc* **126**(47): 15340-15341.
- Denisov, I. G., Y. V. Grinkova, A. A. Lazarides and S. G. Sligar (2004). "Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size." *J Am Chem Soc* **126**(11): 3477-3487.
- DiMaio, D., D. Guralski and J. Schiller (1986). "Translation of open reading frame E5 of bovine papillomavirus is required for its transforming activity." *Proc Natl Acad Sci U S A* **83**(6): 1797-1801.
- Drummond-Barbosa, D., R. Vaillancourt, A. Kazlauskas and D. DiMaio (1995). "Ligand-independent activation of the platelet-derived growth factor beta receptor: requirements for bovine papillomavirus E5-induced mitogenic signaling." *Mol Cell Biol* **15**(5): 2570-2581.
- Freeman-Cook, L., A. Edwards, A. Dixon, K. Yates, L. Ely, D. Engelman and D. Dimaiio (2005). "Specific locations of hydrophilic amino acids in constructed transmembrane ligands of the platelet-derived growth factor beta receptor." *J Mol Biol* **345**(4): 907-921.
- Frolov, V. A., A. V. Shnyrova and J. Zimmerberg (2011). "Lipid polymorphisms and membrane shape." *Cold Spring Harb Perspect Biol* **3**(11): a004747.
- Gabriel, N. E. and M. F. Roberts (1984). "Spontaneous formation of stable unilamellar vesicles." *Biochemistry* **23**(18): 4011-4015.
- Goldstein, D., M. Finbow, T. Andresson, P. McLean, K. Smith, V. Bubb and R. Schlegel (1991). "Bovine papillomavirus E5 oncoprotein binds to the 16K component of vacuolar H(+)-ATPases." *Nature* **352**(6333): 347-349.
- Harroun, T. A., M. Koslowsky, M. P. Nieh, C. F. de Lannoy, V. A. Raghunathan and J. Katsaras (2005). "Comprehensive examination of mesophases formed by DMPC and DHPC mixtures." *Langmuir* **21**(12): 5356-5361.
- Hodgkinson, P. and L. Emsley (1999). "The accuracy of distance measurements in solid-state NMR." *J Magn Reson* **139**(1): 46-59.
- Jaakola, V. P., M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman and R. C. Stevens (2008). "The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist." *Science* **322**(5905): 1211-1217.

- Jermutus, L., L. A. Ryabova and A. Pluckthun (1998). "Recent advances in producing and selecting functional proteins by using cell-free translation." Curr Opin Biotechnol **9**(5): 534-548.
- Kang, C. and Q. Li (2011). "Solution NMR study of integral membrane proteins." Curr Opin Chem Biol **15**(4): 560-569.
- Ketchum, R. R., W. Hu and T. A. Cross (1993). "High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR." Science **261**(5127): 1457-1460.
- Klein, O., G. Polack, T. Surti, D. Kegler-Ebo, S. Smith and D. DiMaio (1998). "Role of glutamine 17 of the bovine papillomavirus E5 protein in platelet-derived growth factor beta receptor activation and cell transformation." J Virol **72**(11): 8921-8932.
- Lancaster, W. and C. Olson (1982). "Animal papillomaviruses." Microbiol Rev **46**(2): 191-207.
- Landau, E. M. and J. P. Rosenbusch (1996). "Lipidic cubic phases: a novel concept for the crystallization of membrane proteins." Proc Natl Acad Sci U S A **93**(25): 14532-14535.
- Liu, W., E. Crocker, D. Siminovitch and S. Smith (2003). "Role of side-chain conformational entropy in transmembrane helix dimerization of glycophorin A." Biophys J **84**(2 Pt 1): 1263-1271.
- Luchette, P. A., T. N. Vetman, R. S. Prosser, R. E. Hancock, M. P. Nieh, C. J. Glinka, S. Krueger and J. Katsaras (2001). "Morphology of fast-tumbling bicelles: a small angle neutron scattering and NMR study." Biochim Biophys Acta **1513**(2): 83-94.
- Luo, W., S. D. Cady and M. Hong (2009). "Immobilization of the influenza A M2 transmembrane peptide in virus envelope-mimetic lipid membranes: a solid-state NMR investigation." Biochemistry **48**(27): 6361-6368.
- Lyukmanova, E. N., Z. O. Shenkarev, A. S. Paramonov, A. G. Sobol, T. V. Ovchinnikova, V. V. Chupin, M. P. Kirpichnikov, M. J. Blommers and A. S. Arseniev (2008). "Lipid-protein nanoscale bilayers: a versatile medium for NMR investigations of membrane proteins and membrane-active peptides." J Am Chem Soc **130**(7): 2140-2141.
- MacKenzie, K., J. Prestegard and D. Engelman (1997). "A transmembrane helix dimer: structure and implications." Science **276**(5309): 131-133.
- Marcotte, I. and M. Auger (2005). "Bicelles as model membranes for solid- and solution-state NMR studies of membrane peptides and proteins." Concepts in Magnetic Resonance Part A **24A**(1): 17-37.
- Marius, P., M. R. de Planque and P. T. Williamson (2012). "Probing the interaction of lipids with the non-annular binding sites of the potassium channel KcsA by magic-angle spinning NMR." Biochim Biophys Acta **1818**(1): 90-96.
- Martin, P., W. Vass, J. Schiller, D. Lowy and T. Velu (1989). "The bovine papillomavirus E5 transforming protein can stimulate the transforming activity of EGF and CSF-1 receptors." Cell **59**(1): 21-32.
- Mattoon, D., K. Gupta, J. Doyon, P. Loll and D. DiMaio (2001). "Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein." Oncogene **20**(29): 3824-3834.
- McGeorge, G., D. W. Alderman and D. M. Grant (1999). "Resolution enhancement in ¹³C and ¹⁵N magic-angle turning experiments with TPPM decoupling." J Magn Reson **137**(1): 138-143.
- Nieh, M. P., V. A. Raghunathan, C. J. Glinka, T. A. Harroun, G. Pabst and J. Katsaras (2004). "Magnetically alignable phase of phospholipid "bicelle" mixtures is a chiral nematic made up of wormlike micelles." Langmuir **20**(19): 7893-7897.

- Petti, L. and D. DiMaio (1994). "Specific interaction between the bovine papillomavirus E5 transforming protein and the beta receptor for platelet-derived growth factor in stably transformed and acutely transfected cells." *J Virol* **68**(6): 3582-3592.
- Petti, L., L. Nilson and D. DiMaio (1991). "Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein." *EMBO J* **10**(4): 845-855.
- Poget, S. F. and M. E. Girvin (2007). "Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better." *Biochim Biophys Acta* **1768**(12): 3098-3106.
- Raman, P., V. Cherezov and M. Caffrey (2006). "The Membrane Protein Data Bank." *Cell Mol Life Sci* **63**(1): 36-51.
- Russell, R. and D. Eggleston (2000). "New roles for structure in biology and drug discovery." *Nat Struct Biol* **7 Suppl**: 928-930.
- Sanders, C. R., 2nd and J. P. Schwonek (1992). "Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR." *Biochemistry* **31**(37): 8898-8905.
- Sanders, C. R. and J. K. Myers (2004). "Disease-related misassembly of membrane proteins." *Annu Rev Biophys Biomol Struct* **33**: 25-51.
- Sanders, C. R. and J. K. Nagy (2000). "Misfolding of membrane proteins in health and disease: the lady or the tiger?" *Curr Opin Struct Biol* **10**(4): 438-442.
- Sanders, C. R. and R. S. Prosser (1998). "Bicelles: a model membrane system for all seasons?" *Structure* **6**(10): 1227-1234.
- Sawasaki, T., T. Ogasawara, R. Morishita and Y. Endo (2002). "A cell-free protein synthesis system for high-throughput proteomics." *Proc Natl Acad Sci U S A* **99**(23): 14652-14657.
- Schlegel, R., M. Wade-Glass, M. Rabson and Y. Yang (1986). "The E5 transforming gene of bovine papillomavirus encodes a small, hydrophobic polypeptide." *Science* **233**(4762): 464-467.
- Singer, S. J. and G. L. Nicolson (1972). "The fluid mosaic model of the structure of cell membranes." *Science* **175**(4023): 720-731.
- Smith, S. and B. Bormann (1995). "Determination of helix-helix interactions in membranes by rotational resonance NMR." *Proc Natl Acad Sci U S A* **92**(2): 488-491.
- Smith, S., M. Eilers, D. Song, E. Crocker, W. Ying, M. Groesbeek, G. Metz, M. Ziliox and S. Aimoto (2002). "Implications of threonine hydrogen bonding in the glycoporphin A transmembrane helix dimer." *Biophys J* **82**(5): 2476-2486.
- Smith, S., R. Jonas, M. Braiman and B. Bormann (1994). "Structure and orientation of the transmembrane domain of glycoporphin A in lipid bilayers." *Biochemistry* **33**(20): 6334-6341.
- Smith, S., D. Song, S. Shekar, M. Groesbeek, M. Ziliox and S. Aimoto (2001). "Structure of the transmembrane dimer interface of glycoporphin A in membrane bilayers." *Biochemistry* **40**(22): 6553-6558.
- Sreekrishna, K., R. G. Brankamp, K. E. Kropp, D. T. Blankenship, J. T. Tsay, P. L. Smith, J. D. Wierschke, A. Subramaniam and L. A. Birkenberger (1997). "Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*." *Gene* **190**(1): 55-62.

- Surti, T., O. Klein, K. Aschheim, D. DiMaio and S. Smith (1998). "Structural models of the bovine papillomavirus E5 protein." *Proteins* **33**(4): 601-612.
- Takegoshi, K., S. Nakamura and T. Terao (2001). "C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR." *Chem Phys Lett* **344**(5-6): 631-637.
- Tate, C. G. (2001). "Overexpression of mammalian integral membrane proteins for structural studies." *FEBS Lett* **504**(3): 94-98.
- Tribet, C., R. Audebert and J. L. Popot (1996). "Amphipols: polymers that keep membrane proteins soluble in aqueous solutions." *Proc Natl Acad Sci U S A* **93**(26): 15047-15050.
- van Meer, G. (1998). "Lipids of the Golgi membrane." *Trends Cell Biol* **8**(1): 29-33.
- van Meer, G., D. R. Voelker and G. W. Feigenson (2008). "Membrane lipids: where they are and how they behave." *Nat Rev Mol Cell Biol* **9**(2): 112-124.
- Verardi, R., L. Shi, N. J. Traaseth, N. Walsh and G. Veglia (2011). "Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method." *Proc Natl Acad Sci U S A* **108**(22): 9101-9106.
- Vold, R. R., R. S. Prosser and A. J. Deese (1997). "Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides." *J Biomol NMR* **9**(3): 329-335.
- Wagner, S., L. Baars, A. J. Ytterberg, A. Klussmeier, C. S. Wagner, O. Nord, P. A. Nygren, K. J. van Wijk and J. W. de Gier (2007). "Consequences of membrane protein overexpression in Escherichia coli." *Mol Cell Proteomics* **6**(9): 1527-1550.
- Wallin, E. and G. von Heijne (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." *Protein Sci* **7**(4): 1029-1038.
- Wang, J., S. Kim, F. Kovacs and T. A. Cross (2001). "Structure of the transmembrane region of the M2 protein H(+) channel." *Protein Sci* **10**(11): 2241-2250.
- Warschawski, D. E., A. A. Arnold, M. Beaugrand, A. Gravel, E. Chartrand and I. Marcotte (2011). "Choosing membrane mimetics for NMR structural studies of transmembrane proteins." *Biochim Biophys Acta* **1808**(8): 1957-1974.
- Warschawski, D. E., M. Traikia, P. F. Devaux and G. Bodenhausen (1998). "Solid-state NMR for the study of membrane systems: the use of anisotropic interactions." *Biochimie* **80**(5-6): 437-450.
- Watts, A., I. J. Burnett, C. Glaubitz, G. Grobner, D. A. Middleton, P. J. Spooner, J. A. Watts and P. T. Williamson (1999). "Membrane protein structure determination by solid state NMR." *Nat Prod Rep* **16**(4): 419-423.
- Watts, A. and P. J. Spooner (1991). "Phospholipid phase transitions as revealed by NMR." *Chem Phys Lipids* **57**(2-3): 195-211.
- Windisch, D., S. Hoffmann, S. Afonin, S. Vollmer, S. Benamira, B. Langer, J. Burck, C. Muhle-Goll and A. S. Ulrich (2010). "Structural role of the conserved cysteines in the dimerization of the viral transmembrane oncoprotein E5." *Biophys J* **99**(6): 1764-1772.
- Wu, B., E. Y. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov and R. C. Stevens (2010). "Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists." *Science* **330**(6007): 1066-1071.

2

MATERIALS AND METHODS

2.1 Reagents and Chemicals

All reagents and chemicals used in this study unless stated otherwise were purchased from Fisher Scientific (Loughborough, Leicestershire, UK) or Sigma-Aldrich (Gillingham, Dorset, UK). All lipids were obtained from Avanti Polar Lipids (Alabaster, AL, USA). All chemicals were of the highest grade available for use in analytical and chemical biology studies and used without any further purification.

2.2 Buffers

All buffers used in this study were prepared from distilled water (dH₂O) (MilliQ) and were filtered through 0.22 µm filters (Millipore). Phosphate buffer solution (20 mM sodium phosphate, 20 mM NaCl, 3 mM NaN₃ at pH 7.4) was used to prepare samples for solid state nuclear magnetic resonance (ssNMR) and circular dichroism (CD) measurements. Phosphate buffered detergent solution (20 mM sodium phosphate, 20 mM NaCl, 3 mM NaN₃ and 34 mM octyl-glucoside (OG) at pH 7.4) was used in the preparation of ssNMR samples using the detergent removal method. HEPES buffer solution (20 mM HEPES, 20 mM NaCl, 3 mM NaN₃ at pH 6.8/7.4) and TRIS buffer (20 mM TRIS, 20 mM NaCl, 3 mM NaN₃ at pH 7.4) were used to prepare samples for ¹H and ³¹P ssNMR experiments. HEPES buffer solution (20 mM HEPES, 20 mM NaCl, 3 mM NaN₃ at pH 6.8/7.4) was used to prepare samples for solution-state NMR experiments.

2.3 Peptide design

2.3.1 Glycophorin A (GpA) peptides

Three peptides derived from the transmembrane (TM) domain of Glycophorin A (GpA) (Jokinen, Andersson et al. 1985) were synthesised with the sequence KKITLIIFGVMAGVIGTILLISYGIKK and containing additional non-native end terminal lysine residues (indicated by underlining). Peptides were N-terminally acetylated and C-terminally amidated to aid solubility (Morozova and Weiss 2010). Peptides were synthesised using solid phase 9-fluorenylmethyl carbamate (Fmoc) chemistry (Carpino and Han 1972) at the Yale University W.M KECK Facility (New Haven, CT, USA). Uniformly $^{13}\text{C}/^{15}\text{N}$ labelled valine and glycine residues were introduced at position 80 (GpA_V), position 83 (GpA_G), or at both positions (GpA_{VG}), and unlabelled peptide was also prepared (GpA_U), as shown in **Table 2.1**.

Table 2.1 List of isotopically labelled GpA peptides used in this study

KKITLIIFG <u>V</u> ⁸⁰ MAGVIGTILLISYGIKK	GpA _V
KKITLIIFGVMAG ⁸³ VIGTILLISYGIKK	GpA _G
KKITLIIFG <u>V</u> ⁸⁰ MAG ⁸³ VIGTILLISYGIKK	GpA _{VG}
KKITLIIFGVMAGVIGTILLISYGIKK	GpA _U

2.3.2 Bovine Papillomavirus E5 peptides

Five peptides corresponding to the TM domain of the BPV E5 protein (Goldstein, Finbow et al. 1991; Mattoon, Gupta et al. 2001) were prepared with the sequence KKKFLGLVAAMQLLLLLFLLFFLVYWDHK containing additional non-native end terminal lysines (indicated by underlining) to aid solubility. These peptides were synthesised in a similar manner to GpA peptides at the Yale University W.M KECK Facility. Uniformly ^{13}C and ^{15}N -labelled Fmoc protected amino acids purchased from Cambridge Isotope Laboratories were incorporated at Leu 24 (BPV E5_L), Phe 28 (BPV E5_F), Tyr 31 (BPV E5_Y),

and unlabelled (BPV E5_U) was also prepared. BPV E5_{V2} was specifically designed for solution NMR studies (the justification of which is given in **Section 3.2**) and was therefore labelled at multiple positions as shown below in **Table 2.2**

Table 2.2 List of isotopically labelled BPV E5 peptides used in this study

KKKFLGLVAAMQLLLLLFL ²⁴ LLFFLVYWDHK	BPV E5 _L
KKKFLGLVAAMQLLLLLFL ²⁸ LVYWDHK	BPV E5 _F
KKKFLGLVAAMQLLLLLFL ³¹ YWDHK	BPV E5 _Y
KKKFLGLVAAMQLLLLLFLFFLVYWDHK	BPV E5 _U
KKKFLG ¹¹ LV ¹³ A ¹⁴ AMQL ¹⁹ LLLFL ²⁴ LLFFLV ³⁰ YWDHK	BPV E5 _{V2}

2.4 Peptide purification using reverse-phase high performance liquid chromatography

Synthetic peptides were supplied as crude reaction products, and as such the crude peptides were subsequently purified by reverse-phase high performance liquid chromatography (RP-HPLC), using protocols previously reported for hydrophobic peptides (Lew and London 1997; Fisher and Engelman 2001). Typically, for each RP-HPLC run, 3-5 mg of crude peptide was weighed out and dissolved in 200 µL trifluoroacetic acid (TFA) (Sigma), 200 µL trifluoroethanol (TFE), 400 µL isopropanol (IPA) and drop wise addition of 70:30 HPLC grade water: isopropanol equilibration buffer to a final volume of 2 mL.

Samples were injected onto a semi-preparative Jupiter C4 5u (300 Å, 250 × 10.0 mm) RP-HPLC column (Phenomenex, Macclesfield, Cheshire, UK) connected to a purpose built two pump HPLC system (Jasco UK, Great Dunmow, Essex, UK). A flow rate of 1.5 mL min⁻¹ was used throughout. A two component gradient (as indicated in **Table 2.3** for GpA and **Table 2.4** for BPV E5) was used in the separation, with solvent A composed of HPLC grade water (Fisher Scientific) and solvent B composed of isopropanol (Fisher Scientific),

each containing 0.1% TFA. The column was initially equilibrated with 70% solvent A, 30% solvent B. The organic phase was linearly increased to a final condition of 0% solvent A, 100% solvent B. The elution was monitored at a wavelength of 280 nm at which aromatic amino acids within the peptides absorb UV light. Multiple runs of RP-HPLC were required to obtain an adequate amount of pure peptide for subsequent experiments with a typical yield of ~30%. Fractions containing purified GpA or BPV E5 peptides were collected and analysed by mass spectrometry (as described in section 2.5).

Table 2.3 Gradient used for the purification of GpA peptides by RP-HPLC

Time (mins)	% Solvent A	% Solvent B
0	70	30
15	70	30
35	20	80
65	0	100
75	0	100
85	70	30

Table 2.4 Gradient used for the purification of BPV E5 peptides by RP-HPLC

Time (mins)	% Solvent A	% Solvent B
0	70	30
15	70	30
50	0	100
70	0	100
90	70	30

2.5 Mass Spectrometry

2.5.1 Electrospray Ionisation (ESI) Mass spectrometry

Fractions obtained from RP-HPLC were analysed using an electrospray ionisation time of flight (ESI-MS) MicroTOF mass spectrometer (Bruker Daltonics, Coventry, UK) to identify which contained pure peptide and to confirm purity of the fractions collected. HPLC fractions were prepared for analysis by addition of 10 μ L of 10% formic acid to 90 μ L of purified peptide. The addition of formic acid to a final concentration of 1% aids flight of peptides in the mass spectrometer. Samples prepared for ESI mass spectrometry were analysed by direct infusion of the sample into the spectrometer and spectra recorded in positive ion mode, measuring between 500 and 3000 m/z (mass/charge) for an average of 2 min. Individual collected spectra were averaged and deconvoluted using the data analysis software provided by the manufacturer (Bruker Daltonics) to obtain the mass of the singly charged species in the sample. Those HPLC fractions which contained pure (>95%) peptide were pooled and lyophilised before storage in TFE.

2.5.2 Matrix Assisted Laser desorption ionisation (MALDI) Mass spectrometry

Crude peptides and selected purified fractions were also analysed by matrix-assisted laser desorption ionisation (MALDI) mass spectrometry, which is a soft ionisation technique allowing the detection of singly charged species as opposed to the multiply charged species observed by ESI MS. Samples obtained from HPLC purification were prepared by combining 5 μ L of purified peptide with 5 μ L of matrix solution (10 mg/mL alpha-cyano-4-hydroxycinnamic acid (α -CHCA) in 50% acetonitrile (ACN), 50% H₂O and 0.1% TFA). 1 μ L of the sample/matrix solution mix was spotted onto a MALDI plate (Bruker Daltonics) and allowed to air dry for 30 minutes before being loaded into the MALDI mass spectrometer (Bruker MALDI-ToF). MALDI-MS spectra were recorded in positive ion and linear mode between the mass range of 2000 to 5000 m/z, calibrated externally with polyethyleneglycol (PEG) 2000.

2.6 Peptide preparation and characterisation

2.6.1 Lyophilisation

Upon confirmation of peptide purity all fractions containing pure peptide were pooled into a 100 mL round bottom flask to which a small quantity of distilled H₂O was added. The flask was swirled in a Dewar of liquid nitrogen to form a thin layer around the bottom of the flask before being placed on dry ice and attached to a vacuum line. The flask was left under high vacuum until all organic solvent had been removed. The resulting lyophilised peptide powder/film was then dissolved in TFE and aliquoted into small glass vials stored at -20 °C ready for use.

2.6.2 Protein concentration determination

The concentration of purified peptides was estimated using UV absorbance at 280 nm using the Beer-Lambert Law (**Equation 7**) where A is the absorbance at 280 nm, ϵ is the extinction coefficient ($M^{-1} cm^{-1}$), l is the path length of the cuvette (cm) and c the concentration ($mol L^{-1}$) (Grimsley and Pace 2004).

$$A = \epsilon lc \quad (7)$$

For each peptide the extinction coefficient at 280 nm was calculated (Gill and Vonhippel 1989) by inputting the amino acid sequence into the tool ProtParam on the ExPASy website, <http://www.expasy.org> (Wilkins, Gasteiger et al. 1999). For all Glycophorin A peptides the extinction coefficient (ϵ_{280}) was calculated to be $1490 M^{-1} cm^{-1}$ and for all BPV E5 peptides ϵ_{280} was calculated to be $6990 M^{-1} cm^{-1}$.

The UV absorbance of 50 μ L of purified peptide in TFE was measured using a UV/Vis spectrophotometer (Biomate, Thermo Scientific) and the resulting protein concentration was estimated. Whilst TFE is optically transparent at these wavelengths, the weak polarity of the solvent can have an effect on the absorptive energy in UV and may therefore have an effect on ϵ)

2.6.3 SDS-PAGE analysis of purified peptides

Purified peptides were visualised by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) in order to estimate molecular weight and oligomeric state. Peptides were separated using NuPAGE 12% Bis-Tris precast gels in an XCell Sure Lock™ vertical mini-cell electrophoresis system, run in MES SDS running buffer (Invitrogen). 15-20 µL of each sample was typically loaded onto the gel and run alongside 5 µL of SeeBlue® Plus2 Pre-Stained Standard (Invitrogen) molecular weight marker. Electrophoresis was generally carried out for 30 minutes at room temperature at 120 A and 200 V.

Samples were prepared by taking the required amount of stock protein in TFE and aspirating off all solvent before being resuspended in SDS sample loading buffer (125 mM Tris-HCl, 20% glycerol, 4% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol, pH 6.8) before being boiled in a heating block for 10 mins.

2.6.4 Coomassie staining

Following gel electrophoresis, separated protein bands were visualised by Coomassie staining (Fazekas de St Groth, Webster et al. 1963). Coomassie dye interacts with positive protein amine groups through Van der Waals attractions and ionic interactions between sulfonic acid groups (Tal, Silberstein et al. 1985). Gels were soaked in 50 mL of fixing solution (50% methanol, 10% acetic acid) for approximately 30 minutes and then transferred to 50 mL of Coomassie staining solution (0.025% Coomassie G250, 10% acetic acid) with gentle shaking for between 6-12 hours at room temperature. Gels were then de-stained (20% methanol, 7% acetic acid) for 3-6 hours with several changes of de-staining solution to ensure a minimal background so that protein bands could be visualised, after which the de-staining solution was removed and the gel rinsed with and stored in dH₂O.

2.6.5 Silver staining

Following gel electrophoresis, proteins were also visualised by silver staining (Switzer, Merrill et al. 1979). Silver staining is more sensitive than Coomassie staining and can detect 0.5-5 ng protein concentrations, as compared to 5-50 ng for Coomassie (Chevallet, Luche et al. 2006). Gels were placed in a fixing solution (60 mL 50% acetone,

1.5 mL 50% TCA, 25 μ L formaldehyde) for 15 minutes with gentle shaking. Gels were then washed with three quick washes in dH₂O before being soaked in dH₂O for 5 minutes. After soaking, gels were washed a further three times in dH₂O. Gels were then briefly soaked in a sodium thiosulphate solution (10 mg Na₂S₂O₃ in 60 mL dH₂O) for one minute before being washed quickly three times with dH₂O and being soaked in stain solution (0.26% (w/v) silver nitrate, 0.37% formaldehyde in 60 mL dH₂O) for 8 minutes. Following staining gels were washed twice briefly in dH₂O before being placed in a developer solution (1.2g Na₂CO₃, 25 μ L formaldehyde, 25 mg Na₂S₂O₃ in 60 mL dH₂O) for 5-15 seconds whilst bands developed. The development of protein bands was quenched by soaking gels with 1% acetic acid, before being rinsed and stored in dH₂O.

2.7 Peptide reconstitution into lipid vesicles

2.7.1 Detergent solubilisation of peptides

Synthetic GpA and E5 peptides used in this study were solubilised in a number of detergents for optimisation of the detergent removal method for production of solid state NMR samples. Dodecylphosphocholine (DPC), *n*-Dodecyl β -D-maltoside (DDM) and dimyristoyl-*sn*-glycerol-3-phosphocholine (DHPC) was purchased from Avanti Polar lipids (Alabaster AL, USA) and β -D-glucopyranoside (OG) from Fisher Scientific (UK).

In order to correctly solubilise synthetic peptides, detergents were used at concentrations above their critical micelle concentration (CMC), above which detergent micelles are formed. **Table 2.5** gives the molecular weight (MW) and CMC for all detergents used in this study.

Table 2.5 List and properties of detergents used in this study

Detergent	Molecular Weight (MW)	Critical Micelle Concentration (CMC) mM	Aggregation number
β -D-glucopyranoside (OG)	292	19-25	84
Dodecylphosphocholine (DPC)	352	1.5	50-60
<i>n</i> -Dodecyl β -D-maltoside (DDM)	511	0.2	98
dimyristoyl- <i>sn</i> -glycero-3-phosphocholine (DHPC)	454	15	30

2.7.2 Detergent removal

Various reconstitution conditions were tested in order to form optimal proteoliposomes for solid state NMR analysis, lipid vesicles with well folded peptide inserted. The first preparation method tested for formation of proteoliposomes was referred to as the "detergent removal method" (Rigaud and Lévy 2003). In this method a total of 5 mg of purified, lyophilised GpA peptide was dissolved in 5 mL of detergent buffer solution (20 mM sodium phosphate, 20 mM NaCl, 3 mM NaN₃ and 34 mM octyl-glucoside (OG) at pH 7.4) to which 200 μ L of TFE and 4 μ L TFA was added to aid solubilisation of the peptide. TFE was removed by evaporation under nitrogen gas. To this peptide detergent buffer solution, 4.75 mL of 2 mg/mL 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) lipid (Avanti Polar lipids) and 250 μ L cholesterol (Sigma) (5%) in chloroform was added at a 8.6 : 0.75 : 1 lipid:cholesterol:peptide molar ratio. The lipid/peptide/detergent solution was vortexed briefly to form mixed micelles and was allowed to mix with gentle agitation at room temperature for 30 minutes. The solution was then transferred to 4 °C with shaking for 20 minutes before being diluted with 10 mL of phosphate buffer (in order to dilute the detergent concentration and reduce the CMC) and left for another 30 minutes, the solution was then further diluted with 20 mL of phosphate buffer and left shaking for an hour at room temperature. At each dilution step, a 20 μ L aliquot was taken for analysis by electron microscopy (EM).

Detergent was removed using SM2 Bio-Beads (BioRad). Bio-beads are small polystyrene macro porous beads that can be used to remove small detergent molecules (Rigaud,

Levy et al. 1998). Detergent molecules are small enough to enter the pores within the Bio-beads and become trapped due to hydrophobic interactions whilst lipid and protein molecules are too large to enter and as such detergent is selectively removed.

Before mixing with the sample, the Bio-beads were first washed in ethanol followed by three washes with dH₂O. 250 mg of beads were added to the lipid/peptide/detergent solution and left with agitation for two hours at 4°C. This step was repeated, after which a further 500 mg of beads were added and the solution left overnight to ensure full removal of detergent. The following day biobeads were removed by decanting and the resulting proteoliposome solution produced by detergent removal using biobeads was transferred into fresh 1.5 mL microcentrifuge tubes in preparation for centrifugation and transfer into a solid state NMR rotor as outlined in section **2.7.4**.

2.7.3 Co-solubilisation

A more straightforward means to produce proteoliposomes was also tested, and is referred to as the “co-solubilisation method”. In this method, DMPC lipid was dissolved in TFE and combined with cholesterol dissolved in chloroform at a cholesterol to DMPC molar ratio of 0.08:1 (5% w/w ratio). For samples containing the doubly labelled GpA peptide (GpA_{VG}), 5.5 mg of peptide was dissolved in 1 mL of TFE and mixed with 5.5 mL of 2 mg/mL DMPC in TFE (11 mg total) and 275 µL of 2 mg/mL cholesterol in chloroform (0.55 mg total) at a 8.6 : 0.75 : 1 lipid:cholesterol:peptide molar ratio, or a 2 : 0.1 : 1 w/w ratio. For samples containing singly labelled GpA peptides, 2.2 mL of 2.2 mg/mL GpA_V and 1.9 mL of 2.6 mg/mL GpA_G dissolved in TFE, were mixed together at a 1:1 molar ratio (5 mg of each peptide for 10 mg total) before being combined with 10 mL of 2 mg/mL DMPC (20 mg total) and 500 µL of 2 mg/mL cholesterol in chloroform (1 mg total) at the same lipid to cholesterol to protein molar ratio and w/w ratio as the doubly labelled GpA sample.

For BPV E5 samples, using singly labelled ¹³C/¹⁵N labelled Leu 24 and Phe 28 peptides, 754 µL of 3.6 mg/mL BPV E5_L and 1.1 mL of 3.6 mg/mL BPV E5_F dissolved in TFE, were mixed together at a 1:1 molar ratio (4 mg of each peptide for 8 mg total) before being combined with 5 mL of 4 mg/mL DMPC (20 mg total) and 500 µL of 2 mg/mL cholesterol in chloroform (1 mg total) at a 13.5: 1.1 : 1 lipid:cholesterol:protein molar ratio, or a 2.5 : 0.1 : 1 w/w ratio.

For the sample using $^{13}\text{C}/^{15}\text{N}$ labelled Phe 28 and Tyr 31 (BPV E5_F and BPV E5_Y), the same amount of protein was used as above but a much higher cholesterol content was tested. 1.1 mL of 3.6 mg/mL BPV E5_F and 1.5 mL of 2.6 mg/mL BPV E5_Y were dissolved in TFE and mixed together at a 1:1 molar ratio (4 mg of each peptide for 8 mg total) before being combined with 5 mL of 4 mg/mL DMPC (20 mg total) and 1.7 mL of 4 mg/mL cholesterol in chloroform (7 mg total) to yield a 14 : 8.4 : 1 lipid:cholesterol:protein molar ratio and 2.5 : 0.8 : 1 w/w ratio.

Lipid/cholesterol/peptide mixtures were mixed using a vortex mixer and lyophilized using a high vacuum line, and the resulting films were then kept under vacuum overnight in order to remove all traces of organic solvent. The films were resuspended in 4 mL of buffer (20 mM sodium phosphate, 20 mM NaCl, 3 mM NaN₃, pH 7.4) and sonicated briefly before being subjected to five rapid freeze thaw cycles to form proteoliposomes.

2.7.4 Transfer of samples to solid state NMR rotor

Proteoliposomes formed by either by detergent removal or co-solubilisation methods were then centrifuged at 19,000 rpm (45,000×*g*) for 30 minutes and the supernatant removed. To remove residual water, the pellets obtained were transferred to 1.5 mL microcentrifuge tubes (Beckman Coulter, High Wycombe, UK) and centrifuged again for an hour at 70,000 rpm (267,000×*g*) in an Optima™ TLX ultracentrifuge (Beckman Coulter). The resulting gel-like, hydrated proteoliposomes were then transferred to 4 mm MAS rotors (Bruker) ready for solid state NMR measurements.

2.8 Characterisation of proteoliposomes

2.8.1 Circular Dichroism (CD) spectroscopy

Circular dichroism (CD) is a spectroscopic technique that is commonly used to characterise the secondary structure of proteins. CD is used to measure the difference in absorbance of left and right-handed circularly polarised light in order to obtain a CD spectrum. This spectrum can be compared to those obtained for α -helical, β -sheet, and

random coil proteins in order to estimate the secondary structure (Kelly and Price 2000; Bulheller, Rodger et al. 2007). CD experiments were carried out at room temperature (RT) (~25 °C) using a Jasco J-815 spectropolarimeter (Jasco UK, Great Dunmow, Essex, UK) and 1 mm path-length quartz cuvettes (Starna; Optiglass Ltd., Hainault, UK) requiring a sample volume of 200 μ L. Spectra were recorded in the far UV region between 190 and 260 nm, with a data pitch of 0.2 nm, a 1 nm bandwidth, 50 nm min⁻¹ scanning speed and a response time of 2 s.

Samples for CD were typically prepared with a final peptide concentration of 0.25 mg/mL in buffer (20 mM sodium phosphate, 20 mM NaCl, 3 mM NaN₃, pH 7.4), and 8 or 16 individual spectra were collected and averaged. In cases where the high tension [HT] voltage rose above 600 mV at wavelengths below 200 nm, indicating excessive light scattering and potentially unreliable results, the data were truncated to 200 nm. For each sample, a blank containing all sample constituents apart from the peptide was subtracted to obtain final spectrum.

2.8.2 Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy

Attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR) is a technique which is used to obtain an infrared spectrum of absorption, emission, photoconductivity or Raman scattering of membrane proteins in a lipid environment (Vigano, Manciu et al. 2000). A schematic of the ATR-FTIR experiment is given in **Figure 2.1**. ATR-FTIR spectra were recorded on a Jasco FTIR-470 spectrometer equipped with a liquid nitrogen cooled MCT detector purged with dried compressed air. Stacked bilayers were used for ATR spectroscopy by pipetting 400 μ L of proteoliposome suspension onto a trapezoidal zinc selenide (ZnSe) crystal (Specac, Orpington, Kent, UK) and dried under a stream of diffused nitrogen gas. Spectra were recorded between 4000-1000 wavenumbers (cm⁻¹) as an average of 1000 interferograms. A schematic of the ATR-FTIR experiment is given in Figure 2.2. An initial spectrum was recorded followed by deuterium exchange measurements, for which the stacked bilayers were then exposed to D₂O saturated nitrogen gas over of the dried film. A spectrum was recorded 1 hour following exposure to D₂O and a second measurement was taken 22 hours later.

From the spectra recorded, the amide I and II peaks from both the initial and D₂O exposed spectra were compared by integration of the amide peaks using the GRAMS/AI software (Thermo Scientific). The ratio of the amide II to amide I peaks was calculated for each sample to account for sample swelling that can occur following exposure to D₂O, which leads to a reduction in overall absorbance of the signal observed as described previously (Beevers, Damianoglou et al. 2010). Comparison of this ratio for measurements taken before and after D₂O exposure enables the degree of peptide insertion to be calculated. To obtain secondary structure information, spectra were peak fitted using the GRAMS/AI software (Thermo Scientific) as described (Beevers and Kukol 2006).

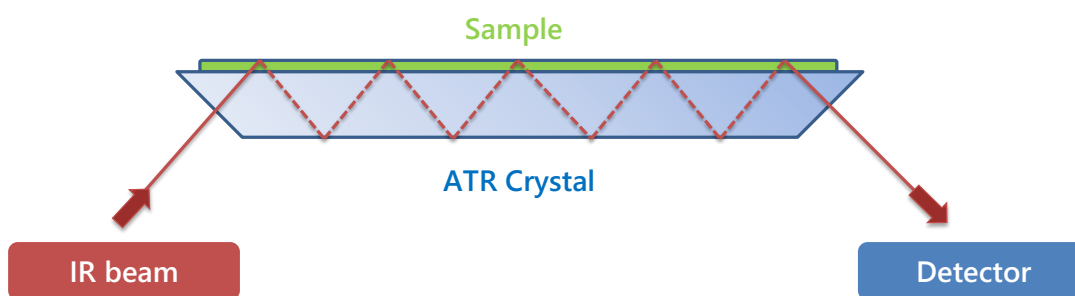


Figure 2.1 Schematic of ATR-FTIR setup

In ATR-FTIR the proteoliposome sample is pipetted onto the top of an ATR crystal and allowed to dry to a film (green strip). This crystal is then placed into the FTIR spectrometer and an I.R beam is fired at the crystal. Internal reflection within the ATR crystal occurs and an evanescent wave that extends beyond the surface of the crystal into the sample is created. The sample absorbs energy which is then passed back into the IR beam, which exits the crystal and is collected by the detector to give an IR interferogram.

2.8.3 Electron Microscopy (EM)

Lipid vesicles were analysed using transmission electron microscopy (TEM) in order to check for vesicle formation or protein aggregation. Proteoliposome solutions prepared using the detergent removal method were diluted to a concentration of 0.04 mg/mL lipid to ensure sufficient physical separation of lipid vesicles observed. Typically 4 μ L of diluted (1 in 50) lipid solution was spotted onto a glow discharged Formvar-Carbon

coated copper 300 mesh grid (Agar Scientific Ltd. Essex, UK) and incubated for 30 seconds. Grids were stained six times with drops (10 μ l) of uranyl acetate solution, and residual stain was removed by blotting off excess staining solution between each step. Grids were then loaded into the transmission electron microscope (JEOL 1200 EXII) for visualisation.

Negative stained vesicles were photographed using the microscopes CCD camera at varying levels of magnification in order to visualise lipid vesicle size, shape and sample uniformity and to see if any reconstitution conditions resulted in protein aggregation.

2.9 Magic Angle Spinning (MAS) Solid state NMR spectroscopy

2.9.1 Solid state NMR experimental procedure

^{13}C cross polarisation (CP) magic angle spinning (MAS) NMR spectra were recorded on a wide bore Bruker Avance III 500 MHz solid state NMR spectrometer (Bruker, Karlsruhe, Germany) operating at 500.1 MHz for ^1H , 125.7 MHz for ^{13}C and 50.68 MHz for ^{15}N . The spectrometer was equipped with a 4 mm triple resonance magic angle spinning (MAS) probe (Bruker) running in double resonance mode for $^1\text{H}/^{13}\text{C}$ experiments and in triple resonance mode for $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ experiments. Samples were cooled with a Bruker BCU Xtreme unit to 258 K (-15°C) to reduce internal lipid motions, and the spinning frequency set to 8.5 kHz \pm 5 Hz maintained by a Bruker BVT 3000 MAS controller unit. External ^{13}C chemical shift referencing was carried out with respect to the carbonyl peak of natural abundance alanine (177.9 ppm) on the DSS scale (Morcombe and Zilm 2003; Harris, Becker et al. 2008). ^{15}N referencing was carried out with respect to external ^{15}N labelled histidine (46.4 ppm). Samples containing DMPC lipid vesicles containing 0 - 0.6 molar ratio cholesterol were prepared in order to study the lipid phase transition at experimental temperatures.

2.9.2 1D $^{31}\text{P}/^1\text{H}$ lipid NMR experiments

Static wide line 1D ^{31}P and MAS ^1H spectra were recorded on a Bruker Avance II+ solid state NMR spectrometer (Bruker, Karlsruhe, Germany) operating at 599.40 MHz for ^1H and 242.64 for ^{31}P and equipped with a 4 mm triple resonance magic angle spinning (MAS) probe running in double resonance mode. For ^1H MAS experiments the spinning speed was set to 5 kHz \pm 5 Hz maintained by a Bruker MAS controller unit. External ^{31}P chemical shift referencing was carried out with respect to the phosphorous peak of adenosine di-hydrogen phosphate (ADP) (0.9 ppm) on the DSS scale and ^1H referencing to H α of natural abundance alanine (4.2 ppm). For 1D ^{31}P experiments a standard Hahn Echo pulse sequence (Rance and Byrd 1983) was used with an echo delay of 50 μs and 80 kHz two pulse-phase modulated (TPPM) proton decoupling during the 40 ms acquisition, and a recycle delay of 5 seconds for 256 co-added transients. Spectra were acquired using a $\pi/2$ (90°) pulse for excitation of ^1H and ^{31}P of 2.5 μs and 4 μs respectively. ^{31}P spectra were acquired with 8k complex data points in F_1 with a spectral window of 412 ppm, and data were Fourier transformed into 16k complex data points. 1D ^1H spectra were recorded for 20k complex points in F_1 using a single proton pulse for 128 co-added transients with a 3 sec recycle delay with a spectral window of 834 ppm, data were Fourier transformed into 65k complex data points and EM line broadening of -1.0 Hz was applied during processing. Spectra were recorded at temperatures ranging from 253-298 K (-20 - 25°C).

2.9.3 1D ^{13}C NMR experiments

1D ^{13}C spectra were recorded with a 10 ms acquisition time, with 1k complex data points in F_1 for 1024 co-added transients, and a spectral window of 397 ppm. A ^1H - ^{13}C cross polarisation (CP) at 73 kHz with a contact time of 1 ms was used, along with 100 kHz proton decoupling during acquisition. Data were Fourier transformed into 65k complex data points and GM line broadening of 20 Hz was applied during processing. Unless otherwise stated, spectra were acquired using a 3 s recycle delay and a $\pi/2$ (90°) pulse for excitation of ^1H and ^{13}C of 2.5 μs and 4 μs respectively.

2.9.4 2D ^{13}C - ^{13}C DARR NMR experiments

Dipolar-assisted rotational recouping (DARR) spectroscopy (Takegoshi, Imaizumi et al. 2000) assists the spin diffusion process by using a combination of physical rotation of the sample and the application of continuous pulses on the proton channel during a mixing time (τ_{mix}) to reintroduce homonuclear dipolar coupling that is normally averaged out by magic angle spinning, thereby allowing the observation of long range ^{13}C - ^{13}C contacts at long mixing times. In the DARR experiments, carbon magnetisation was produced by using a (50-100%) ramped proton CP pulse of 1 ms with RF field strengths of 100 kHz for ^1H and 80 kHz for ^{13}C and transferred using dipolar-assisted rotational resonance (DARR). Proton decoupling was performed during signal acquisition at ~ 100 kHz by using the SPINAL-64 (Fung, Khitrin et al. 2000) (small phase incremental alteration with 64 steps) method employed during signal acquisition periods of 10 ms in the direct dimension and 5 ms in the indirect dimension, with decoupling field strengths of approximately 83 -100 Hz. Experiments were conducted with a range of mixing times (typically 10 - 600 ms) in order to probe short and long distance spin interactions. All 2D ^{13}C - ^{13}C DARR spectra were acquired with 994 complex data points in F_2 . Measurements of the mixture of GpA peptides (GpA_V & GpA_G) were performed with 339 F_1 increments and 308 co-added transients, while measurements of the doubly-labelled GpA (GpA_{VG}) were performed with 120 increments in F_1 and 112 co-added transients. For all 2D DARR experiments the spectral window for F_2 and F_1 was set to 397 and 270 ppm respectively. The data were Fourier transformed to 4k (F_2) \times 2k (F_1) and GM (Lorentz-Gauss) line broadening of -1.0 Hz in F_2 and QSINE line broadening (SSB=0.3) applied. Spectra obtained were processed and analysed using Bruker Topspin 2.1 software. Cross peak integrals (obtained from cross peak volumes in 2D spectra) were plotted against τ_{mix} to obtain build-up curves of cross peak intensity for intra- and inter-helical correlations.

2.9.5 2D ^{13}C - ^{15}N NMR experiments

In order to probe through-space carbon and nitrogen coupling between labelled amino acids at the GpA dimer interface, Transfer Echo DOuble Resonance (TEDOR) experiments were used (Fyfe, Mueller et al. 1992; Jaroniec, Filip et al. 2002). Experiments were conducted with a range of mixing times (typically 5-10 ms) in order to probe both

short and long-range spin interactions. All 2D ^{13}C - ^{15}N TEDOR spectra were acquired with 3k complex data points in F_2 and 20 points in F_1 . The spectral windows for F_2 and F_1 were set to 397 and 49 ppm, respectively. Measurements of the mixture of GpA peptides (GpA_V & GpA_G) were acquired with 3840 co-added transients. The data were Fourier transformed to 4k (F_2) \times 128 (F_1) and GM (Lorentz-Gauss) line broadening of -1.0 Hz in F_2 and QSINE (SSB = 0.3) line broadening of 0.3 Hz applied.

2.10 Solution NMR experiments

2.10.1 Bicelle preparation

Isotropic bicelles (Sanders and Schwonek 1992; Vold, Prosser et al. 1997) prepared from long chain phospholipids and short chain detergents were prepared for solution NMR studies as follows; 20 mg of 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) was dissolved in 2 mL of TFE and 10 mg of DMPC in 1 mL of TFE to form 10 mg/mL stock solutions for producing bicelles. For $q = 0.25$ bicelles, where $q = [\text{DMPC}]/[\text{DHPC} - 15\text{mM}]$ (Vold, Prosser et al. 1997), 681 μL of DHPC in TFE was mixed with 380 μL of DMPC to prepare a lipid bicelle suspension. Bicelles with a $q = 0.33$ were prepared by combining 681 μL of DHPC and 500 μL of DMPC. Once the two components were mixed, TFE was evaporated off using a stream of nitrogen and samples were placed in a desiccator under high vacuum overnight in order to remove all traces of organic solvent. The resulting lipid film was then rehydrated in 1 mL of HEPES/TRIS buffer.

For BPV E5_{v2} samples, 321 μL of purified peptide (4 mg/ mL) in TFE was pipetted into a glass vial. TFE was removed by a stream of nitrogen gas and the sample was placed in a desiccator under high vacuum overnight to remove all traces of organic solvent. To the resulting peptide film, 180 μL of $q = 0.25$ or $q = 0.33$ bicelles was added followed by brief sonication to resuspend the film and vortexing before being subjected to five freeze/thaw cycles (liquid nitrogen/40 °C water bath) to form bicelles with a final peptide concentration of 1 mM, a lipid to protein molar ratio of 30:1 (40:1 w/w) and a total amphiphile concentration of 2%. Samples were then transferred to 3 mm bore NMR tubes (Bruker) for analysis by NMR.

2.10.2 Solution NMR experimental procedure

Solution NMR spectra were recorded on a Bruker Avance II 700 MHz solution state NMR spectrometer (Bruker, Karlsruhe, Germany) operating at 700.1 MHz for ^1H , 70.94 MHz for ^{15}N and 176.04 MHz for ^{13}C . The spectrometer was equipped with a cryoprobe (Bruker) running in triple resonance mode for $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ experiments. Experiments were typically run at 313 K (40 °C) using samples in 3 mM bore NMR tubes (Bruker). ^1H chemical shift referencing was carried out with respect to water (4.7 ppm) on the DSS scale (Wishart, Bigam et al. 1995). Samples contained 10% D_2O which was used to lock onto the spectrometer field during acquisition.

2.10.3 1D ^1H NMR spectroscopy

1D ^1H spectra were recorded with a 400 ms acquisition time, 1.25 s recycle delay, 16 co-added transients, and a typically a $7.5\ \mu\text{s}\ \pi/2$ pulse for excitation of ^1H with a power level of 5.9 dB. Spectra were acquired with 32k complex data points in F_1 and a spectral window of 20 ppm. Data were Fourier transformed into 65k complex data points and GM line broadening of -0.2 Hz was applied during processing

2.10.4 2D heteronuclear NMR spectroscopy

2D ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectra (Schleucher, Schwendinger et al. 1994) were recorded with 2k data complex data points in F_2 , 64 increments in F_1 , and 16 co-added transients per plane. A spectral width of 15 ppm in F_2 and 40 ppm in F_1 was used. Spectra were recorded at a range of temperatures (25-50 °C), but typically 40 °C was used. Spectra were acquired using a $\pi/2$ (90°) pulse for excitation of ^{15}N of $2.5\ \mu\text{s}$ and $4\ \mu\text{s}$ with a power level of -1.30 dB, with 16.80 dB ^{13}C decoupling. The raw data were Fourier transformed ($2\text{k}\ (F_2) \times 256\ (F_1)$ complex data points) and QSINE line broadening ($\text{SSB} = 0.3$) in both F_2 and F_1 applied before peak picking and subsequent analysis in Topspin 2.1 software.

2.10.5 3D heteronuclear NMR spectroscopy

3D ^{15}N edited ^1H - ^{15}N -heteronuclear TOCSY-HSQC (Hwang and Shaka 1995) (Total Correlation Spectroscopy) experiments with spin lock times of 40, 60 and 70 ms were recorded to aid assignment of the backbone amide chemical shifts in BPV E5_{v2}. ^{15}N edited ^1H - ^{15}N NOESY-HSQC (Nuclear Overhauser Effect Spectroscopy) experiments with mixing times of 20-300 ms were also recorded to investigate intra- and intermolecular interactions between peptide monomers. Spectra were acquired as pseudo-3D spectra with 2k data complex data points in F_3 , 256 increments in F_1 , and only one increment in F_2 . The spectral widths for F_3 , F_2 and F_1 were set to 15, 40 and 15 ppm, respectively. A total of 256 co-added transients were acquired for HSQC-TOCSY experiments and 96 co-added transients for HSQC-NOESY experiments. HSQC-TOCSY raw data was Fourier transformed into 2k (F_3) \times 1 (F_2) \times 512 (F_1) complex data points and QSINE line broadening of 551 Hz (F_3) and 0.3 Hz (F_2) applied. For HSQC-NOESY experiments, raw data was Fourier transformed into 2k (F_3) \times 1 (F_2) \times 256 (F_1) complex data points and QSINE line broadening (SSB =1) (F_3) and 0.3 Hz in F_2 and F_1 applied. Experiments were typically acquired at 40 °C and all 3D experimental data were processed in Bruker Topspin 2.1.

2.11 References

- Beevers, A. J., A. Damianoglou, J. Oates, A. Rodger and A. M. Dixon (2010). "Sequence-Dependent Oligomerization of the Neu Transmembrane Domain Suggests Inhibition of "Conformational Switching" by an Oncogenic Mutant." *Biochemistry* **49**(13): 2811-2820.
- Beevers, A. J. and A. Kukol (2006). "Secondary structure, orientation, and oligomerization of phospholemman, a cardiac transmembrane protein." *Protein Sci* **15**(5): 1127-1132.
- Bulheller, B. M., A. Rodger and J. D. Hirst (2007). "Circular and linear dichroism of proteins." *Phys Chem Chem Phys* **9**(17): 2020-2035.
- Carpino, L. A. and G. Y. Han (1972). "9-Fluorenylmethoxycarbonyl Amino-Protecting Group." *J Org Chem* **37**(22): 3404-&.
- Chevallet, M., S. Luche and T. Rabilloud (2006). "Silver staining of proteins in polyacrylamide gels." *Nat Protoc* **1**(4): 1852-1858.
- Fazekas de St Groth, S., R. G. Webster and A. Datyner (1963). "Two new staining procedures for quantitative estimation of proteins on electrophoretic strips." *Biochim Biophys Acta* **71**: 377-391.
- Fisher, L. E. and D. M. Engelman (2001). "High-yield synthesis and purification of an alpha-helical transmembrane domain." *Anal Biochem* **293**(1): 102-108.

- Fung, B. M., A. K. Khitrin and K. Ermolaev (2000). "An improved broadband decoupling sequence for liquid crystals and solids." *J Magn Reson* **142**(1): 97-101.
- Fyfe, C. A., K. T. Mueller, H. Grondey and K. C. Wongmoon (1992). "Dipolar Dephasing between Quadrupolar and Spin-1/2 Nuclei - Redor and Tedor Nmr Experiments on Vpi-5." *Chem Phys Lett* **199**(1-2): 198-204.
- Gill, S. C. and P. H. Vonhippel (1989). "Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data." *Anal Biochem* **182**(2): 319-326.
- Goldstein, D., M. Finbow, T. Andresson, P. McLean, K. Smith, V. Bubb and R. Schlegel (1991). "Bovine papillomavirus E5 oncoprotein binds to the 16K component of vacuolar H(+)-ATPases." *Nature* **352**(6333): 347-349.
- Grimsley, G. R. and C. N. Pace (2004). "Spectrophotometric determination of protein concentration." *Curr Protoc Protein Sci* **Chapter 3**: Unit 3 1.
- Harris, R. K., E. D. Becker, S. M. Cabral De Menezes, P. Granger, R. E. Hoffman and K. W. Zilm (2008). "Further conventions for NMR shielding and chemical shifts IUPAC recommendations 2008." *Solid State Nucl Magn Reson* **33**(3): 41-56.
- Hwang, T. L. and A. J. Shaka (1995). "Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients." *Journal of Magnetic Resonance, Series A* **112**(2): 275-279.
- Jaroniec, C. P., C. Filip and R. G. Griffin (2002). "3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly C-13, N-15-labeled solids." *J Am Chem Soc* **124**(36): 10728-10742.
- Jokinen, M., L. C. Andersson and C. G. Gahmberg (1985). "Biosynthesis of the major human red cell sialoglycoprotein, glycophorin A. O-Glycosylation." *J Biol Chem* **260**(20): 11314-11321.
- Kelly, S. M. and N. C. Price (2000). "The use of circular dichroism in the investigation of protein structure and function." *Curr Protein Pept Sci* **1**(4): 349-384.
- Lew, S. and E. London (1997). "Simple procedure for reversed-phase high-performance liquid chromatographic purification of long hydrophobic peptides that form transmembrane helices." *Anal Biochem* **251**(1): 113-116.
- Mattoon, D., K. Gupta, J. Doyon, P. Loll and D. DiMaio (2001). "Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein." *Oncogene* **20**(29): 3824-3834.
- Morcombe, C. R. and K. W. Zilm (2003). "Chemical shift referencing in MAS solid state NMR." *J Magn Reson* **162**(2): 479-486.
- Morozova, D. and M. Weiss (2010). "On the role of acylation of transmembrane proteins." *Biophys J* **98**(5): 800-804.
- Rance, M. and R. A. Byrd (1983). "Obtaining High-Fidelity Spin-1/2 Powder Spectra in Anisotropic Media - Phase-Cycled Hahn Echo Spectroscopy." *J Magn Reson* **52**(2): 221-240.
- Rigaud, J. and D. Lévy (2003). "Reconstitution of membrane proteins into liposomes." *Methods Enzymol* **372**: 65-86.
- Rigaud, J. L., D. Levy, G. Mosser and O. Lambert (1998). "Detergent removal by non-polar polystyrene beads - Applications to membrane protein reconstitution and two-dimensional crystallization." *European Biophysics Journal with Biophysics Letters* **27**(4): 305-319.
- Sanders, C. R., 2nd and J. P. Schwonek (1992). "Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR." *Biochemistry* **31**(37): 8898-8905.

- Schleucher, J., M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. Sorensen and C. Griesinger (1994). "A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients." J Biomol NMR **4**(2): 301-306.
- Switzer, R. C., 3rd, C. R. Merrill and S. Shifrin (1979). "A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels." Anal Biochem **98**(1): 231-237.
- Takegoshi, K., T. Imaizumi and T. Terao (2000). "One- and two-dimensional ^{13}C - ^1H / ^{15}N - ^1H dipolar correlation experiments under fast magic-angle spinning for determining the peptide dihedral angle ϕ ." Solid State Nucl Magn Reson **16**(4): 271-278.
- Tal, M., A. Silberstein and E. Nusser (1985). "Why does Coomassie Brilliant Blue R interact differently with different proteins? A partial answer." J Biol Chem **260**(18): 9976-9980.
- Vigano, C., L. Manciu, F. Buyse, E. Goormaghtigh and J. M. Ruyschaert (2000). "Attenuated total reflection IR spectroscopy as a tool to investigate the structure, orientation and tertiary structure changes in peptides and membrane proteins." Biopolymers **55**(5): 373-380.
- Vold, R. R., R. S. Prosser and A. J. Deese (1997). "Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides." J Biomol NMR **9**(3): 329-335.
- Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel and D. F. Hochstrasser (1999). "Protein identification and analysis tools in the ExPASy server." Methods Mol Biol **112**: 531-552.
- Wishart, D. S., C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley and B. D. Sykes (1995). " ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR." J Biomol NMR **6**(2): 135-140.

3

PEPTIDE DESIGN, PREPARATION AND CHARACTERISATION

3.1 Introduction

Membrane proteins can be extremely difficult to express in bacterial host systems and equally challenging to subsequently purify due to their inherent hydrophobic nature. As such, a commonly used strategy for studying small transmembrane proteins is the use of peptides, prepared synthetically using solid phase 9-fluorenylmethyl carbamate (Fmoc) chemistry, that have been derived from the transmembrane protein of interest but are much easier to produce for biological studies (Oates, Hicks et al. 2008; Beevers and Dixon 2010; King, Oates et al. 2011) In this work, the structures of two membrane proteins have been investigated in a lipid environment using the above strategy, namely the major sialoglycoprotein found in the membrane of red blood cells (Glycophorin A, or GpA) and the product of the smallest known viral oncogene (the E5 protein from bovine papillomavirus (Surti, Klein et al. 1998). Both proteins are known to form strong α -helical homodimers via protein-protein interactions in the transmembrane (TM) domain (Popot and Engelman 1990). GpA was selected to act as well-characterized “standard” that could be used to validate the sample preparation and experimental methods used in this work since it has been heavily studied in the past (Smith, Jonas et al. 1994; MacKenzie, Prestegard et al. 1997; Smith, Song et al. 2001) and a good deal of structural information exists for this protein. The lessons learned from analysis of GpA were then been applied to study of the E5 protein, for which no three-dimensional structure exists.

3.2 Peptide design and purification for ssNMR analyses

The first step in designing the peptides used in this study was the incorporation of any existing information on the structures of the transmembrane homodimers into a molecular model for the system. Any information about the secondary structure, residues that lie at the homodimer interface and residues that may interact with the lipid was incorporated into models of the GpA and E5 homodimers. In the case of GpA, a solution NMR structure of the TM homodimer is available (PDB reference 1AFO, (MacKenzie, Prestegard et al. 1997)) and was used in this work, as shown in **Figure 3.1**.

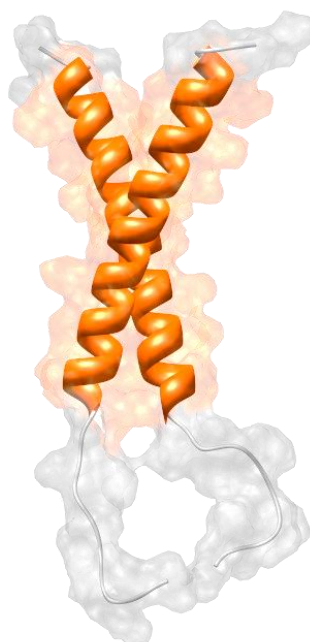


Figure 3.1 3D structure of the alpha-helical homodimer transmembrane Glycophorin A (GpA)

3D structure of GpA produced from pdb file (1AFO) (MacKenzie, Prestegard et al. 1997) deposited at the Protein Databank, structure obtained by solution NMR in DPC micelles, the transmembrane region is highlighted in orange.

Once a working model for the GpA TM homodimer was obtained, a synthetic peptide corresponding to the TM domain of human Glycophorin A, containing the sequence shown in **Figure 3.1**, was synthesised at the KECK Facility (Yale University,

USA). Using synthetic peptides allows the incorporation of isotopically labelled amino acids in key positions (e.g. within the dimer interface) in the peptide. For GpA, Val 80 and Gly 83 have previously been reported to pack tightly at the GpA homodimer interface (Smith and Bormann 1995) which is also confirmed by the solution NMR structure, as shown in **Figure 3.2** where this is highlighted. These residues have also been shown to be important in stabilisation of the dimer (MacKenzie, Prestegard et al. 1997; Smith, Song et al. 2001), as mutation of Gly 83 to Ile results in a substantial decrease in GpA dimer formation (Lemmon, Flanagan et al. 1992; Lemmon, Treutlein et al. 1994).

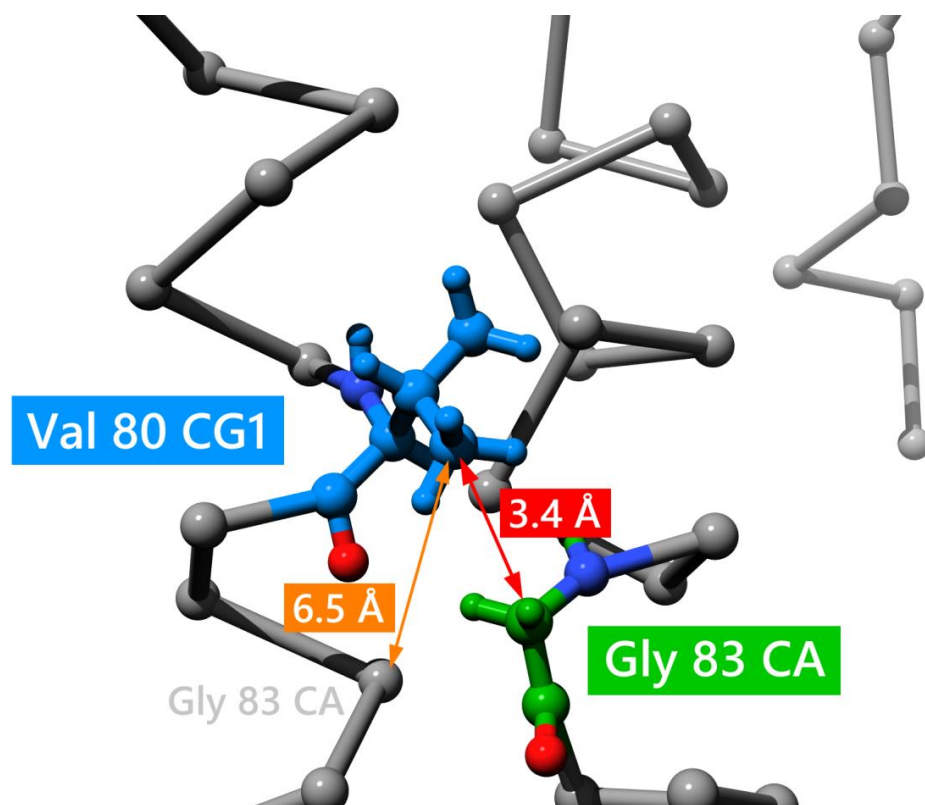


Figure 3.2 Ball and stick representation of the GpA homodimer interface

3D model produced from solution NMR structure (1AFO, Mackenzie et al., 1997). Valine 80 and Glycine 83 are both amino acids at the homodimer interface which have been shown to interact and pack tightly together at the dimer interface, and have been highlighted above.

Therefore uniformly labelled $^{13}\text{C}/^{15}\text{N}$ amino acids were inserted at positions Val 80 and Gly 83 (GpA_{VG}) (**Figure 3.3**) in synthetic peptides prepared for this study. A number of GpA peptides were synthesised for this project as detailed in (**Table 2.1, Chapter 2**), containing U- $^{15}\text{N}/^{13}\text{C}$ amino acids at one (GpA_V & GpA_G) or both (GpA_{VG}) of the above named positions in the sequence. Peptides containing only one labelled amino acid per chain were mixed together and used to confirm that any interactions between labelled amino acids were fully attributable to inter-helical interactions as opposed to intra-helical interactions.

GpA_{VG}: KKITLIIFGV₈₀MAG₈₃VIGTILLISYGIKK

GpA_V: KKITLIIFGV₈₀MAGVIGTILLISYGIKK

GpA_G: KKITLIIFGVMAG₈₃VIGTILLISYGIKK

Figure 3.3. Isotopically labelled GpA TM domain peptide sequences

Uniformly labelled $^{13}\text{C}/^{15}\text{N}$ labelled amino acids in bold type

In the case of the E5_{TM} dimer, no three-dimensional structural data was available, since the structure has not been solved to date, so a model of the E5 homodimer was created (**Figure 3.4**) by using the CHI molecular dynamics software (Adams, Arkin et al. 1995). The CHI software uses *in vacuo* computational modelling combined with molecular dynamics to search for energetically favourable inter-helical interactions by systematic rotation of the helices at varying degrees of crossing angle between the central plane of axis for each helix. During the search of interactions for TM domain dimers, the two helices were simultaneously rotated about their central axis in 30° increments from 0 to 360°. Molecular dynamic simulations are then used in order to produce a 3D model of the energy minimised homodimer of the E5_{TM} protein based upon the amino acid sequence.

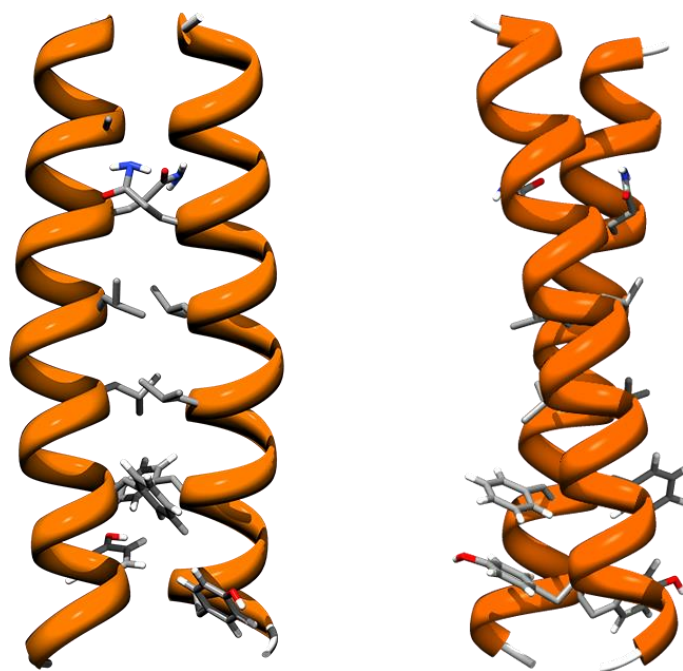


Figure 3.4 Molecular model of BPV E5 transmembrane protein generated from CHI

Model of BPV E5_{TM} putative homodimer interface. Amino acids at the dimer interface are indicated in stick form. Generated using the molecular modelling software CHI to produce an energy minimised structure of the homodimer using the peptide sequence supplied. Face on model to the left, rotation by 90° represented on the right.

As with GpA, when deciding upon which amino acids to label at the proposed dimer interface, we chose to label only one amino acid per chain of the dimer. This way any interactions observed between the labelled amino acids could be fully attributed to an inter-helical interaction with no possibility of it arising through an intra-helical interaction. The CHI generated E5.pdb file was analysed in the molecular modelling software MOLMOL (Koradi, Billeter et al. 1996) in order to obtain a comprehensive list of intra and inter molecular distances between the individual carbon atoms in the E5 homodimer (**Table 3.1**) to identify which amino acid pairings would be predicted to have the smallest inter-helical distances that would be observable by ssNMR. This list of distances, alongside previous knowledge of which amino acids are believed to lie at the dimer interface based upon mutagenesis data (DiMaio, Guralski et al. 1986) (Klein, Polack et al. 1998), were used to pick which amino acids to label. The only amino acids in E5

that (i) were at the interface (ii) had the shortest inter-helical distance pairings ($< 5.5 \text{ \AA}$) (iii) were economically viable to obtain with isotopic labels were leucine 24, phenylalanine 28 and tyrosine 31. Therefore three different peptides with $^{13}\text{C}/^{15}\text{N}$ uniformly isotopically labelled amino acids at these positions were synthesised. Due to the large number of leucines at dimer interface, there was a limited choice of amino acid sites that were available to label. Glutamine 17 which has been shown to be important in forming dimers (Klein, Polack et al. 1998) would be an ideal choice but due to the cost of ordering $^{13}\text{C}/^{15}\text{N}$ isotopically labelled glutamine it was not economically viable.

Table 3.1 Table of the shortest inter-helical distance between amino acids at the E5 dimer interface

Intra-helical distances between labelled pairs are also indicated. Distances were measured between both chains of the BPV E5 homodimer (chain A and Chain B). Distances were measured from chain A-B and also from chain B-A with inter-helical distances coloured in green to red, for shorter to longer range interactions. Distances generated from CHI molecular model of BPV E5 using MOLMOL modelling software.

Residues	Inter (\AA)			Intra (\AA)		
	A - B	B - A	Average	A-B	B-A	Average
24 Leu CD1 - Phe 28 CG	4.00	5.82	4.91	7.00	9.20	8.10
24 Leu CD1 - Phe 28 CD2	3.66	5.50	4.58	7.80	9.90	8.85
28 Phe CD1 - Tyr 31 CD1	4.93	3.98	4.46	7.70	8.80	8.25
28 Phe CD1 - Tyr 31 CE1	4.89	3.82	4.36	8.50	8.90	8.70
28 Phe CE1 - Tyr 31 CD1	4.49	3.71	4.10	8.80	9.10	8.95
28 Phe CE1 - Tyr 31 CE1	4.08	3.19	3.64	9.60	10.10	9.85
28 Phe CE1 - Tyr 31 CZ	5.19	3.51	4.35	10.70	11.30	11.00
28 Phe CZ - Tyr 31 CE1	4.93	3.68	4.31	10.90	11.20	11.05
28 Phe CZ - Tyr 31 CZ	5.87	3.75	4.81	11.90	12.44	12.17

Using these three peptides allowed for the possibility to measure any inter-helical interactions from Phe 28 to both Leu 24 and Tyr 31(see **Figure 3.5**). The sequences of each peptide with respective isotopic labels are shown below in **Figure 3.6**.

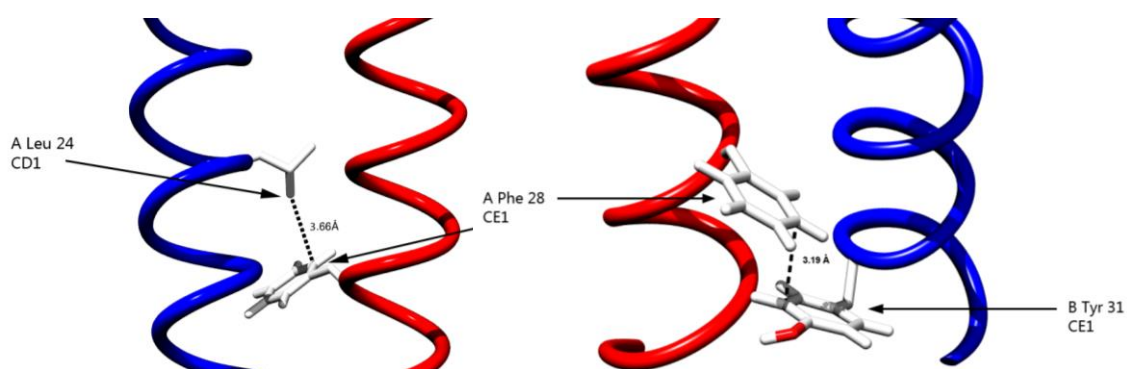


Figure 3.5 Molecular model of selected E5 dimer interfacial regions generated using CHI

Using the amino acid sequence a 3D model of the E5 homodimer was generated from which using the software MOLMOL amino acids with the shortest inter-helical distances with long intra helical distances were selected. Three sites were identified which were suitable for labelling and the following pairings were chosen; leucine 24 and phenylalanine 28 and phenylalanine 28 and tyrosine 31. Carbon atoms from the selected amino acids with the shortest distances are indicated.

$E5_L$: KKKFLGLVAAMQLLLLLFL²⁴LLFFLVYWDHK

$E5_F$: KKKFLGLVAAMQLLLLLFL²⁸LVYWDHK

$E5_Y$: KKKFLGLVAAMQLLLLLFL³¹LVYWDHK

Figure 3.6 Isotopically labelled E5 TM domain peptide sequences

The synthetic peptides chosen for this study were supplied as crude products and required further purification to remove unwanted contaminants such as truncated peptides and protecting groups left over from the synthesis process, and this was achieved using reversed-phase HPLC (RP-HPLC) (for details on the column and conditions used, see **Section 2.4**). Owing to their hydrophobic nature, these synthetic peptides can be difficult to purify as they can bind strongly to the stationary phase of a RP-HPLC column and may be very difficult to elute off. This leads to broad unresolved peaks that lack clear baseline separation. To minimise this effect, the conditions for peptide purification were optimised by changing the gradient of organic solvent that was used as the mobile phase and the flow rate at which the mobile phase was passed over the column when eluting purified peptides

Representative RP-HPLC chromatograms for both GpA and BPV E5 peptides are shown in **Figure 3.7** and **3.8**, respectively. For each peptide, the corresponding chromatograms were almost identical from run to run, irrespective of labelling, demonstrating the reproducibility. As shown in **Figure 3.7**, a number of peaks were eluted during the linear isopropanol (IpOH)/H₂O gradient phase and a single peak relating to purified GpA eluted during the isocratic phase of the run (mobile phase composed of 100% ACN or isopropanol) at 62 minutes. **Figure 3.8** shows the chromatogram for the purification of E5 which eluted during the linear gradient of solvent B at 47 minutes.

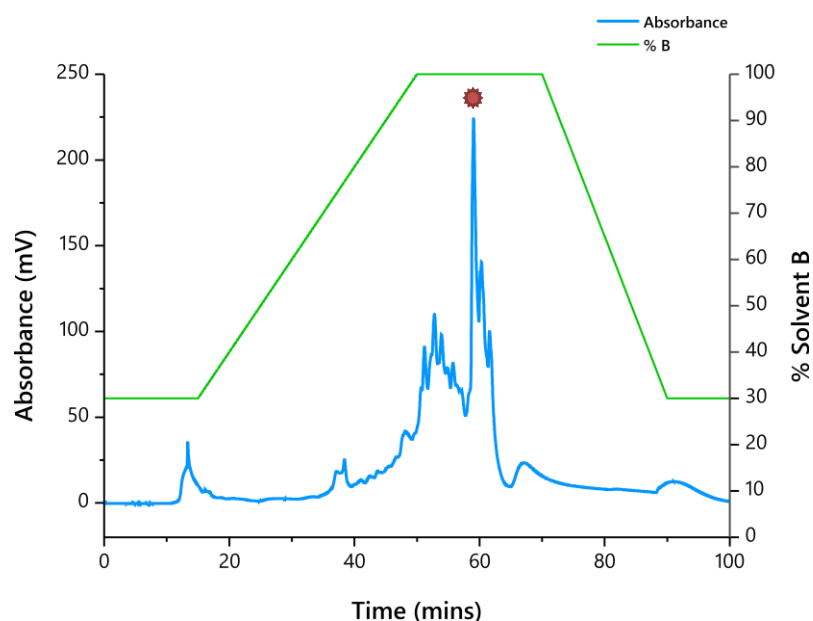


Figure 3.7 Representative RP-HPLC chromatogram of crude GpA purification

GpA peptide was purified by RP-HPLC using a Phenomenex Jupiter C4 column. Separation of crude peptide was monitored using a UV/Vis detector set to 280 nm (blue line). Crude peptide was purified by increasing the percentage of solvent B (IpOH) over time (green line). The peak at 62 minutes (indicated by *) corresponds to pure GpA peptide.

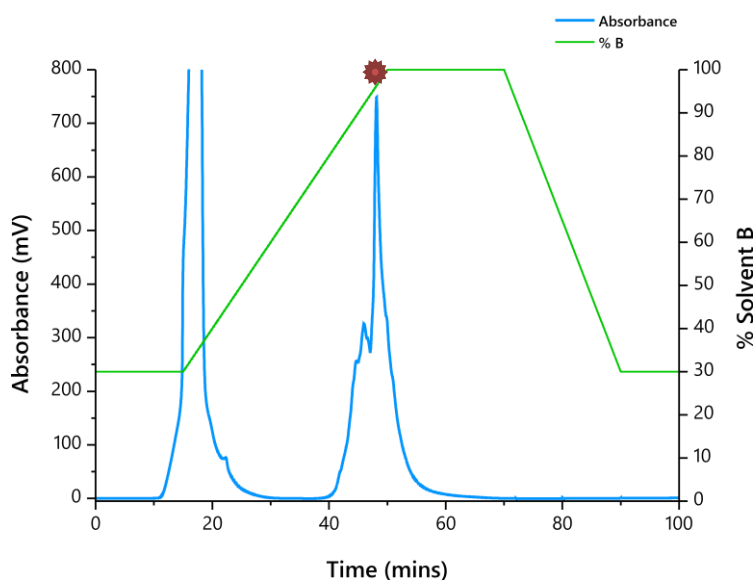


Figure 3.8 Representative RP-HPLC chromatogram of crude BPV E5 purification

BPV E5 peptide was purified by RP-HPLC using a Phenomenex Jupiter C4 column. Separation of crude peptide was monitored using a UV/Vis detector set to 300 nm (blue line). Crude peptide was purified by increasing the percentage of solvent B (IpOH) over time (green line). The peak at 47 minutes (indicated by *) corresponds to pure BPV E5 peptide.

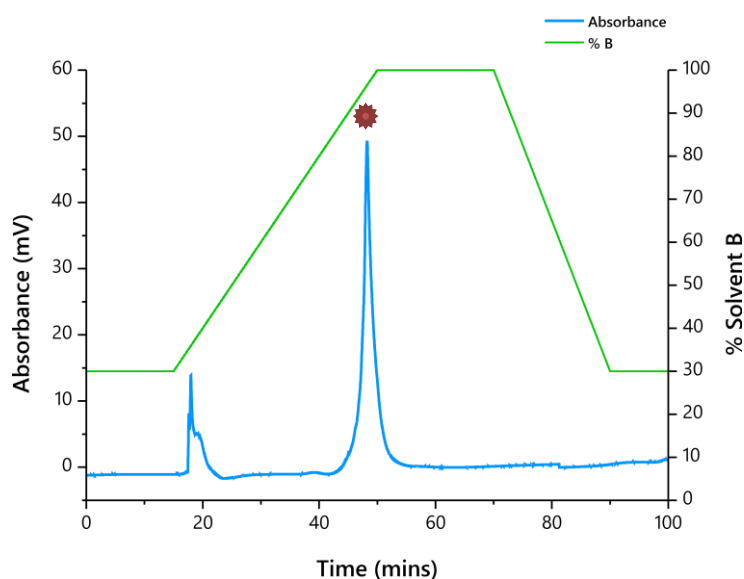


Figure 3.9 RP-HPLC chromatogram of purified BPV E5 peptide

BPV E5 peptide purified by RP-HPLC was re-injected onto a Phenomenex Jupiter C4 column in order to confirm purity. Purified peptide was run using the same conditions as crude peptide and the run monitored using a UV/Vis detector set to 300 nm (blue line). The peak at 47 minutes (indicated by *) corresponds to pure BPV E5 peptide with no contaminant shoulder peaks.

Fractions obtained from RP-HPLC believed to contain pure peptide were analysed using an electrospray ionisation (ESI) MicroTOF mass spectrometer (Bruker, UK) to confirm identity and purity. The purest fractions for a given peptide were pooled, and a final HPLC chromatogram (**Figure 3.9**) and mass spectrum was acquired before lyophilisation. In the resulting mass spectra, multiple charge states of each peptide were observed and summarised in **Table 3.2** for all peptides. These values correlate to +4, +3, +2 and +1 charge states, respectively, of the expected mass of purified GpA and BPV E5 peptides.

Table 3.2 Summary of observed charge states for peptides used in this study analysed by ESI-MS

List of all peptides purified and tested by ESI-MS to confirm peptide identity and purity. Table indicates the observed m/z charge states observed by ESI-MS for each peptide used in this study along with the predicted mass of each synthetic peptide.

Synthetic peptide	+4	+3	+2	Observed mass (+1)	Predicted mass
GpA doubly labelled	735.96	980.94	1470.90	2940.10	2941.76
GpA ¹³ C/ ¹⁵ N Val 80	735.20	979.92	1469.35	2936.78	2938.77
GpA ¹³ C/ ¹⁵ N Gly 83	734.46	978.93	1467.85	2933.80	2934.47
GpA unlabelled	733.91	978.07	---	2931.20	2933.45
BPV E5 ¹³ C/ ¹⁵ N Leu 24	918.53	1224.36	1836.03	3669.09	3671.65
BPV E5 ¹³ C/ ¹⁵ N Phe 28	919.30	1225.39	1837.57	3672.34	3674.65
BPV E5 ¹³ C/ ¹⁵ N Tyr 31	919.10	1225.47	1838.19	3672.39	3674.65
BPV E5 unlabelled	916.75	1222.00	---	3662.00	3664.60

The mass spectra obtained suggest a difference in mass between observed and expected by an average of 2 Da. This difference can be accounted for by calibration of the mass spectrometer. The raw mass spectra also suggest some degree of impurity is present in the E5 peptide fractions, most likely consisting of deletion products arising from the solid phase synthesis of the peptides. A sodium (Na⁺) adduct is seen in the mass spectrum of GpA peptides, whilst the E5 peptides show no sodium adducts but appear to have minor peaks arising from deletion products arising from the synthesis process. Deconvoluted mass spectra for each of the peptides used in this study are shown in **Figure 3.10** for GpA peptides and **Figure 3.11** for BPV E5.

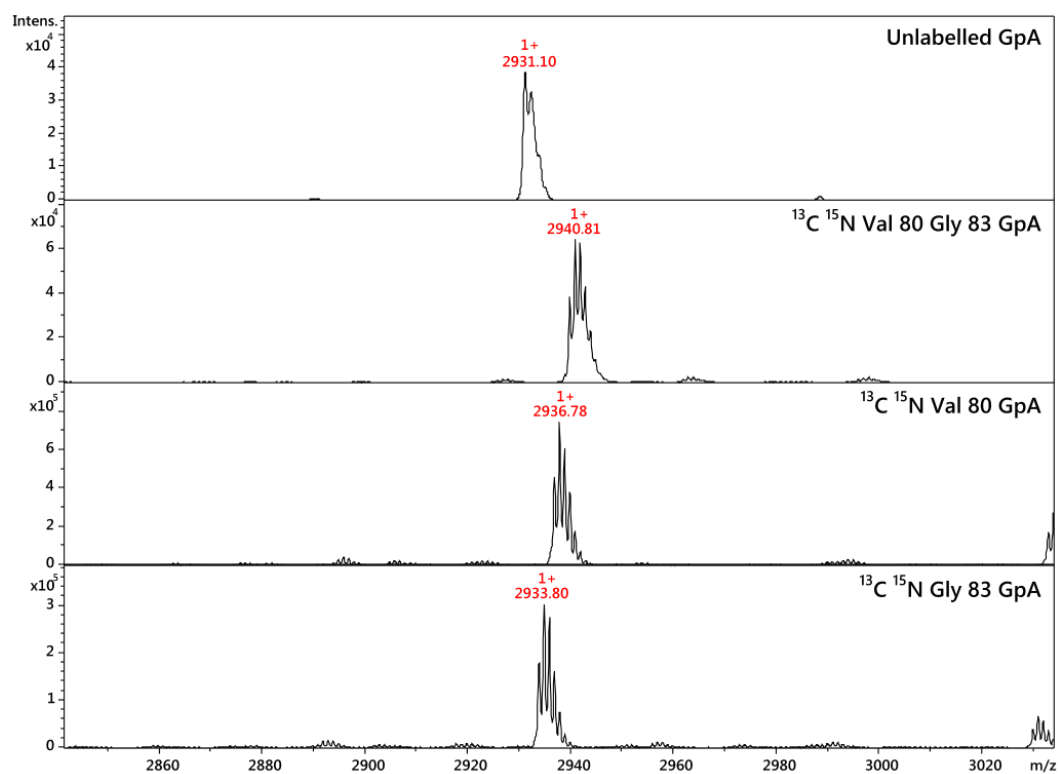


Figure 3.10 GpA Deconvoluted ESI-MicroTOF mass spectra

Isotopic distribution of deconvoluted mass spectra obtained from Micro-TOF ESI-MS of fractions of RP-HPLC purified unlabelled GpA and U- $^{13}\text{C}/^{15}\text{N}$ doubly labelled GpA_{VG} and singly labelled GpA_V and GpA_G peptides.

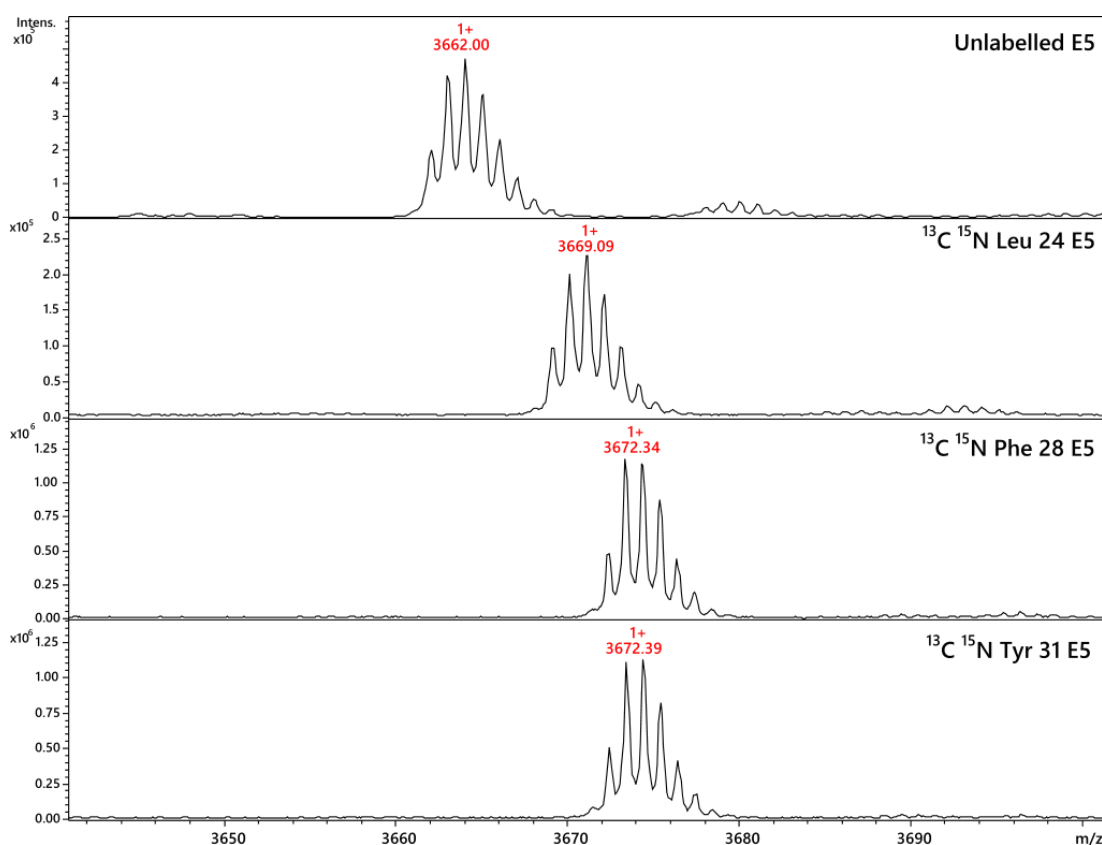


Figure 3.11 BPV E5 Deconvoluted ESI-MicroTOF mass spectra

Isotopic distribution of deconvoluted mass spectra obtained from Micro-TOF ESI-MS of fractions of RP-HPLC purified unlabelled BPV E5 and U-¹³C/¹⁵N E5_L, E5_F and E5_Y peptides

Following lyophilisation, purified peptide was solubilised in TFE and the concentration was estimated using UV/Vis spectrometry and the Beer-Lambert law (**Equation 1** in Methods, **Section 2.6**). The extinction coefficients were calculated using the amino acid sequence for each corresponding peptide and the ProtParam tool at the ExPASy website (<http://expasy.org/>) yielding values of 1490 mol⁻¹ cm⁻¹ for GpA peptides and 6990 mol⁻¹ cm⁻¹ for BPV E5 peptides.

3.3 Reconstitution of GpA peptide into lipid vesicles

In order to obtain structural information for the GpA and BPV E5 TM domain peptides in a lipid environment using solid-state NMR, the preparation of proteoliposomes was required. The choice of lipid to use in these studies was based on several ssNMR reports for small transmembrane proteins in the literature, with the most commonly used lipid being DMPC in a number of cases when studying small transmembrane peptides (Ketchum, Hu et al. 1993; Smith, Song et al. 2001; Wang, Kim et al. 2001; Cady, Mishanina et al. 2009). However, a single component bilayer is not representative of a native bilayer, as biological membranes are typically composed of several different lipid types. The structure of the erythrocyte membrane in which GpA is located, is composed of up to 52% membrane proteins, 40% lipid bilayer and 8% carbohydrates (Smith 1987). The outer monolayer lipid composition of the bilayer consisting of uncharged Phosphatidyl choline (PC) and Sphingomyelin (SM) and the inner monolayer consisting of charged Phosphatidyl ethanolamine (PE) and Phosphatidyl serine (PS). In addition to these lipids unesterified cholesterol is also found between the bilayers, dictating membrane fluidity and permeability (Borochov, Abbott et al. 1979). Therefore, whilst in keeping with published literature DMPC was chosen for initial studies, in order to slowly move towards a more complex membrane environment, the decision was made to also add cholesterol to the lipid vesicles. Cholesterol is a major component of biological lipid membranes making up as much as 10-50 % of the total composition of the Golgi body and plasma membrane (van Meer, Voelker et al. 2008). At physiological temperatures, cholesterol intercalates between phospholipid tails causing the vesicles to become more rigid. This rigidification is believed to have a positive effect on NMR measurements due to the reduced internal Brownian motions of lipids and peptides embedded in the lipid bilayer, resulting in improved spectral resolution and better line shape.

Once the lipid composition was decided upon, the most suitable method for peptide reconstitution into the bilayer needed to be established. Two basic methods for protein reconstitution into bilayers for structural studies were found in the literature at the time the project was started, and both were tested for the GpA TM peptide in this work. The first method (which will be referred to as the Detergent Removal method) is based on the slow removal of detergent from a detergent-solubilised peptide/lipid solution using Bio-beads

(Rigaud, Levy et al. 1998; Rigaud and Lévy 2003). Bio-beads are small porous polystyrene beads that trap and selectively remove detergent from solution due to the size of the pores being large enough to only allow detergent molecules to enter. Once trapped inside the bead, the hydrophobic effect prevents detergent release back into solution. This slow removal of detergent molecules from solution drives the formation of lipid vesicles as the hydrophobic tails of the lipid molecules come together to form bilayers. At the same time, the hydrophobic peptide partitions into the hydrophobic core of the newly-formed bilayer in order to maintain the non-polar environment it requires.

3.3.1 CD for screening of detergents

Due to the hydrophobic nature of the GpA_{TM} peptide, the most suitable detergent for solubilising GpA was screened that would promote the correct secondary structure (α -helical) and oligomeric state (dimer) of solubilised peptide in order to facilitate correct insertion of peptide into lipid vesicles and to avoid any peptide aggregation. A number of different detergents were screened, namely β -D-glucopyranoside (OG) a non-ionic detergent which was chosen due to its high CMC, allowing for efficient removal by Bio-beads, Dodecylphosphoglycol (DPG) a mild anionic detergent and *n*-Dodecyl β -D-maltoside (DDM) another non-ionic detergent similar to OG but with increased alkyl chain length and lower CMC. Circular dichroism (CD) spectroscopy was then used to measure the secondary structure of the detergent solubilised peptide. CD is a simple biophysical technique for determination of secondary structure of proteins and peptides via the differential absorption of circularly polarised light. **Figure 3.12** shows the CD spectrum obtained by solubilisation of GpA peptide in the organic solvent TFE which promotes alpha helicity through promoting hydrogen bonding. The results of the detergent screen using CD are shown in **Figure 3.13**. Using CD it was observed that, when solubilised in 20 DDM, the peptide took on a β -sheet secondary structure as evidenced by the peak at 217 nm, therefore DDM was discounted. Solubilisation in 20 mM DPC gave a slightly more α -helical spectrum, and 15 mM OG co-solubilised in TFE gave the most α -helical spectrum out of all detergents tested as seen by the negative minima at 210 and 222 nm, slightly red shifted from the expected 208/222 nm minima for alpha helical proteins due to being embedded in a detergent environment.

Further analysis of the CD spectra was carried out by fitting the data obtained using the DichroWeb website (Whitmore and Wallace 2004; Whitmore and Wallace 2008) using the K2D algorithm with reference set 4 (190-240 nm) the results of which are shown in **Figure 3.14**. These results indicated that OG was the best detergent for solubilisation of GpA peptide, with co-solubilisation of OG in TFE and peptide in TFE giving rise to an alpha helix content of 68% whilst OG alone gave 63% helicity. With DPC this figure dropped to 48% and out of the three detergents tested DDM gave the lowest value of 43% helicity and an increasing percentage of beta sheet structure suggesting that these two detergents were unsuitable for use in preparing GpA samples. OG also has the beneficial property of a high critical micelle concentration (23 mM), which allows for efficient removal by Biobeads and as such was the detergent we chose to use for preparing GpA samples for ssNMR.

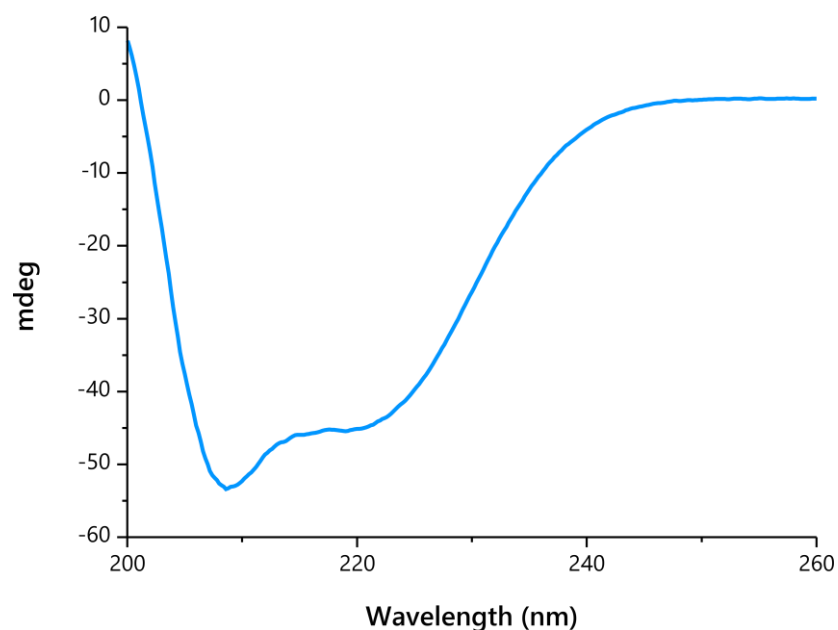


Figure 3.12 CD spectra of GpA peptide dissolved in TFE

CD spectrum recorded between 200-260 nm obtained by solubilising 0.25 mg of GpA peptide in TFE. Negative peaks at 208 and 220 nm indicate an α -helical structure of GpA.

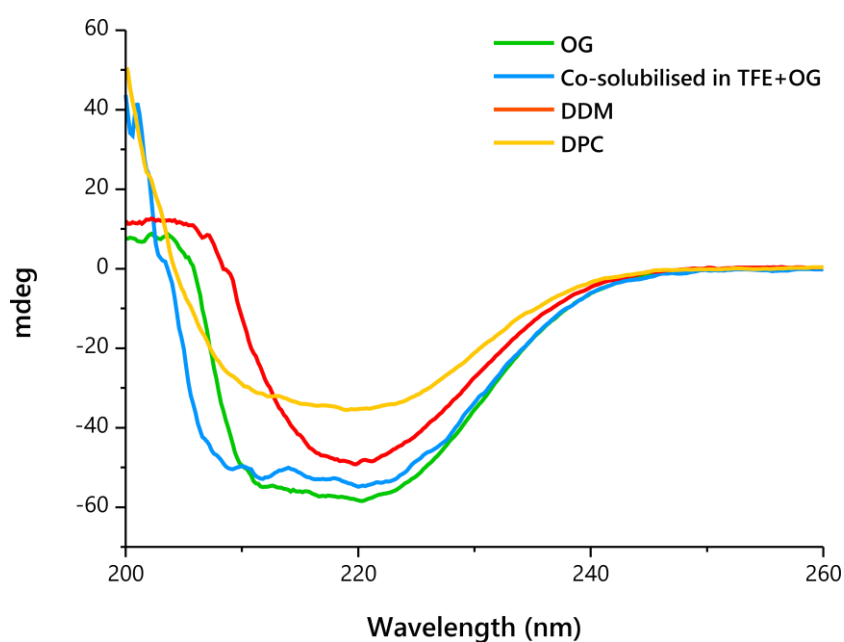


Figure 3.13 CD spectra of GpA peptide dissolved in varying detergents

CD spectra recorded between 200-260 nm obtained by solubilising 0.25 mg of GpA peptide in varying detergents in sodium phosphate buffer. Negative peaks at 210 and 224 nm (red shifted in detergent environment) indicate an α -helical structure for GpA in detergent micelles in OG solubilised samples (green line). A single negative peak at 217 nm indicates a β -sheet structure for DDM (red line)

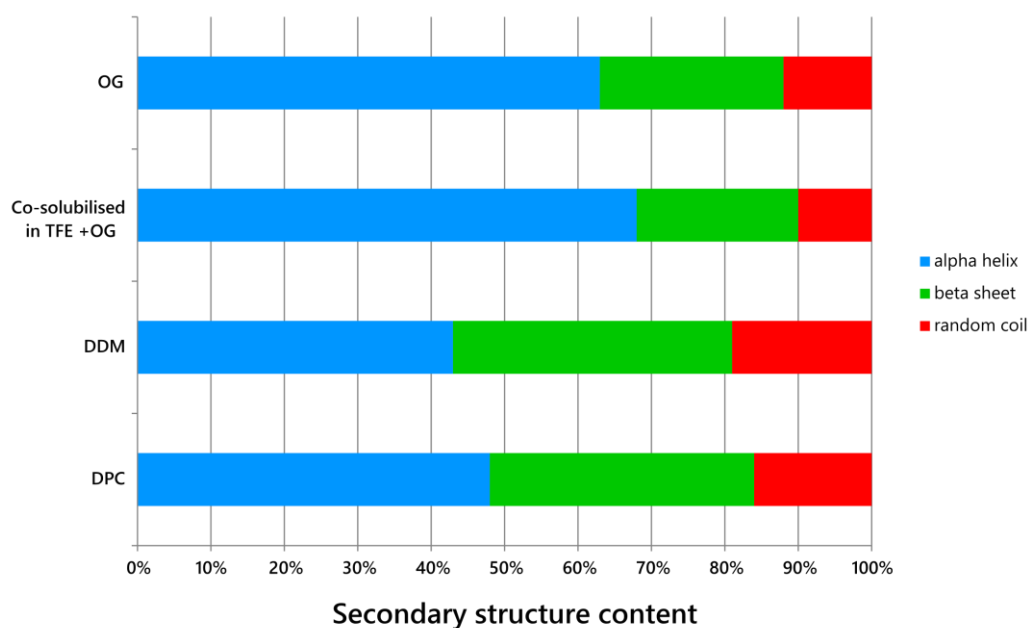


Figure 3.14 Secondary structure analysis of CD data

CD data recorded for GpA TM peptide solubilised in OG, TFE +OG, DDM and DPC were analysed using the DichroWeb website. Data was fit using the CDSSTR method and reference set 4 to give an estimation of the secondary structure content, alpha helix (blue), beta sheet (red) and random coil (green) content are shown for each of the detergents used.

The second method (which will be referred to as the co-solubilisation method) substitutes an organic solvent such as trifluoroethanol (TFE), which can be easily and almost completely removed by lyophilisation, for detergent as the initial peptide and lipid solubilisation agent. TFE also promotes helix formation for the GpA peptide (**Figure 3.12**). TFE-solubilised peptide and lipid are mixed, followed by removal of TFE by evaporation to form a film. This film is then rehydrated in aqueous buffer, following which the rehydrated film is briefly sonicated resulting in the formation of small unilamellar vesicles (SUV). Once fully resuspended in buffer, the solution is then freeze thawed in a liquid nitrogen/water bath for several cycles. This freeze-thaw cycling causes the vesicles to fuse and form large multi lamellar (LMV) and giant unilamellar vesicles (GUV) containing inserted peptide. An overview of the process for proteoliposome formation using both the detergent removal and co-solubilisation methods is shown in **Figure 3.15**.

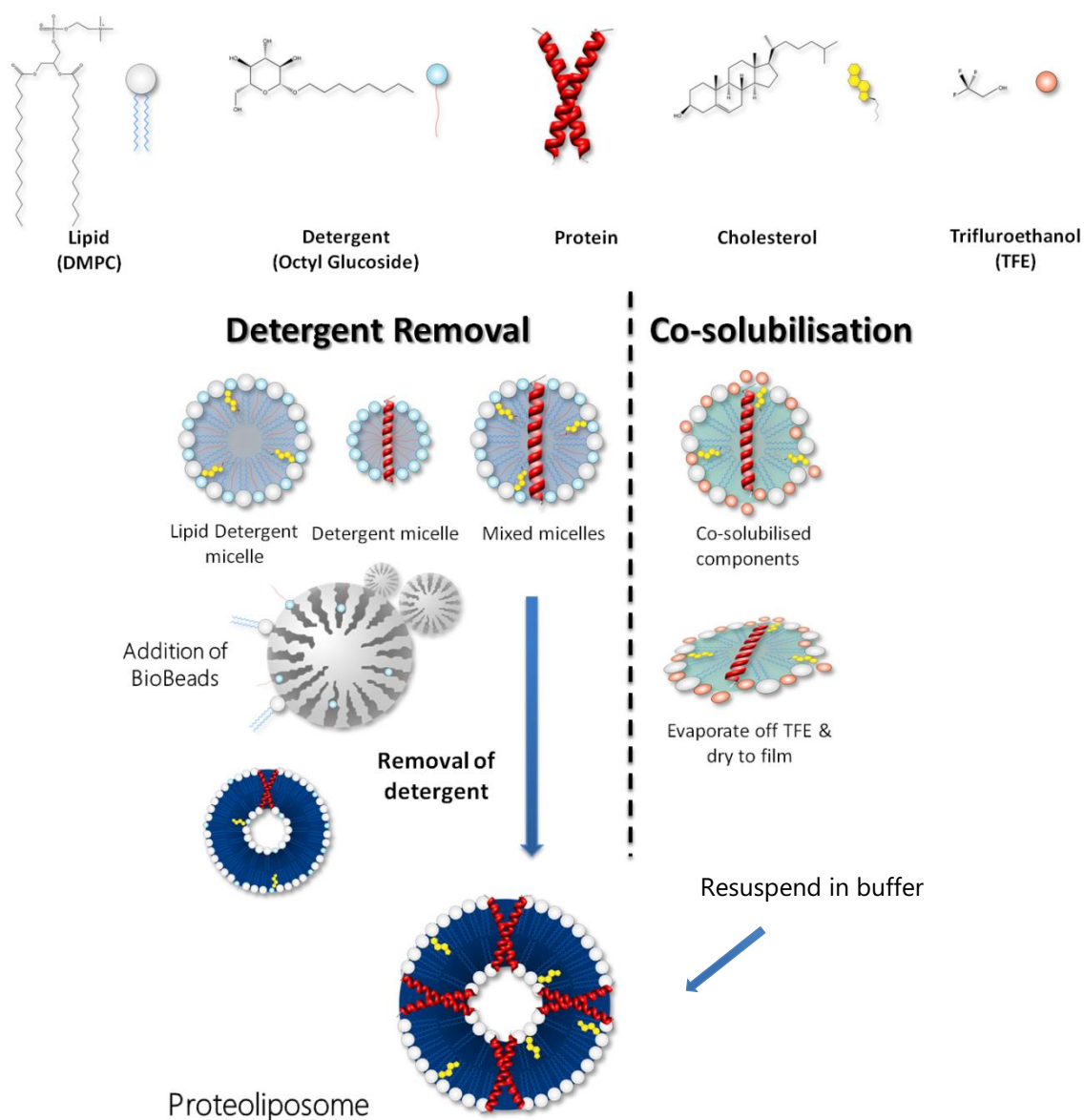


Figure 3.15 Overview of reconstitution protocols

DMPC lipid, OG detergent, GpA peptide and cholesterol were used to make ssNMR samples. For the detergent removal method, lipid and peptide were dissolved in detergent, which results in the formation of various micelles as shown above. Mixed micelles are produced upon mixing lipid and peptide together in detergent. To remove the detergent, Bio beads were added stepwise and samples agitated overnight at 4°C. The removal of detergent leads to the formation of proteoliposomes. For the co-solubilisation method, lipid and protein were dissolved in TFE and cholesterol in chloroform, mixed together and dried to a film before being resuspended in buffer to form proteoliposomes

3.3.2 Electron Microscopy

For samples prepared using the detergent removal method, the integrity of the vesicles formed was investigated using transmission electron microscopy (TEM). TEM is a powerful technique that allows visualisation of microscopic structures on a nanometre scale, as such it is ideal for screening lipid vesicles. We used EM in order to directly visualise the proteoliposomes produced by the reconstitution method in order to ascertain if our conditions favoured insertion of peptide in liposomes or if they resulted in unwanted protein aggregation. **Figure 3.16** below shows the observed images relating to the samples we prepared for solid state NMR from aliquots taken at various points during the reconstitution process. From the images obtained we can see that at the initial stage, whilst detergent is still present in the solution, small vesicles are formed (less than 100 nm). These small mixed detergent/lipid/peptide micelles increase in size as detergent is gradually removed upon stepwise addition of Biobeads, resulting in the formation of large proteoliposomes (>1000 nm) as shown in **Figure 3.16 C**. By directly viewing the proteoliposomes produced, we can see that the conditions used are likely to result in insertion of peptide into lipid vesicles. Previous experiments aimed at identifying ideal conditions (data not shown) identified conditions that resulted purely in peptide aggregation. Although the majority of the EM grid surveyed for the NMR sample contained liposomes, a small proportion contained what appears to be aggregated peptide/lipid (**Figure 3.16 D**). However, from EM images alone it is difficult to know the composition of the aggregates observed. Following EM, further biophysical characterisation of the proteoliposomes produced was carried out.

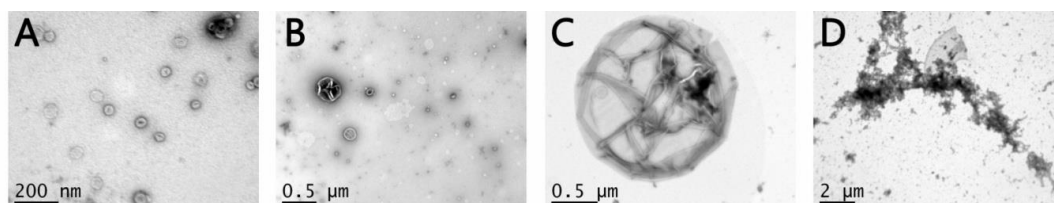


Figure 3.16 Images obtained by TEM using negative staining

Samples relating to A) mixed micelles formed during initial stage of detergent removal, B) proteoliposomes produced upon full removal of detergent, C) large proteoliposome formed in the final sample and D) aggregation observed in final sample. Scale bar relates to size.

In order to confirm whether the reconstituted GpA_{TM} peptide had correctly folded into its native α -helical secondary structure when inserted into DMPC liposomes using each of the protocols summarised in **Figure 3.15**, we again used CD. A CD spectrum for an aliquot of proteoliposomes prepared for solid state NMR was collected before proteoliposomes were pelleted in the ultracentrifuge. As shown in **Figure 3.17**, the results obtained show a typical α -helical spectrum with negative maxima at 210 and 224 nm. This is a slight red shift from what is typically seen for α -helices (208 and 222 nm) and can be attributed to the lipid environment in which the peptide is inserted (Oates, Hicks et al. 2008). As such the results obtained by the CD spectrum indicated a α -helical secondary structure for GpA reconstituted into lipid vesicles using the detergent removal method.

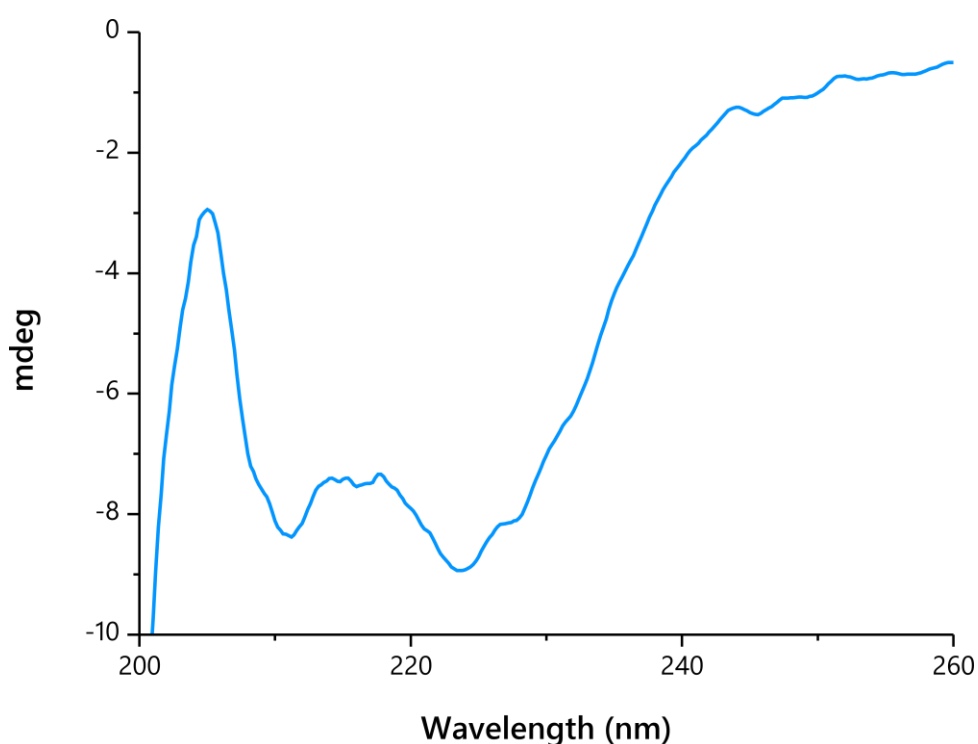


Figure 3.17 CD spectrum of GpA proteoliposome solution obtained by Capillary CD

A CD spectrum obtained using 5 μ l of GpA proteoliposome solution prepared using the detergent removal method for solid state NMR. Negative peaks at 210 and 224 nm (red shifted in lipid environment) indicate an α -helical structure of GpA in the lipid bilayer. Light scattering from the sample results in unreliable data below 205 nm, as such the negative peak at 200 nm is an artefact of this light scattering.

GpA samples prepared for ssNMR using the co-solubilisation method were also tested by CD to determine secondary structure (**Figure 3.18**). The sample prepared by this method showed an improvement in the spectra obtained due to an increase in alpha helical content from 59 to 63% as indicated by fitting the CD spectra obtained using DichroWeb, as such suggested an improvement over using the detergent removal method for sample preparation.

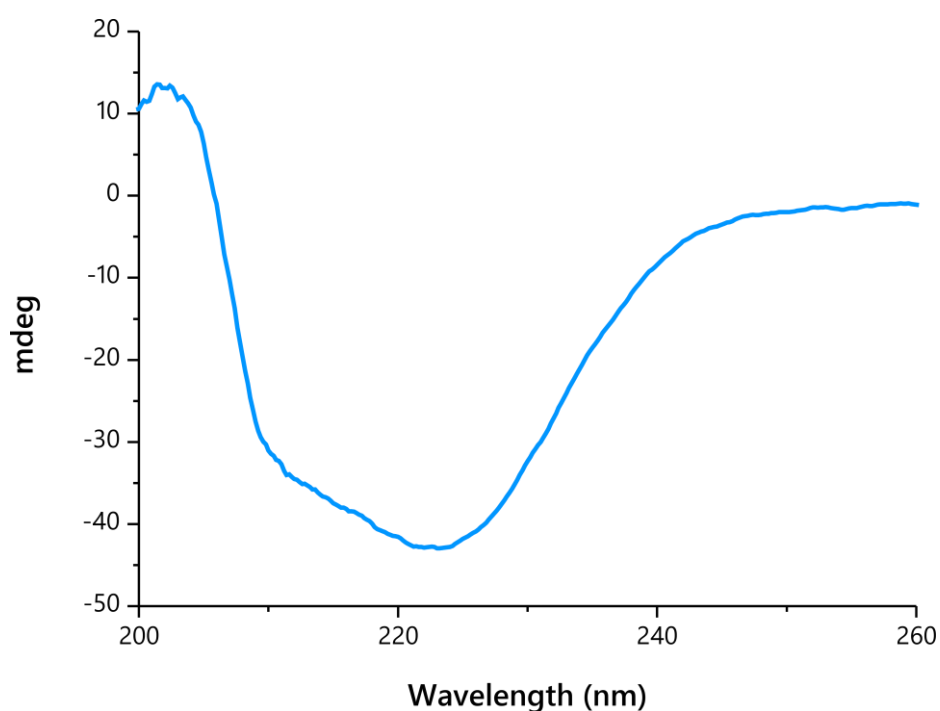


Figure 3.18 CD spectrum of GpA proteoliposome solution prepared using the co-solubilisation method

Negative peaks at 210 and 224 nm (red shifted in lipid environment) indicate an α -helical structure of GpA in the DMPC lipid bilayer. Light scattering from the sample results in unreliable data below 200nm.

3.3.3 ATR-FTIR analysis of reconstituted peptides

Attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR) was also used to confirm the secondary structure and insertion of GpA_{TM} into DMPC lipid bilayers for samples prepared for ssNMR by the detergent removal method. ATR-FTIR is a well-established technique for assessing secondary structure and orientation of TM domains in stacked lipid bilayers (see section **2.8.3**). The amide I and amide II bands

that link amino acid residues within a peptide have characteristic absorption wavelengths based on the amide I C=O bond stretching and amide II N-H bending (Herrebout, Clou et al. 2003). From the spectrum obtained in **Figure 3.18**, for GpA samples prepared using the detergent removal method, the position of the Amide I peak at 1680 cm^{-1} indicates the GpA_{TM} peptide has β -sheet secondary structure while the presence of a peak at 1650 cm^{-1} also indicates α -helical secondary structure in the peptide inserted in DMPC liposomes. GpA samples prepared by the co-solubilisation method were also analysed by FTIR as seen in **Figure 3.19** in order to compare the two sample preparation methods. Using co-solubilisation appeared to give better signal to noise in the spectrum obtained in comparison to the sample prepared by the detergent removal method. The Amide I peak was centred on 1630 cm^{-1} which again would indicate some β -sheet secondary structure. Due to high amount of noise in the data obtained these peaks cannot be correctly fitted in order to gain a more quantitative result as to the percentage helicity and beta sheet structure. This is an issue that has consistently been observed for the GpA peptide, as repeating experiments did not seem to improve the data and the peptide has also been shown to give poor results when using other techniques such as CD in comparison to other hydrophobic peptides.

Using FTIR we can also assess the insertion of peptide into bilayers, by monitoring changes in the Amide II peak upon exposure to deuterium oxide. When dissolved in water, hydrogen atoms attached to the amide nitrogen can exchange with those attached to water molecules, however if the protein is inserted into a membrane this exchange is very slow or does not occur. By exposing the dried proteoliposome film to D₂O saturated nitrogen gas, we can monitor any exchange that occurs by observing the magnitude of the Amide II peak at 1550 cm^{-1} over a period of time. If the amide hydrogens exchange with deuterium, the wavelength of absorption changes (due to the increased mass of deuterium) and the Amide II peak will shift.

Both methods of sample preparation resulted in peptides that appear to be inserted into lipid bilayers. This can be seen in **Figures 3.19** and **3.20** from the initially recorded spectrum (blue solid line) which was then followed by the recording of a second spectrum (indicated by the green line) which shows that there was no change in the intensity of the Amide II band after exposing the sample to D₂O for over 22 hours,

suggesting that there was no exchange and the GpA_{TM} peptide is fully inserted into the lipid membrane and therefore protected from deuterium exchange.

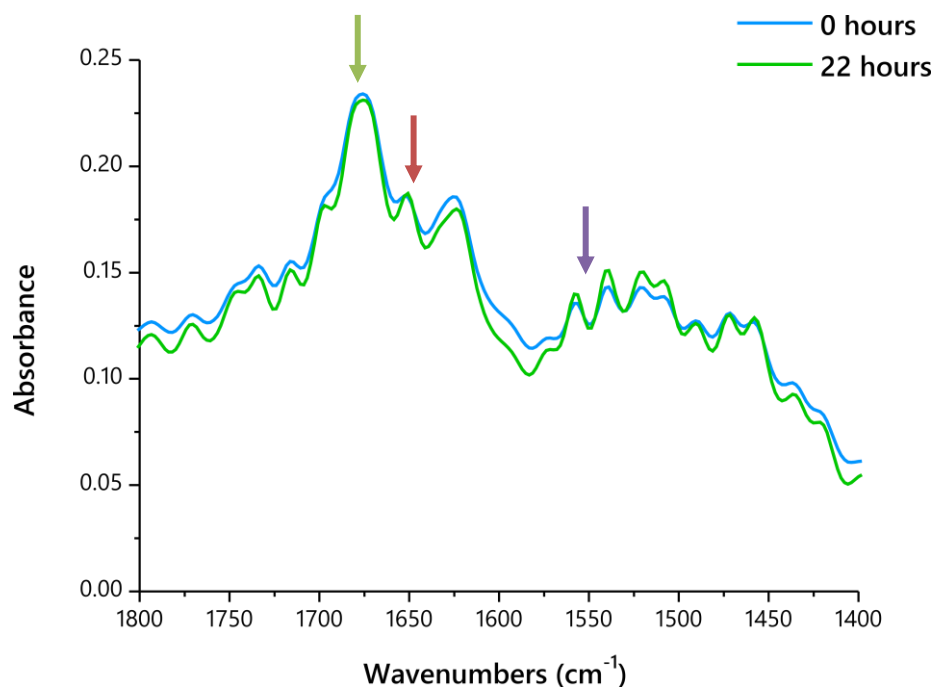


Figure 3.19 FTIR spectrum of the GpA peptide in DMPC liposomes containing 5% cholesterol prepared using the detergent removal method

The blue line represents the initial recorded spectrum, and the green line was recorded 22 hours after exchange with D₂O by passing saturated nitrogen gas over the sample in a sealed chamber. The position of the Amide I peak at 1680 cm⁻¹ (green arrow) indicates β -sheet secondary structure, the presence of a peak at 1650 cm⁻¹ (red arrow) also suggests some α -helix secondary structure. The magnitude of the Amide II peak at 1550 cm⁻¹ (purple arrow) after 22 hours of D₂O exposure indicated complete bilayer insertion.

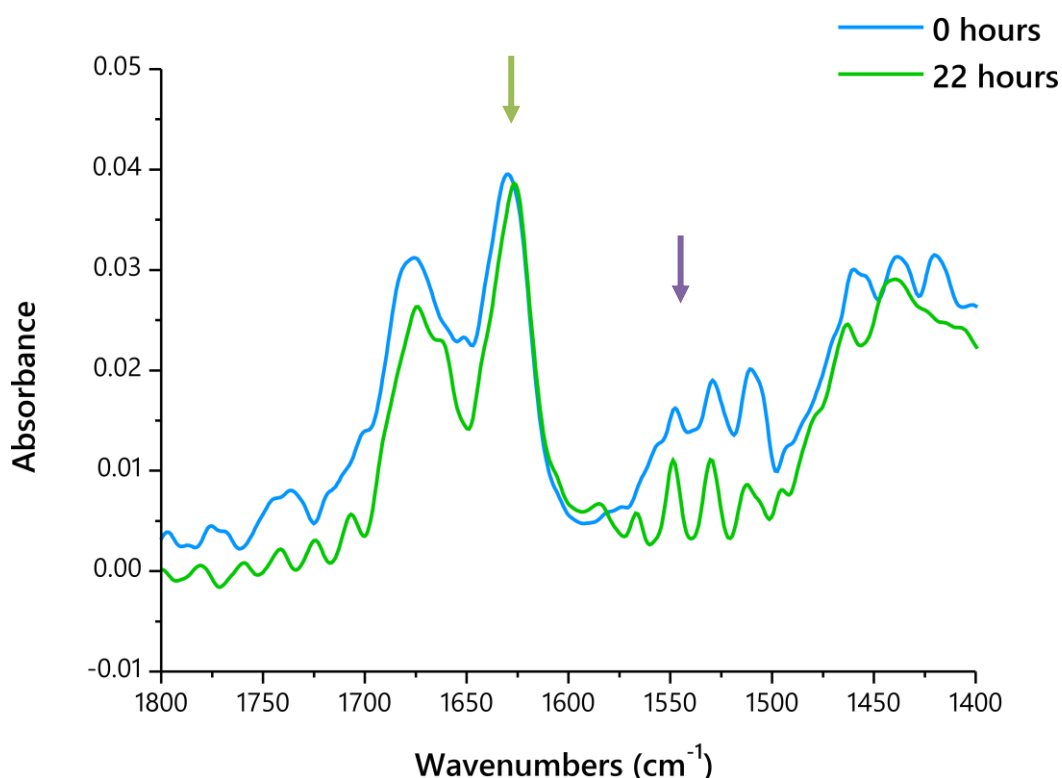


Figure 3.20 FTIR spectrum of the GpA peptide in DMPC liposomes containing 5% cholesterol prepared using the co-solubilisation method

The solid blue line represents the initial recorded spectrum, and the green line was recorded 22 hours after exchange with D₂O. The position of the Amide I peak at 1630 cm⁻¹ (green arrow) indicates some β -sheet secondary structure, and the magnitude of the Amide II peak at 1550 cm⁻¹ (purple arrow) after 22 hours of D₂O exposure indicated complete bilayer insertion.

Using the results described thus far for the reconstitution of the GpA peptide into bilayers, it was decided that the co-solubilisation method produced samples containing helical peptides fully inserted in lipid bilayers without requiring the much longer sample preparation times and increased beta sheet content that appeared to occur with the detergent removal method as evident also by a comparison of 1D ¹³C spectra obtained using the detergent removal method and the co-solubilisation method **Figure 3.21** shows an improvement in the overall homogeneity of the sample prepared. The secondary species observed in the sample prepared using detergent removal method were significantly reduced in the samples produced by co-solubilisation, and

the signal to noise was also improved. Therefore, the co-solubilisation method was also used for BPV E5 samples.

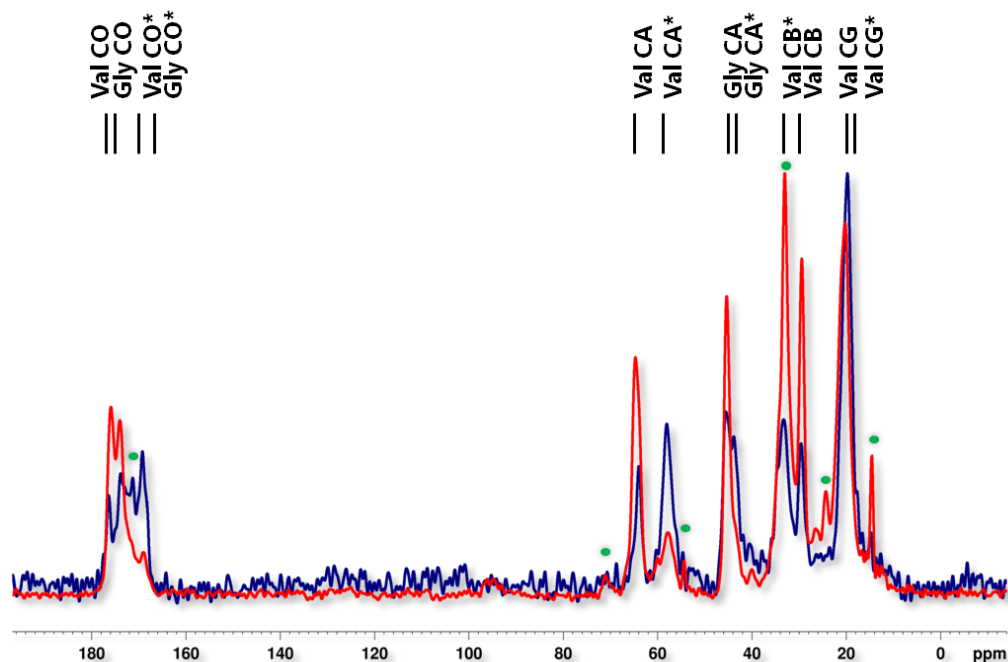


Figure 3.21 1D ^{13}C spectra obtained from samples prepared by detergent removal (blue) and co-solubilisation (red)

1D ^{13}C spectra obtained from samples prepared by detergent removal (blue) and co-solubilisation (red) with a 2:1 (w/w) LPR. Peaks from labelled valine and glycine carbon atoms are indicated, peaks arising from secondary species indicated with *, DMPC lipid peaks indicated by green dots.

3.3.4 Co-solubilisation of E5 peptides

For E5, samples prepared using this method were also tested by FTIR to confirm that the samples were suitable for further analysis by ssNMR. As seen in **Figure 3.22**, E5 samples prepared using the co-solubilisation method also resulted in the formation of liposomes with alpha helical peptides as indicated by the Amide I peak at 1650 cm^{-1} , although this peak has a slight shoulder at 1675 cm^{-1} which also has a contribution to the Amide I region. This shoulder can be attributed to residual trifluoroacetic acid (TFA) from the peptide purification process that interferes with this result, as is evidenced by the reduction in this peak following exposure to D_2O saturated nitrogen gas. From the area beneath the Amide II peak the insertion of E5 into DMPC liposomes can be

calculated. It can be seen that the Amide II peak does not shift position following 22 hours of exposure to D₂O but the magnitude of the peak does decrease, as does that of the overall spectrum. This decrease in overall intensity can be attributed to sample swelling. To account for sample swelling the ratio of the area beneath the Amide II peak can be normalised to that of the area beneath the Amide I peak, any decrease in this ratio following D₂O exchange would suggest that the peptide was not fully inserted and exchange with deuterium was occurring. For BPV E5 the initial $A_{\text{AmideII}}:A_{\text{AmideI}}$ ratio at 0 hours was 0.53, after 22 hours this ratio was 0.25, this represents a value of 52.8% exchange. This value may have been affected by the presence of the TFA shoulder to the Amide I peak, making integration difficult resulting in a higher apparent exchange value.

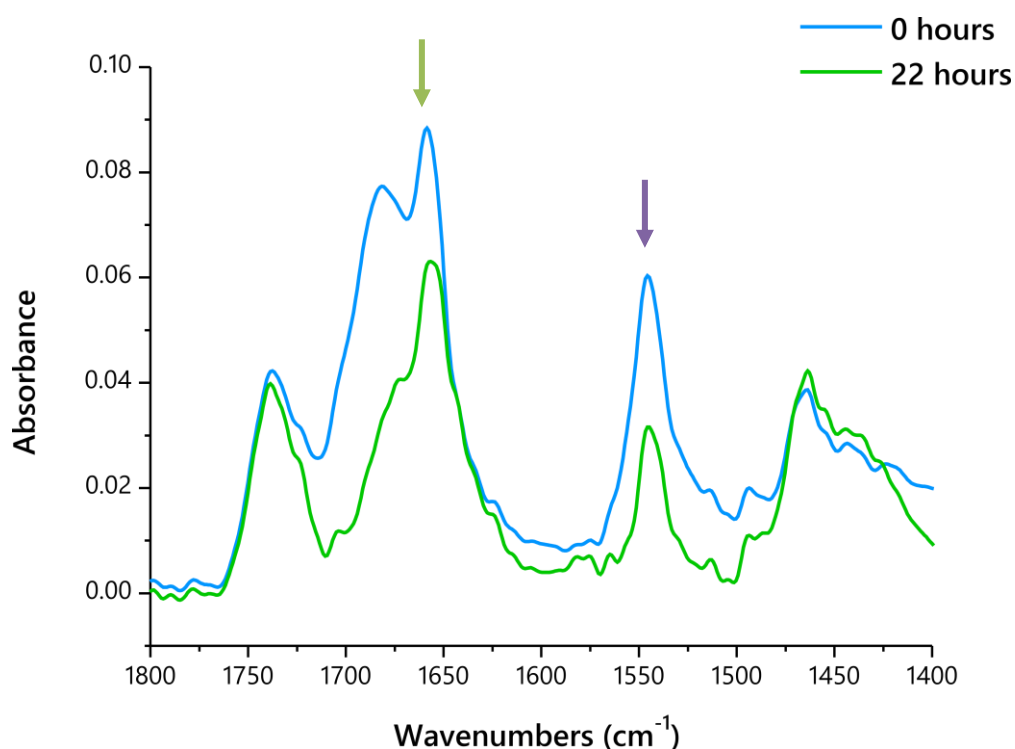


Figure 3.22 FTIR spectrum of the E5_{TM} peptide in DMPC liposomes prepared using the co-solubilisation method

The blue line represents the initial recorded spectrum, and the green line was recorded 22 hours after exchange with D₂O saturated nitrogen passed over the sample in a sealed chamber. The position of the Amide I peak at 1650 cm⁻¹ (green arrow) indicates α -helical secondary structure, and the ratio of the magnitude of the Amide II peak at 1550 cm⁻¹ (purple arrow) to Amide I after 22 hours of D₂O exposure indicated bilayer insertion with some exchange.

In order to further confirm whether the reconstituted BPV E5 peptide had correctly folded into its native α -helical secondary structure when inserted into DMPC liposomes, we again used CD spectroscopy. A CD spectrum for an aliquot of proteoliposomes prepared for solid state NMR was collected before proteoliposomes were pelleted in the ultracentrifuge. As shown in **Figure 3.23**, the results obtained show a typical α -helical spectrum with negative maxima at 210 and 224 nm. This is a slight red shift from what is typically seen for α -helices (208 and 222 nm) and can be attributed to the lipid environment in which the peptide is inserted. As such the results obtained by CD spectroscopy indicated a predominantly α -helical secondary structure for BPV E5 reconstituted into lipid vesicles.

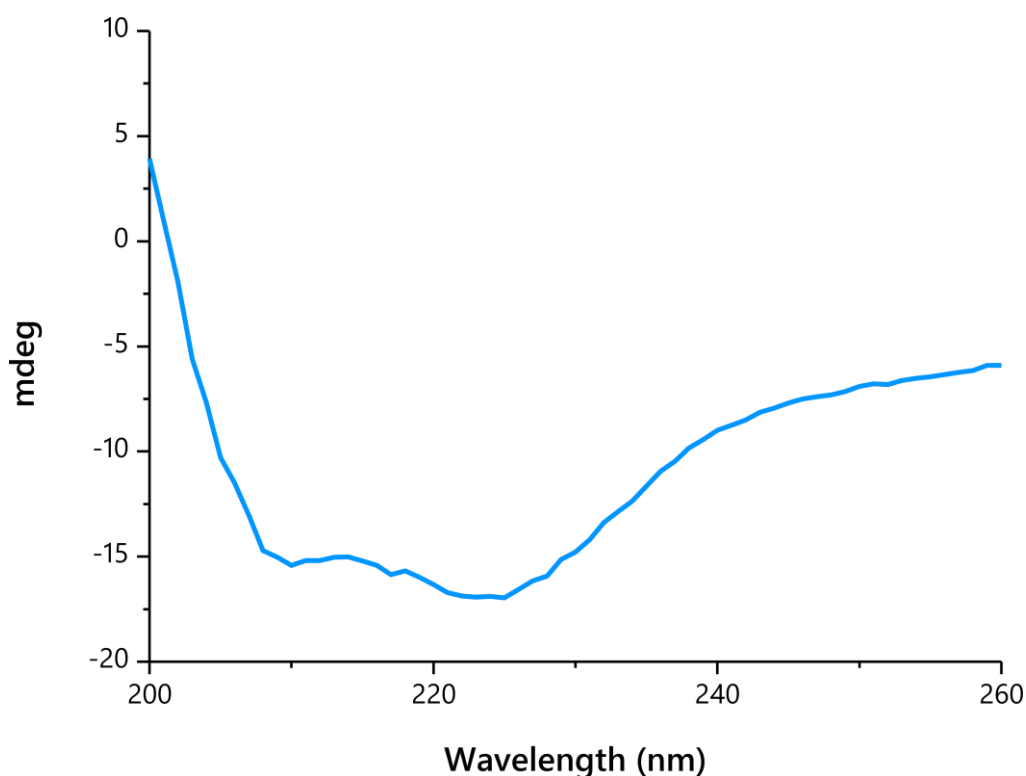


Figure 3.23 CD spectrum of BPV E5 proteoliposome solution prepared for solid state NMR using the co-solubilisation method

Negative peaks at 210 and 224 nm (red shifted in lipid environment) indicate an α -helical structure of GpA in the lipid bilayer. Light scattering from the sample results in unreliable data below 205 nm, as such the negative peak at 200 nm is an artefact of this light scattering.

3.4 Summary

In this chapter, results are presented for two methods which were used to prepare proteoliposome samples for ssNMR. The two methods tested were (i) a well-established detergent removal method (Rigaud, Levy et al. 1998) and (ii) a method based upon co-solubilisation using organic solvents. Both methods were tested, but initially samples were prepared using the more popular detergent removal method.

The detergent removal method was, at the start of this study one of well-established method for reconstitution of membrane proteins into lipid vesicles and currently is still a popular method when expressing membrane proteins, which typically require that they be detergent solubilised in order to extract them from cell membranes. Typically the detergent removal method utilises some form of dialysis in order to remove detergent molecules from solution resulting in the formation of lipid vesicles into which hydrophobic peptides partition. This can be done using dialysis membranes, but the small volumes used in such sample preparations make this method difficult and as such Bio-beads can be used. Bio-beads also result in faster detergent removal, thereby cutting down the time required to produce samples.

Bio-beads were used to prepare samples for ssNMR when using the detergent removal method which was optimised in a number of ways such as the stepwise addition of Bio-beads in order to more slowly remove detergent and produce well formed proteoliposomes. Proteoliposomes formed in this manner were then characterised using a number of bio-physical techniques such as CD and FTIR, as well as electron microscopy (EM) in order to confirm the insertion and fold of peptides reconstituted into lipid vesicles and to ensure the conditions used did not result in aggregation of peptides. From the bio-physical characterisation it was seen that whilst the ssNMR samples prepared using the detergent removal method resulted in peptides that were inserted into lipid vesicles, and the majority of peptide reconstituted into proteoliposomes was α -helical there was also a small population of peptide that was incorrectly folded into β -sheet form. This beta sheet population was likely due to aggregation of peptide as a result of the detergent removal process as it was less visible in earlier steps of the reconstitution process. The most likely explanation would be the aggregates form due to the cholesterol interactions in the sample, as unlike the other

components of the solution was much harder to solubilise. This was also seen to manifest itself in the form of a secondary species resulting in a second set of ssNMR signals suggesting a second chemical environment for the peptide. Therefore, another method for sample preparation was investigated.

Using the co-solubilisation method we prepared samples by solubilising lipid/cholesterol and peptide in organic solvent, which was then fully removed, followed by resuspension in buffer to form proteoliposomes. Just as with samples prepared by the detergent removal method, the samples prepared by co-solubilisation were also characterised using CD, FTIR and NMR spectra. Samples prepared with this method showed an improvement in the form of a reduced beta-sheet population by CD spectroscopy and improved overall spectral quality by FTIR spectra although the results for GpA from FTIR suggested that both methods resulted in the formation of a β -sheet secondary structure. It has been shown that CD is more sensitive to alpha helical secondary structure than FTIR and this could explain why CD indicates that the co-solubilisation method lead to an increase in helical content. This improvement in reduction in β -sheet secondary species in samples produced using the co-solubilisation method when compared to GpA samples prepared by the detergent removal method, was most evident when analysing each sample by ssNMR. From 1D ^{13}C experiments (see **Chapter 4, Section 4.2**) it could be seen that samples prepared using the co-solubilisation method resulted in improved signal to noise and a much reduced β -sheet secondary species.

The co-solubilisation method not only resulted in more homogenous samples with improved secondary structure but also resulted in a lower loss of peptide from the sample as seen with the detergent removal method, which due to having to decant bio-beads from resuspended solution also appeared to result in loss of sample. Overall the co-solubilisation method was found to be a much simpler sample preparation method to use and was the method of choice for the preparation of BPV E5 and GpA ssNMR samples.

3.5 References

- Adams, P. D., I. T. Arkin, D. M. Engelman and A. T. Brunger (1995). "Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban." *Nat Struct Biol* **2**(2): 154-162.
- Beevers, A. J. and A. M. Dixon (2010). "Helical membrane peptides to modulate cell function." *Chem Soc Rev* **39**(6): 2146-2157.
- Borochoy, H., R. E. Abbott, D. Schachter and M. Shinitzky (1979). "Modulation of erythrocyte membrane proteins by membrane cholesterol and lipid fluidity." *Biochemistry* **18**(2): 251-255.
- Cady, S. D., T. V. Mishanina and M. Hong (2009). "Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding." *J Mol Biol* **385**(4): 1127-1141.
- DiMaio, D., D. Guralski and J. Schiller (1986). "Translation of open reading frame E5 of bovine papillomavirus is required for its transforming activity." *Proc Natl Acad Sci U S A* **83**(6): 1797-1801.
- Herrebout, W., K. Clou, H. Desseyn and N. Bleton (2003). "Vibrational characterization of the peptide bond." *Spectrochim Acta A Mol Biomol Spectrosc* **59**(1): 47-59.
- Ketchum, R. R., W. Hu and T. A. Cross (1993). "High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR." *Science* **261**(5127): 1457-1460.
- King, G., J. Oates, D. Patel, H. A. van den Berg and A. M. Dixon (2011). "Towards a structural understanding of the smallest known oncoprotein: investigation of the bovine papillomavirus E5 protein using solution-state NMR." *Biochim Biophys Acta* **1808**(6): 1493-1501.
- Klein, O., G. Polack, T. Surti, D. Kegler-Ebo, S. Smith and D. DiMaio (1998). "Role of glutamine 17 of the bovine papillomavirus E5 protein in platelet-derived growth factor beta receptor activation and cell transformation." *J Virol* **72**(11): 8921-8932.
- Koradi, R., M. Billeter and K. Wuthrich (1996). "MOLMOL: a program for display and analysis of macromolecular structures." *J Mol Graph* **14**(1): 51-55, 29-32.
- Lemmon, M. A., J. M. Flanagan, H. R. Treutlein, J. Zhang and D. M. Engelman (1992). "Sequence specificity in the dimerization of transmembrane alpha-helices." *Biochemistry* **31**(51): 12719-12725.
- Lemmon, M. A., H. R. Treutlein, P. D. Adams, A. T. Brunger and D. M. Engelman (1994). "A dimerization motif for transmembrane alpha-helices." *Nat Struct Biol* **1**(3): 157-163.
- MacKenzie, K., J. Prestegard and D. Engelman (1997). "A transmembrane helix dimer: structure and implications." *Science* **276**(5309): 131-133.
- Oates, J., M. Hicks, T. R. Dafforn, D. DiMaio and A. M. Dixon (2008). "In vitro dimerization of the bovine papillomavirus E5 protein transmembrane domain." *Biochemistry* **47**(34): 8985-8992.
- Popot, J. and D. Engelman (1990). "Membrane protein folding and oligomerization: the two-stage model." *Biochemistry* **29**(17): 4031-4037.
- Rigaud, J. and D. Lévy (2003). "Reconstitution of membrane proteins into liposomes." *Methods Enzymol* **372**: 65-86.
- Rigaud, J. L., D. Levy, G. Mosser and O. Lambert (1998). "Detergent removal by non-polar polystyrene beads - Applications to membrane protein reconstitution and two-dimensional crystallization." *European Biophysics Journal with Biophysics Letters* **27**(4): 305-319.

- Smith, J. E. (1987). "Erythrocyte membrane: structure, function, and pathophysiology." Vet Pathol **24**(6): 471-476.
- Smith, S. and B. Bormann (1995). "Determination of helix-helix interactions in membranes by rotational resonance NMR." Proc Natl Acad Sci U S A **92**(2): 488-491.
- Smith, S., R. Jonas, M. Braiman and B. Bormann (1994). "Structure and orientation of the transmembrane domain of glycophorin A in lipid bilayers." Biochemistry **33**(20): 6334-6341.
- Smith, S., D. Song, S. Shekar, M. Groesbeek, M. Ziliox and S. Aimoto (2001). "Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers." Biochemistry **40**(22): 6553-6558.
- Surti, T., O. Klein, K. Aschheim, D. DiMaio and S. Smith (1998). "Structural models of the bovine papillomavirus E5 protein." Proteins **33**(4): 601-612.
- van Meer, G., D. R. Voelker and G. W. Feigenson (2008). "Membrane lipids: where they are and how they behave." Nat Rev Mol Cell Biol **9**(2): 112-124.
- Wang, J., S. Kim, F. Kovacs and T. A. Cross (2001). "Structure of the transmembrane region of the M2 protein H(+) channel." Protein Sci **10**(11): 2241-2250.
- Whitmore, L. and B. A. Wallace (2004). "DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data." Nucleic Acids Res **32**: W668-W673.
- Whitmore, L. and B. A. Wallace (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases." Biopolymers **89**(5): 392-400.



SOLID STATE NMR

ANALYSIS OF

RECONSTITUTED GPA

4.1 Introduction

The first section of this chapter describes the use of the doubly labelled GpA peptide initially used for study by solid-state NMR (ssNMR), and a comparison between the sample preparation methods used in order to produce samples for ssNMR. Initial one- and two-dimensional data were used to determine whether through space couplings could be observed between the labelled amino acids and identify some of the challenges involved in sample preparation. Following this, the use of singly labelled peptides and the rationale behind this method is described, in addition to further evidence for the observation of through space inter-helical interactions between labelled amino acids. Finally some preliminary data making use of isotopically labelled nitrogen atoms obtained using z-filtered TEDOR experiments are also presented.

GpA was selected as a model transmembrane peptide to compare sample preparation methods developed for studying small transmembrane peptides in lipid environments. GpA was chosen since it has been well characterised by solution NMR (with a 3D structure already deposited at the protein databank (pdb file: 1AFO, (MacKenzie, Prestegard et al. 1997)) and also by ssNMR (Smith, Jonas et al. 1994; Smith, Song et al. 2001; Smith, Eilers et al. 2002). This makes GpA ideal for use as a model system for our ssNMR experiments and also allows us to exploit the strategy of using synthetic peptides for this study. Synthetic peptides allow for the selective uniform labelling of individual amino acids at key positions in the GpA peptide, and the interaction of these labelled amino acids upon dimerisation of the peptides in lipid vesicles can be studied by ssNMR.

Due to the inherent broadness of peaks in solid state NMR, the uniform isotopic labelling of amino acids in protein samples (especially those studied in fully hydrated

lipid bilayers) is less commonplace due to the spectral crowding that occurs, making assignment of resonances difficult. The problem of spectral crowding is often combated by the use of 1, 3 ^{13}C labelled glucose in the case of expressed membrane proteins. This yields specific amino acid labelling patterns which, when used in conjunction with various 3D experiments (NCOCX, NCACX etc.), makes the assignment process easier (Higman, Flinders et al. 2009). Our aim was to use selective labelling at a TM homodimer interface in order to observe through space couplings between chosen amino acids, thereby reducing spectral crowding. To test this method, we used the well characterised membrane protein GpA to see if we could observe interactions between those amino acids predicted to lie at the dimer interface as previously published (MacKenzie, Prestegard et al. 1997; Smith, Song et al. 2001). The rationale behind selecting the amino acids to label is given in the previous chapter (**Chapter 3.1.1**). Once a working protocol was established for studying small transmembrane proteins in this manner, we hoped to apply the same labelling techniques to the Bovine Papillomavirus E5 protein for which no current 3D or atomic structure currently exists.

4.2 ^{13}C NMR of doubly labelled GpA_{VG} peptide reconstituted using the detergent removal method

Magic angle spinning (MAS) solid state NMR (MAS ssNMR) was used to study the structure of the homodimeric GpA_{TM} peptide in a lipid bilayer environment, using a peptide (GpA_{VG}) corresponding to the TM region of GpA into which uniformly ^{13}C and ^{15}N -labelled valine 80 and glycine 83 (**Figure 4.1**) residues were incorporated. This peptide was reconstituted into DMPC lipid vesicles using the detergent removal method as described in the **Chapter 2 Methods section. 2.7.2**. Solid state NMR is advantageous as it allows the study of membrane proteins in a "native-like" lipid environment, unlike solution NMR which most commonly makes use of detergent micelles for solubilisation of hydrophobic membrane proteins. These micelles, due to their spherical shape and small diameter, can cause curvature stress to the embedded transmembrane protein resulting in distortion of the membrane protein structure. For systems that yield spectra containing few peaks (such as selectively-labelled peptide), one-dimensional (1D) NMR

can be used to gain initial structural information about the environment of labelled amino acids through measurement of chemical shifts.

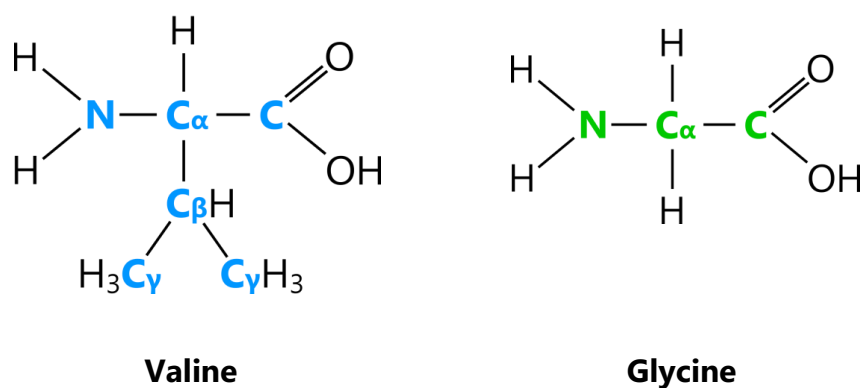


Figure 4.1 Chemical structure of valine and glycine.

Chemical structures of U-¹³C/¹⁵N labelled valine and glycine are shown, with isotopically labelled carbons and nitrogens indicated in colour.

4.2.1 1D ¹³C ssNMR spectra of doubly labelled (GpA_{VG}) peptide

One-dimensional ¹³C NMR measurements were used to gain chemical shift information for the GpA peptide in DMPC lipid bilayers for each of the labelled amino acids valine 80 and glycine 83. Due to the inherent low sensitivity of solid state NMR, cross polarisation (CP) is a technique that is used to transfer magnetisation from highly abundant nuclei with a high gyromagnetic ratio (e.g. ¹H) to a less abundant nucleus with a lower gyromagnetic ratio (e.g. ¹³C) in order to increase sensitivity and obtain better signal-to-noise ratios for the low abundance nuclei. The CP experiment uses Hartmann-Hahn matched RF pulses in order to transfer magnetisation. A 1D ¹³C MAS ssNMR spectrum of the GpA_{VG} sample obtained using a cross polarisation (CP) experiment is shown in **Figure 4.2**. All experiments were carried out at low temperature (−15 °C) where DMPC is in the gel phase (Needham and Evans 1988) so that internal motions within the protein, such as side-chain dynamics are reduced to a minimum in order to improve line widths and to allow for efficient polarisation transfer. (Abdine, Verhoeven et al. 2010)

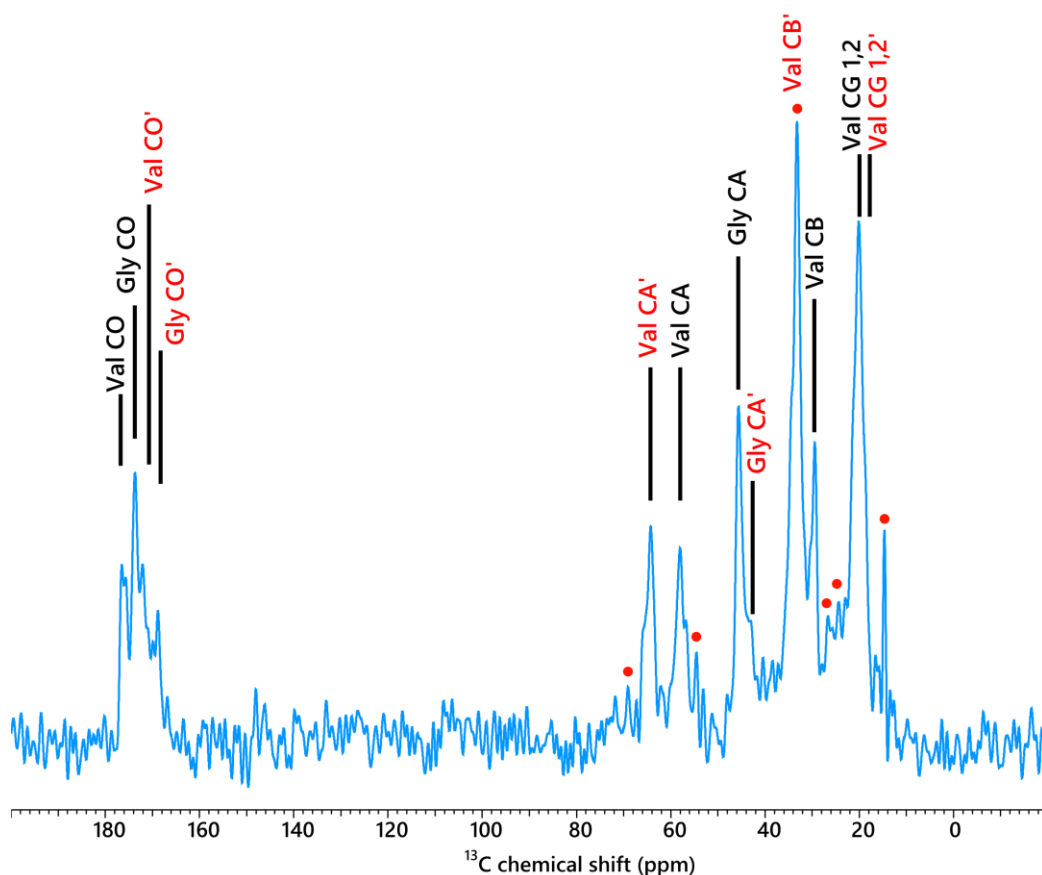


Figure 4.2 1D proton-decoupled ^{13}C CP-MAS spectrum of doubly labelled GpA_{VG} in DMPC liposomes prepared using the detergent removal method

Doubly labelled GpA_{VG} sample prepared using the detergent removal method (as described in Chapter 2.7.2). Spectrum recorded at 500 MHz with 8.5 kHz MAS at 258 K (−15 °C) for 1024 co-added transients. Resonances arising from U- ^{13}C labelled valine and glycine are labelled accordingly with resonances arising from a secondary species labelled in red. Resonances arising from natural abundance ^{13}C lipid are indicated by red circles (•).

In the 1D ^{13}C spectrum of GpA, in addition to resonances arising from labelled amino acids we can observe a number of resonances which arise from the large number of DMPC lipid molecules in the sample which result from natural abundance ^{13}C . As such a natural abundance ^{13}C spectrum of a DMPC and 5% cholesterol only, sample was recorded (**Figure 4.3**) and the resonances observed, assigned based on previously reported chemical shifts for DMPC (Lee and Griffin 1989) (summarised in **Table 4.1**).

The ^{13}C chemical shifts of added cholesterol in the sample were undetermined, but due to the low concentration of cholesterol present in the sample, any contribution was deemed too small to be visible in the obtained spectra.

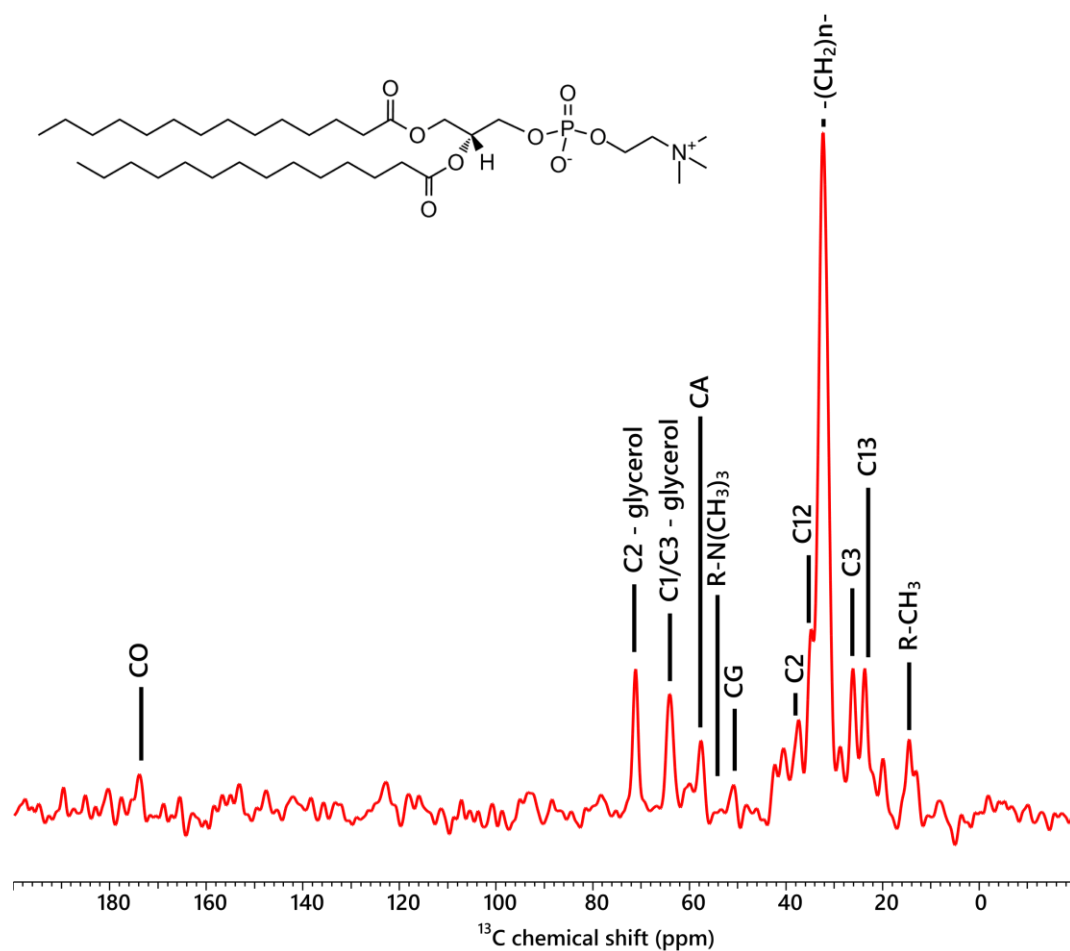


Figure 4.3 1D proton-decoupled ^{13}C CP-MAS spectrum of DMPC with 5% cholesterol

Spectrum of natural abundance ^{13}C signals from DMPC liposomes containing 5% cholesterol. Recorded at 500 MHz with 11 kHz MAS at 258 K (25 °C). A CP contact time of 750 μs was used, a 2.5 second recycle delay with 100 kHz proton decoupling during acquisition was applied and 512 co-added transients recorded. -1.00 Hz GM line broadening was applied before Fourier transformation. Chemical structure of DMPC is shown in the inset.

Table 4.1 List of previously assigned dimyristoyl-*sn*-phosphatidylcholine (DMPC) resonances

¹³C DMPC chemical shift data as assigned from previously published experimental data on DMPC only lipid bilayers studied by MAS ssNMR (Lee and Griffin 1989).

Assignment	¹³ C chemical shift (ppm)
R-CH ₃	14.84
C13 methylene	24.11
C3 methylene	26.07
-(CH ₂) _n -	31.44
C12 methylene	33.39
C2 methylene	35.34
R-N(CH ₃) ₃	55.36
α-methylene	60.73
c1/c3-glycerol	64.15
β methylene	67.08
C2-glycerol	71.86
C=O (carbonyl)	174.50

The resonances arising from U-¹³C labelled valine and glycine within the GpA sample were assigned using published average chemical shift values for valine and glycine as found at the Biological Magnetic Resonance Databank (BMRB) <http://www.bmrwisc.edu/> (Ulrich, Akutsu et al. 2008). The corresponding C_α, C_β, C_{γ1}/C_{γ2} and CO chemical shifts for ¹³C labelled carbon atoms in valine and CO and C_α for glycine incorporated into the GpA peptide, (given in **Table 4.2**) were found to match well with those values published at the BRMB. Whilst resonances observed arising from DMPC appeared as sharp peaks (typically 0.6-1.5 ppm/83-194 Hz peak width at half height), due to the fast internal Brownian motions of the lipid molecules (Saffman and Delbruck 1975), resonances arising from the labelled amino acids were observed to be broader in general (typically 1.2-2.6 ppm/152-337 Hz peak width at half height). Due to

the broadness of these peaks, some of the resonances had considerable overlap. For example, the valine and glycine carbonyl carbon atoms and the valine C γ 1/2 carbons were unresolvable as single peaks and rather gave just a single peak with a slight shoulder. A few resonances arising from labelled amino acids were seen to overlap with those arising from DMPC, namely valine C γ 1/2 and valine C β . All other resonances were reasonably well-resolved with good peak shape and with an average signal to noise ratio of 46.9.

Table 4.2 ^{13}C chemical shift data for labelled GpA Val 80 and Gly 83

Val 80 ^{13}C chemical shifts (ppm)	
CO	176.10
C α	64.77
C β	29.55
C γ 1,2	20.93

Gly 83 ^{13}C chemical shifts (ppm)	
CO	173.81
C α	45.71

In addition to the expected resonances arising from labelled amino acid carbons as discussed above, an additional number of secondary resonances were also observed within the spectra. These had slightly different chemical shifts compared to the peaks of the major species, suggesting the presence of a second species arising from protein in a second environment within the sample. The chemical shifts for this secondary species are given in **Table 4.3**. This secondary set of peaks suggested the presence of protein in an alternative second environment and therefore in order to understand these two environments secondary shift analysis was used.

Table 4.3 ^{13}C chemical shift data for secondary species in labelled GpA Val 80 and Gly 83

Val 80' ^{13}C chemical shifts (ppm)	
CO'	172.73
C α '	58.26
C β '	33.27
C γ 1,2'	19.48

Gly 83' ^{13}C chemical shifts (ppm)	
CO'	169.04
C α '	43.83

4.2.2 Secondary shift analysis

Using the chemical shift index (CSI) method (Spera, Ikura et al. 1991) and data from the Biological Magnetic Resonance Databank (BMRB), the experimental chemical shifts observed for C α , C β , C γ 1,2 and CO can be compared to those values published for random coil amino acids in order to get an approximation of the secondary structure of the protein. By comparing the shifts for the two species with those in the published database, secondary chemical shifts $\Delta\delta_S^i$ can be calculated using **Equation 8**, where δ_{obs}^i is the observed chemical shift and δ_{rc}^i is the random coil value. Since the chemical shift values stated in the BMRB are referenced to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) whereas our experimental shifts are referenced to alanine on the TMS (Tetramethylsilane) scale, these values have to be converted to DSS referenced shifts by adding 2 ppm (Wishart 2011).

$$\Delta\delta_S^i = \delta_{obs}^i - \delta_{rc}^i \quad (8)$$

Chemical shifts for all labelled carbons in GpA are given in **Table 4.4** alongside those for the corresponding average random coil chemical shifts published at the BMRB. Using this method and **Equation 8**, it is suggested that if the secondary shift for C α and CO is positive, that residue is part of an α -helical secondary structure. Should the protein have a β -sheet secondary structure, then the secondary chemical shifts should be negative.

As can be seen in **Table 4.4** and in the graph shown in **Figure 4.4**, the major species yields positive values for the secondary shifts of valine CO, C α and C γ , a negative secondary shift for valine C β , and positive secondary shifts for glycine CO and C α . This would suggest a predicted α -helical secondary structure for the primary species observed in our ssNMR GpA sample. The peaks arising from the secondary species present in our GpA ssNMR sample yields very different secondary shift values, with negative values for valine CO, C α and C γ , a positive value for valine C β , and a negative value for glycine CO (a slightly positive value is observed for glycine C α). These results suggest that the secondary species that we observe in our GpA ssNMR samples prepared using the detergent removal method has β -sheet secondary structure or is a result of aggregated peptide, in an alternative chemical environment as seen in the electron microscopy images in **Chapter 3 Figure 3.16 D**.

Table 4.4 Comparison of experimental GpA chemical shifts compared to random coil values

Averaged GpA chemical shifts from experimentally derived major (top half of table) and secondary species, denoted by ' (bottom half of table). Chemical shift values obtained by extraction of rows from 2D experiments. Chemical shift values were converted to DSS referenced values by the addition of +2 ppm. BMRB average chemical shift values (random coil) then subtracted to give secondary shift values.

Carbon	GpA shifts (TMS)	GpA (DSS)	BMRB random coil (DSS)	Secondary chemical shift
Val CO	176.10	178.10	175.62	2.48
Val CA	64.77	66.77	62.5	4.27
Val CB	29.55	31.55	32.75	-1.20
Val CG1,2	20.93	22.93	21.52	1.41
Gly CO	173.81	175.81	173.85	1.96
Gly CA	45.71	47.71	45.38	2.33
Val CO'	172.73	174.73	175.62	-0.89
Val CA'	58.26	60.26	62.5	-2.24
Val CB'	33.27	35.27	32.75	2.52
Val CG1,2'	19.48	21.48	21.52	-0.04
Gly CO'	169.04	171.04	173.85	-2.81
Gly CA'	43.83	45.83	45.38	0.45

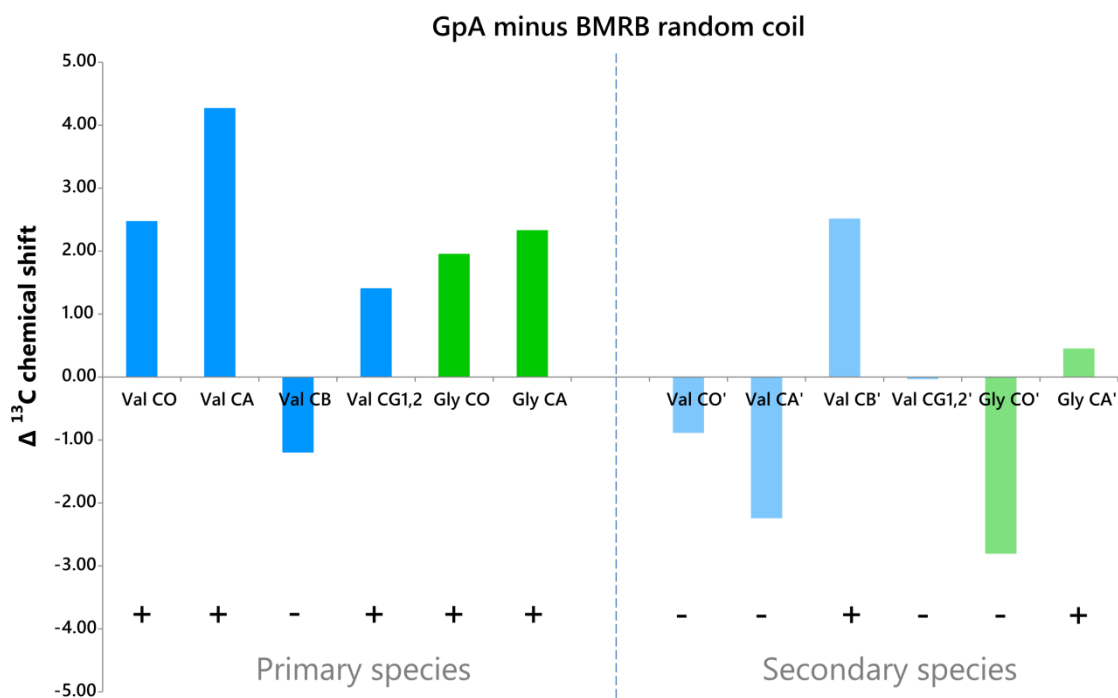


Figure 4.4 Experimentally derived secondary chemical shifts for GpA

Bar chart showing the result of experimentally observed chemical shifts from the major and minor species observed in our GpA ssNMR minus those of random coil shifts from the BMRB. Positive and negative values are indicated at the bottom of the graph. Valine chemical shifts shown in blue and glycine chemical shifts shown in green.

4.2.3 1D ¹³C ssNMR of GpA_{VG} prepared using the co-solubilisation

method

Following on from the 1D ¹³C experiments carried out on the GpA_{VG} ssNMR sample prepared using the detergent removal method (as described in **Chapter 2 section 2.7.2**), the same set of experiments were recorded for the GpA_{VG} peptide reconstituted into DMPC liposomes prepared using an alternative method based on co-solubilisation (as described in **Chapter section 2.7.3**). In agreement with results from biophysical characterisation of samples prepared using the detergent removal and co-solubilisation methods as detailed in **Chapter 3**, the 1D ¹³C CP MAS spectrum shown in **Figure 4.5** indicates that the sample prepared by co-solubilisation of peptide along with lipid and cholesterol resulted in a marked reduction of any observable secondary species within the sample. In addition to the reduction in secondary species, indicating improved sample homogeneity and a reduction in incorrectly folded β-sheet or

aggregated protein, the spectrum obtained also showed an improvement in overall signal to noise ratio (56 versus 46.9 for detergent removal method) and decreased line widths of the protein resonances observed.

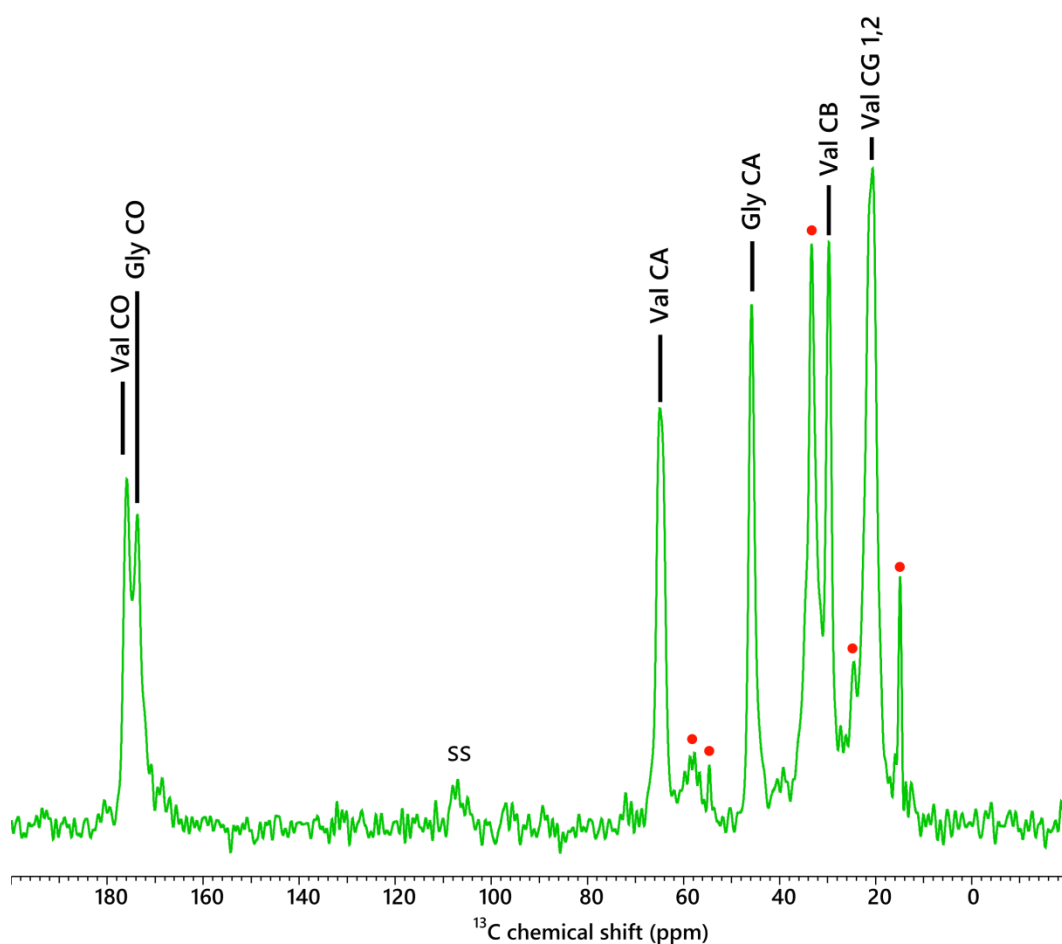


Figure 4.5 1D proton-decoupled ^{13}C CP-MAS spectrum of doubly labelled GpA_{VG} in DMPC liposomes prepared using the co-solubilisation method.

Doubly labelled GpA_{VG} sample prepared using the co-solubilisation method (as described in Chapter 2.7.3). Spectrum recorded at 500 MHz with 8.5 kHz MAS at 258 K (-15 °C) with 1024 co-added transients recorded. Resonances arising from U- ^{13}C labelled valine and glycine are labelled accordingly and those arising from natural abundance ^{13}C from DMPC lipid indicated by a red circle (•), spinning sideband indicated by ss.

4.2.4 2D ^{13}C - ^{13}C DARR ssNMR

Following the acquisition and assignment of the 1D ^{13}C GpA spectrum, two-dimensional (2D) MAS NMR experiments were performed on the sample in order to observe any dipolar coupling between the two labelled amino acids of the doubly labelled GpA_{VG} peptide. 1D protein NMR spectra were not sufficient to resolve all resonances due to overlap of signals, and as such NMR experiments were extended to a second dimension to increase resolution and simplify interpretation of the spectra.

In order to see if couplings between isotopically labelled valine and glycine residues predicted to lie at the GpA homodimer interface could be observed, 2D ^{13}C - ^{13}C Dipolar Assisted Rotational Resonance (DARR) (Takegoshi, Nakamura et al. 2001) experiments were recorded using the hydrated lipid-reconstituted GpA_{VG} peptide sample. A short mixing time (20 ms) 2D ^{13}C - ^{13}C DARR spectrum of doubly labelled GpA_{VG} peptide reconstituted into hydrated DMPC liposomes using the co-solubilisation method is shown in **Figure 4.6**. From the 2D spectrum obtained a number of cross peaks were observed off the diagonal (with strong diagonal signals arising from labelled amino acids in GpA and also from background natural abundance ^{13}C from DMPC lipid). Due to the use of selective labelling, the spectral crowding that can occur when using uniformly labelled protein samples was avoided, leading to well-resolved peaks in both the aliphatic and carbonyl regions. Signals from all labelled carbon atoms can be accounted for, with only cross peaks arising from valine C γ 1/2 atoms being unresolvable as individual separate resonances due to their similar chemical shifts and the broadness of the peaks. From the short (20 ms) mixing time DARR spectrum, a number of cross peaks are observed that arise from intra-residue correlations between the U- ^{13}C carbon atoms within each labelled valine and glycine in the sample. These intra-residue cross peaks can be seen at shorter mixing times due to the short one and two bond ($\sim 1.5/3.0\text{\AA}$) distances between nuclei. As shown in **Figure 4.6**, intra-residue cross peaks could be observed between valine C α – C β , valine C α - C γ 1/2, valine C α - CO, valine C β -C γ 1/2, valine C β - CO and valine C γ 1/2 – CO. For glycine, the sole C α – CO intra-residue cross peak was also observed.

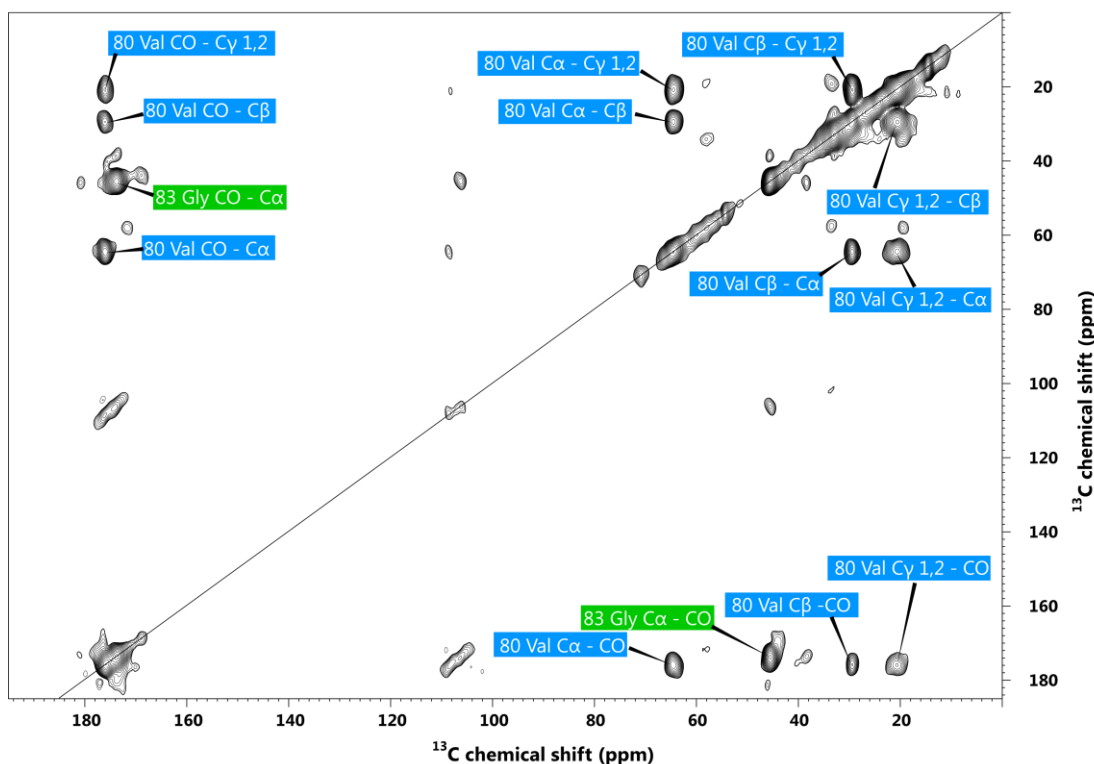


Figure 4.6 20 ms 2D ^{13}C - ^{13}C DARR spectrum of doubly labelled GpA_{VG} in DMPC liposomes.

2D ^{13}C - ^{13}C DARR correlation spectrum of doubly labelled (GpA_{VG}) GpA, acquired over 48 hours with a mixing time of 20 ms. Spectrum was recorded at 500 MHz with 8.5 kHz MAS at 258 K (-15 °C) with 112 scans. At short mixing times only *intra-residue* cross peaks are observed. Cross peaks are labelled according to the amino acid spin system, with cross peaks arising from valine 80 labelled in blue and from glycine 83 in green.

As the DARR mixing time was increased from 20 ms to 400 ms (**Figure 4.7**), the number of cross peaks observed off the diagonal increased, since at longer mixing times magnetisation travels further allowing longer range through space couplings to be observed. As such, in addition to the intra-residue cross peaks, we also observed inter-residue cross peaks from those labelled carbon atoms that are close together in space at the dimer interface. From the chemical shifts, these additional inter-residue cross peaks were attributed to coupling between Val-Cy1/2 to Gly-C α and Val-C β to Gly-C α . **Figure 4.8** shows an enlargement of the aliphatic region of the 2D ^{13}C - ^{13}C DARR spectrum where these additional long range cross peaks can be observed.

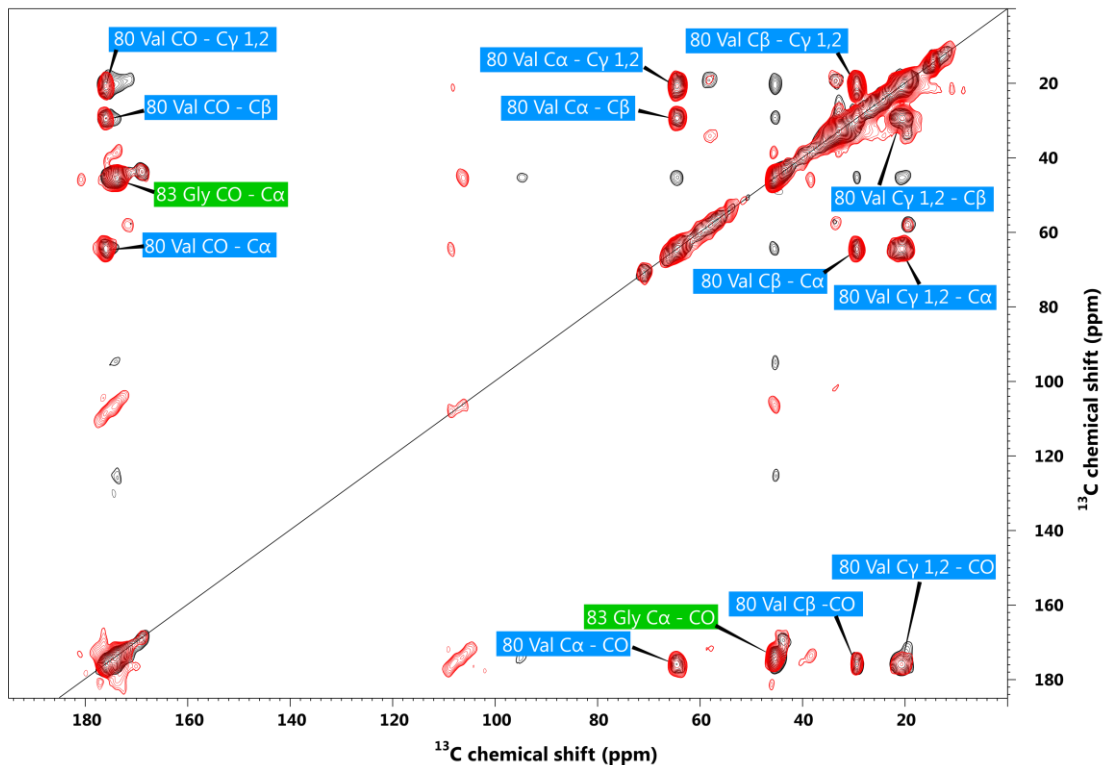


Figure 4.7 Overlay of 20 & 400 ms 2D ^{13}C - ^{13}C DARR spectra of doubly labelled GpA_{VG} in DMPC liposomes.

2D ^{13}C - ^{13}C DARR correlation spectrum of doubly labelled (GpA_{VG}) GpA, acquired over 48 hours with a mixing time of 20 ms (red) and 400 ms (black). Spectra were recorded at 500 MHz with 8.5 kHz MAS at 258 K (-15 °C) with 112 scans. At short mixing times (20 ms) only *intra-residue* cross peaks are observed. At longer mixing times (400 ms), *inter-helical* cross peaks are observed. Cross peaks are labelled according to the amino acid spin system, with cross peaks arising from valine 80 marked in blue and from glycine 83 in green.

The appearance of these additional cross peaks at long mixing times would suggest that there was an interaction between labelled valine and glycine amino acids at the GpA homodimer interface, resulting in coupling between the two residues. This result fits in well with previously published studies of the GpA transmembrane protein homodimer ((MacKenzie, Prestegard et al. 1997; Smith, Song et al. 2001; Smith, Eilers et al. 2002)) which have shown that valine 80 and glycine 83 pack closely together at the dimer interface through van der Waals interactions (MacKenzie, Prestegard et al. 1997). The GpA homodimer interface has also previously been studied in pure DMPC liposomes using ssNMR REDOR experiments in order to study the dimer interface using isolated spin pairs ((Smith, Jonas et al. 1994; Smith, Hamilton et al. 1994; Smith, Song et

al. 2001)). This work also suggested that these two residues are in close contact within the GpA homodimer.

The presence of these inter-residue cross peaks between isotopically labelled valine and glycine, taken together with the CD and 1D ^{13}C NMR results, would suggest that reconstitution of the GpA_{VG} peptide into DMPC/cholesterol liposomes using our co-solubilisation method successfully produced correctly folded alpha-helical protein that forms homodimers within the hydrated lipid bilayers. The addition of cholesterol to our lipid vesicles is a step forward towards making the lipid environment in which the GpA peptide is embedded more biologically relevant than single lipid systems that have been previously reported.

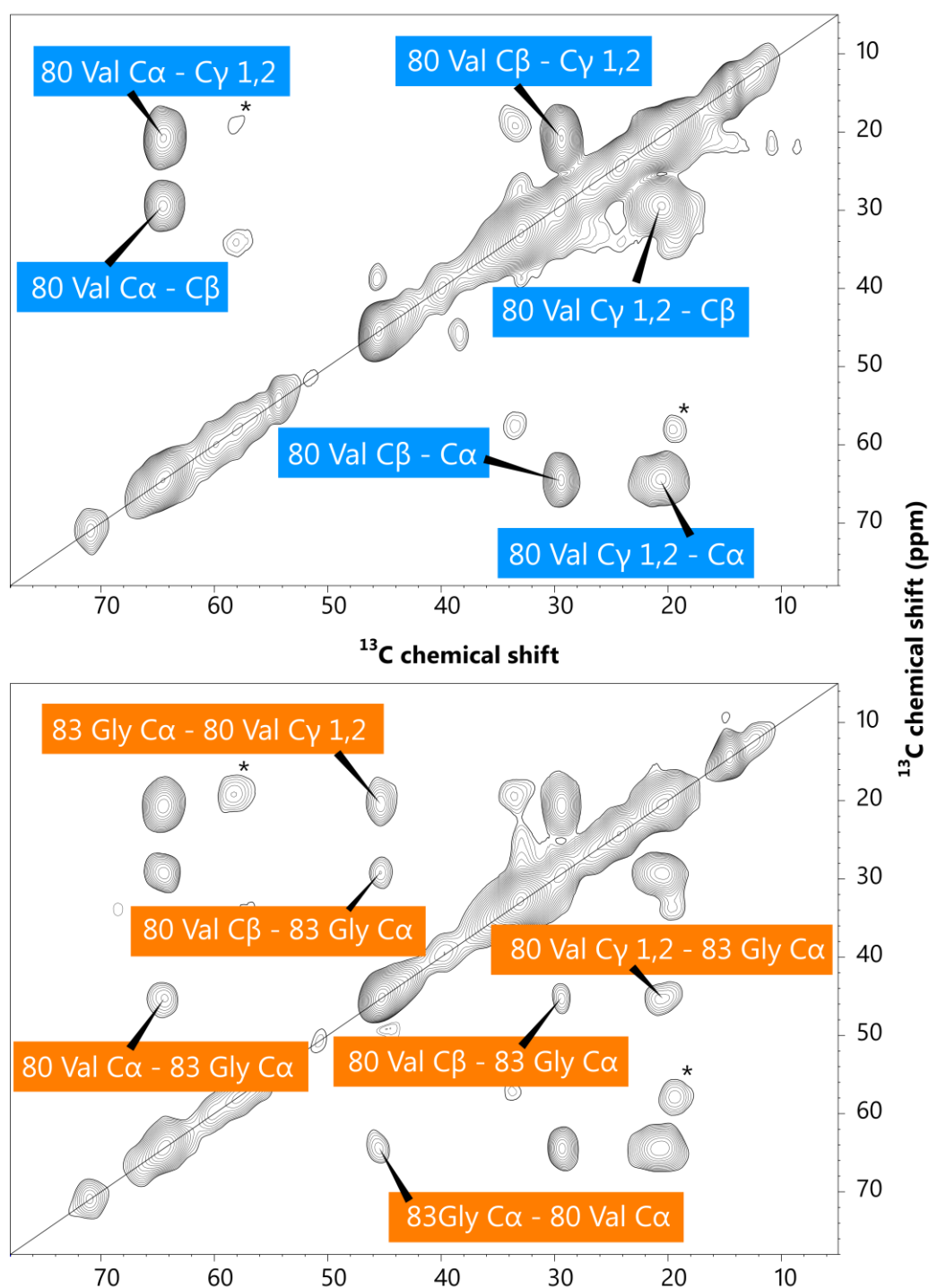


Figure 4.8 Aliphatic regions of 2D ^{13}C - ^{13}C DARR spectra of doubly labelled GpA_{VG} at short and long mixing times.

2D ^{13}C - ^{13}C correlation spectrum of GpA_{VG} labelled at both valine 80 and glycine 83, using mixing times of 20 ms (top panel) and 400 ms (bottom panel). In the short mixing time spectrum, short range *intra*-residue cross peaks assigned are labelled in the top panel, whilst long range *inter*-helical cross peaks, only arising at long mixing times, are labelled in the bottom panel. Secondary species attributed to incorrectly folded protein indicated by *.

4.2.5 Relating cross peaks observed to existing GpA structure

To confirm that the inter-residue cross peaks observed in long mixing time 2D ^{13}C - ^{13}C DARR spectra between valine and glycine carbon atoms were a result of inter-helical dipolar couplings between individual peptide chains, as opposed to intra-helical couplings within the same peptide chain, we compared the average distances between interacting atoms using the previously published data for the solution NMR structure of the GpA dimer in DPC detergent micelles (MacKenzie, Prestegard et al. 1997). Average carbon-carbon inter-atomic distances, along with standard deviations of the twenty best structures taken from the GpA solution structure (pdb file: 1AFO) are shown in **Table 4.5**. Those inter-atomic distances shorter than 5.5 Å are highlighted in bold type, as distances above this are unlikely to be detected due to the limitations of the DARR NMR experiment (Crocker, Patel et al. 2004).

Table 4.5 Average distances between valine 80 and glycine 83 carbon atoms in the published GpA homodimer structure

List of inter-atomic distances in Angstroms, obtained from the solution structure of the DPC detergent micelle-solubilised transmembrane domains of GpA obtained from solution-state NMR (pdb file: 1AFO (MacKenzie et al 1997)). 20 models within the solution NMR structure deposited were averaged to give the mean distance. Those distances within a single helix (intra-helical) are shown on the left and those between helices (inter-helical) on the right. Highlighted in bold type are those distances that are shorter than 5.5Å.

Intra-helix Distances						Inter-helix Distances					
Interacting atoms				Mean Distance (Å)	Standard Deviation	Interacting atoms				Mean Distance (Å)	Standard Deviation
V	C α	G	C α	5.15	0.06	V	Cα	G	Cα	4.56	0.28
V	C α	G	CO	5.82	0.06	V	C α	G	CO	5.76	0.22
V	C β	G	C α	6.50	0.07	V	Cβ	G	Cα	4.62	0.26
V	C β	G	CO	6.99	0.06	V	C β	G	CO	6.03	0.23
V	C γ 1	G	C α	6.54	0.07	V	Cγ1	G	Cα	3.49	0.27
V	C γ 1	G	CO	6.75	0.08	V	Cγ1	G	CO	4.97	0.25
V	C γ 2	G	C α	7.36	0.08	V	Cγ2	G	Cα	5.25	0.22
V	C γ 2	G	CO	8.03	0.08	V	C γ 2	G	CO	6.70	0.21
V	CO	G	Cα	4.54	0.05	V	CO	G	Cα	4.56	0.30
V	CO	G	CO	4.94	0.04	V	CO	G	CO	5.57	0.24

The inter-residue correlations observed in the 400 ms DARR spectrum of GpA_{VG} (Val80 C γ 1 – Gly83 C α ; Val80 C α – Gly83 C α ; and Val80 C β – Gly83 C α) are highlighted in yellow in Table 4.5. From the inter-atomic distances calculated from the solution NMR structure, it can be seen that for all three of the correlations observed, that the inter-helical distance is significantly shorter than the intra-helical distance. For example, for the Val80 C γ 1 – Gly83 C α cross peak, the inter-helical distance in the detergent solubilised dimer is 3.49 Å whereas the intra-helical distance within a single GpA helix is 6.54 Å. This is illustrated in **Figure 4.9**. The same trend is observed for the remaining inter-residue cross peaks. It should be noted that, from the table of distance calculated, it was also expected inter-helical cross peaks between Val80 CO – Gly83 C α and Val80 C γ 1 – Gly83 C α should be observable, but due to the broadness of the cross peaks arising from intra-residue couplings the resolution was not there to be able to differentiate these cross peaks. Another reason for the inability to observe these longer range cross peaks could be due to the effects of dipolar truncation. As the dipolar coupling between two nuclei is related to the distance between them (r^3) the transfer of spin magnetisation between more weakly coupled nuclei is affected by that of more strongly coupled nuclei of the same species. This effect results in the truncation of magnetisation transfer and therefore weaker inter-helical cross peaks are much less likely to be detected when using U-¹³C labelled samples (or amino acids). From the distances calculated and the cross peaks identified in the DARR spectra, we can say with reasonable confidence that the cross peaks assigned likely arise from the shorter inter-helical distances and not the longer (> 5 Å) intra-helical distances between the same atoms, which are unlikely to be observed using DARR ssNMR (Crocker, Patel et al. 2004).

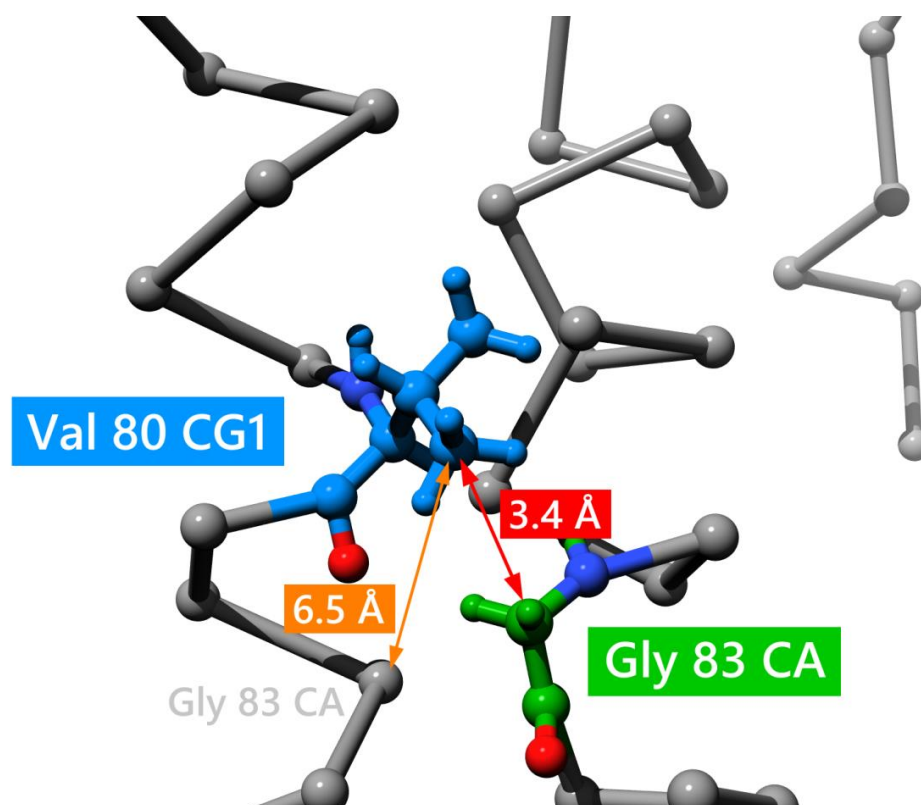


Figure 4.9 Molecular model of human GpA homodimer

Molecular model of the GpA homodimer produced using coordinates from the solution NMR protein data bank file 1AFO (MacKenzie, Prestegard et al. 1997). Protein backbone and labelled amino acid residues are shown in ball and stick form. Labelled valine chain A (blue) and glycine chain B (green) residues have been coloured. The intra-helical distance of 6.5 Å and inter-helical distance of 3.4 Å have both been indicated by orange and red arrows, respectively.

4.3 Alternative peptide labelling scheme.

Although the inter-helical cross peaks observed in the 400 ms 2D ^{13}C - ^{13}C DARR spectrum were between carbon atoms that lie $< 5 \text{ \AA}$ from one another at the homodimer interface in the detergent-solubilised GpA solution structure, the possibility still existed that the observed peaks could be due to intra-helical interactions in DMPC/cholesterol bilayers. Different membrane protein structures have been observed depending on the chemical environment (i.e. detergent vs. lipid) (Poget and Girvin 2007). Furthermore, since the aim of the work was ultimately to apply the sample preparation and ssNMR experimental procedures to proteins for which no structural data exists, such as the Bovine Papillomavirus E5 protein (BPV E5), any approach which relied heavily on existing structural/distance information in order to draw conclusions

about inter- versus intra-helical coupling would be flawed. As such, an alternative peptide labelling scheme was proposed which would give rise to unambiguous assignments.

This alternative labelled scheme (referred to from now on as singly labelled) differed from our original labelling scheme (doubly labelled) in that only one amino acid was labelled within a synthetic peptide (outlined in **Figure 4.10**). Therefore, to investigate Val - Gly couplings as before, two peptides were required, one containing a U $^{13}\text{C}/^{15}\text{N}$ labelled Val 80 and one containing a U $^{13}\text{C}/^{15}\text{N}$ labelled Gly 83. This labelling scheme had a major advantage over the previous doubly labelled scheme in that any inter-residue cross peaks observed between the two isotopically labelled amino acids could be unambiguously attributed to inter-helical coupling. The major disadvantage of this method is that when preparing samples for ssNMR, peptides that have been labelled at different positions have to be mixed together before reconstitution into liposomes. This results in formation of the heterodimers of interest, but also results in formation of homodimers that will yield no inter-residue coupling information. If it is assumed a statistical mix of homo- and heterodimers, then when using singly labelled peptides a loss of 50% signal intensity is expected when compared to using doubly labelled peptides, where all dimers formed are heterodimers.

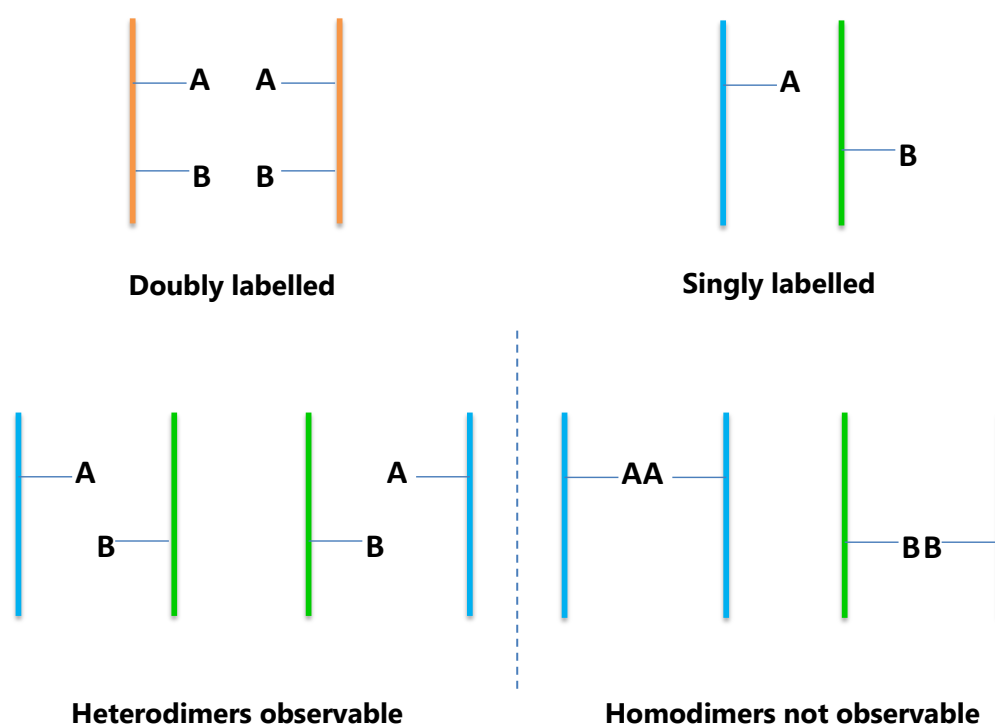


Figure 4.10 Diagrammatic form of alternative peptide labelling scheme

Peptides labelled at sites A and B on double labelled peptides (top left), where both A and B are labelled within a peptide chain. Singly labelled peptides (top right) are labelled at either position A or B. When singly labelled peptides are mixed, statistically there is a 50% chance of forming hetero dimers (bottom left) which are NMR observable and a 50% chance of forming homo dimers (bottom right) which are not observable in DARR spectra.

4.3.1 ssNMR of singly labelled GpA peptides.

Given that the advantage of using singly labelled peptides meant that any cross peaks observed between the labelled amino acid sites could be attributed unambiguously to inter-helical interactions (and thus would be applicable to future work with the BPV E5 peptide), we tested this approach using the GpA TM peptide. We had GpA peptides synthesised (as described in **Chapter section 2.3.1**) that were singly U- $^{13}\text{C}/^{15}\text{N}$ labelled at either valine 80 (GpA_V) or glycine 83 (GpA_G) positions. The sample prepared using this labelling scheme was then compared to that of the previously recorded doubly labelled sample to see if the same inter-helical cross peaks could be observed.

A 1D ^{13}C CP-MAS spectrum of a GpA sample reconstituted into DMPC/cholesterol liposomes using a mixture of singly labelled GpA_V and GpA_G peptides, as described in **Chapter section 2.7.3**, prepared using the co-solubilisation

method is shown in **Figure 4.11**. As shown in this figure, the new singly labelled sample yielded the same ^{13}C chemical shifts and the same number of signals as observed for the doubly labelled GpA_{VG} peptide, and no evidence of a misfolded secondary species. Line widths at half height of 1.2-2.7 ppm/163-350 Hz were observed for protein resonances for the singly labelled sample, which were slightly broader than those observed in the doubly labelled sample prepared by co-solubilisation. The signal to noise ratio was comparable to the doubly labelled sample, but overall signal obtained was, as expected, approximately 46% lower than that of the doubly labelled GPA_{VG} sample using the same number of co-added transients (1024).

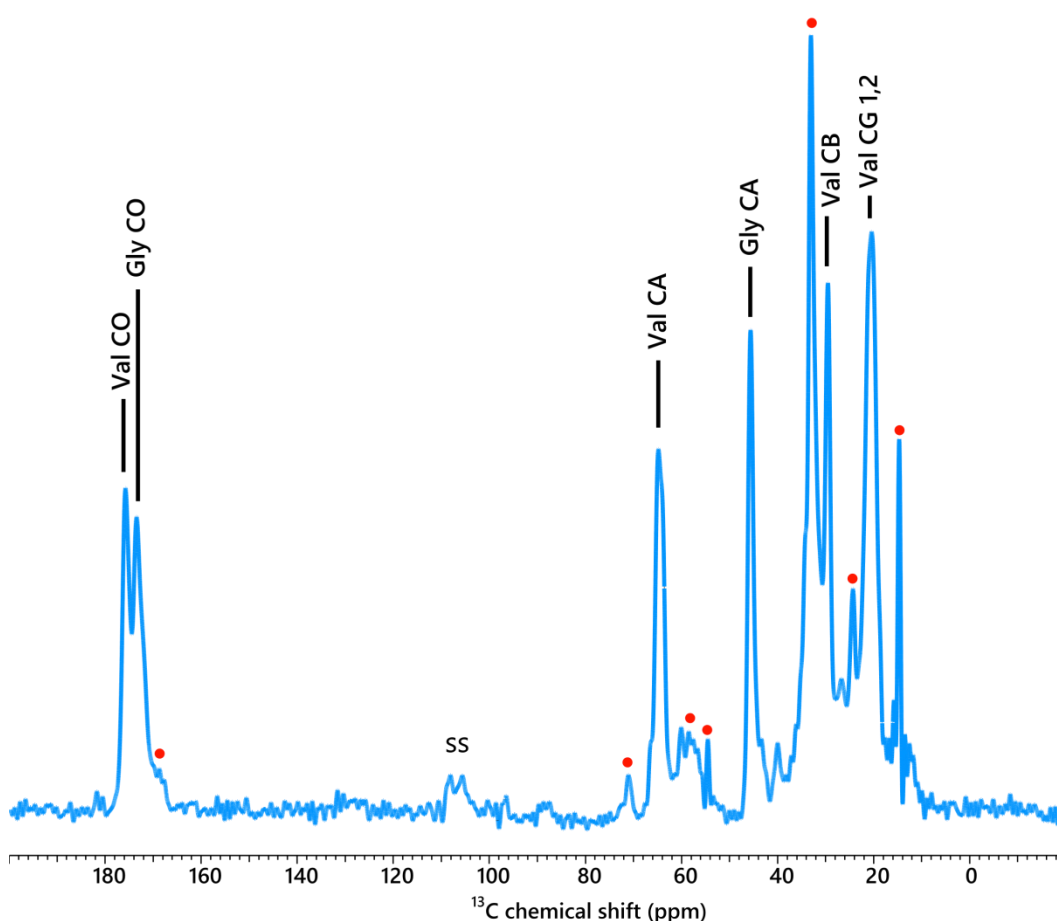


Figure 4.11 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled GpA_V + GpA_G mixture in DMPC liposomes prepared using the co-solubilisation method

Spectrum recorded at 500 MHz with 8.5 kHz MAS at 258 K ($-15\text{ }^{\circ}\text{C}$). Resonances arising from U- ^{13}C labelled valine and glycine are labelled accordingly and those arising from natural abundance ^{13}C from DMPC lipid indicated by \bullet , spinning sideband indicated by ss.

Following 1D experiments, 2D ^{13}C - ^{13}C DARR correlation spectra were again recorded with both short (20 ms) and long (400 ms) mixing times. Due to the decrease in overall signal intensity in comparison to the doubly labelled GpA_{VG} sample, when recording the 2D ^{13}C - ^{13}C DARR spectra we chose to increase the number of scans from 112 to 308 in order to ensure that if any inter-helical cross peaks were present, they would be observable. At short mixing times (data not shown) in the singly labelled GpA sample we observed the same intra-residue cross peaks that were previously present in our doubly labelled sample, albeit at much lower intensity in comparison to our doubly labelled GpA_{VG} sample.

More interestingly, in the 400 ms DARR spectrum of the singly labelled sample, the same three inter-residue cross peaks between valine and glycine were observed (overlaid in **Figure 4.12**) that were seen in the doubly labelled GpA_{VG} sample. These results allowed us to unambiguously and confidently attribute the inter-residue cross peaks to long range, through-space, inter-helical dipolar couplings between the two amino acids at the homodimer interface of GpA. These data also suggest that the inter-nuclear distances over which DARR cross peaks can be observed do not extend much above 4.6 Å for our samples, as only inter-nuclear distances predicted to be less than or equal to this value produced cross-peaks in our DARR spectra.

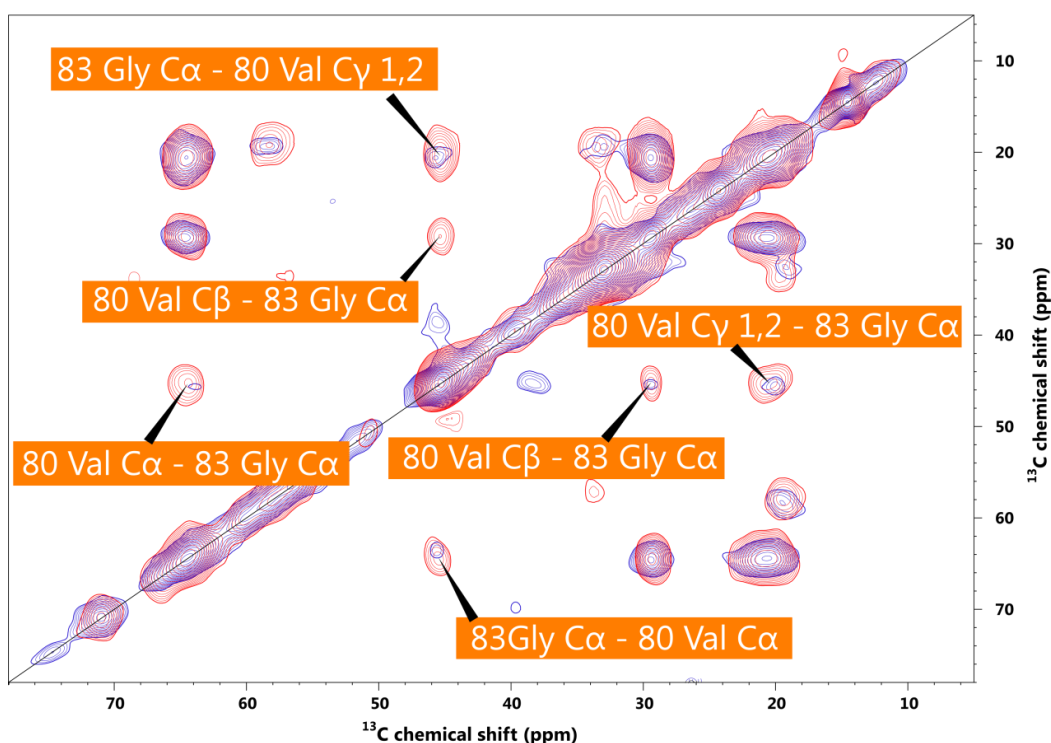


Figure 4.12 Overlaid 400 ms 2D ^{13}C - ^{13}C DARR correlation spectrum of doubly labelled GpA (GpA_{VG}) vs. singly labelled GpA_{V} + GpA_{G}

2D ^{13}C - ^{13}C correlation spectrum with a mixing time of 400 ms for doubly labelled GpA_{VG} (red spectrum) and singly labelled GpA_{V} + GpA_{G} (blue spectrum) peptides reconstituted into DMPC liposomes. Inter-helical cross peaks between labelled valine and glycine are indicated. Spectra were recorded at 8.5 kHz MAS at 258 K ($-15\text{ }^{\circ}\text{C}$) with 112 co-added transients for doubly labelled GpA_{VG} and 280 for singly labelled GpA_{V} + GpA_{G}

4.3.2 Singly labelled GpA build-up curves

In order to further investigate the inter-helical cross peaks observed between valine 80 and glycine 83, a series of 2D ^{13}C - ^{13}C DARR experiments at a range of mixing times (20, 100, 200, 400 and 600 ms) (experimental time for each experiment: 4.5 days) were recorded in order to produce build up trajectories for each cross peak observed from dipolar couplings between labelled carbon atoms in each labelled amino acid. **Figure 4.13** shows the build-up curves for intra- (top panel) and inter-residue cross peak intensities (bottom panel) for the singly labelled GpA sample. Inter-helical cross peak

intensities were normalised to the glycine C α diagonal cross peak intensity for each mixing time, to account for the slight variations between each experiment.

As can be seen from the graphs of intensities vs. mixing time, the cross peak arising from the shortest inter-helical distance (3.49 Å) between Val80 C γ 1 and Gly 83 C α , has the highest intensity at each mixing time and builds up the fastest of all inter-residue cross peaks observed. The next shortest inter-helical distance between Val80 C α and Gly83 C α (4.56 Å) results in a cross peak that has a lower intensity than the Val80 C γ 1 - Gly83 C α cross peak, but has a slightly higher intensity than that of the Val80 C β - Gly83 C α , which has a slightly longer inter helical distance of 4.62 Å. In this way, the DARR-derived build-up curves for inter-residue cross peaks show very good correlation with inter-atomic distance as predicted from the GpA solution structure for inter-helical interactions. For cross peaks arising due to the shorter intra-residue dipolar couplings, maximum cross peak intensity was observed at 100 ms, followed by a steady decrease in intensity for mixing times between 100-600 ms (**Figure 4.13**, top panel) due to decay of magnetisation through transfer into the surrounding system further removed from the protein spin system.

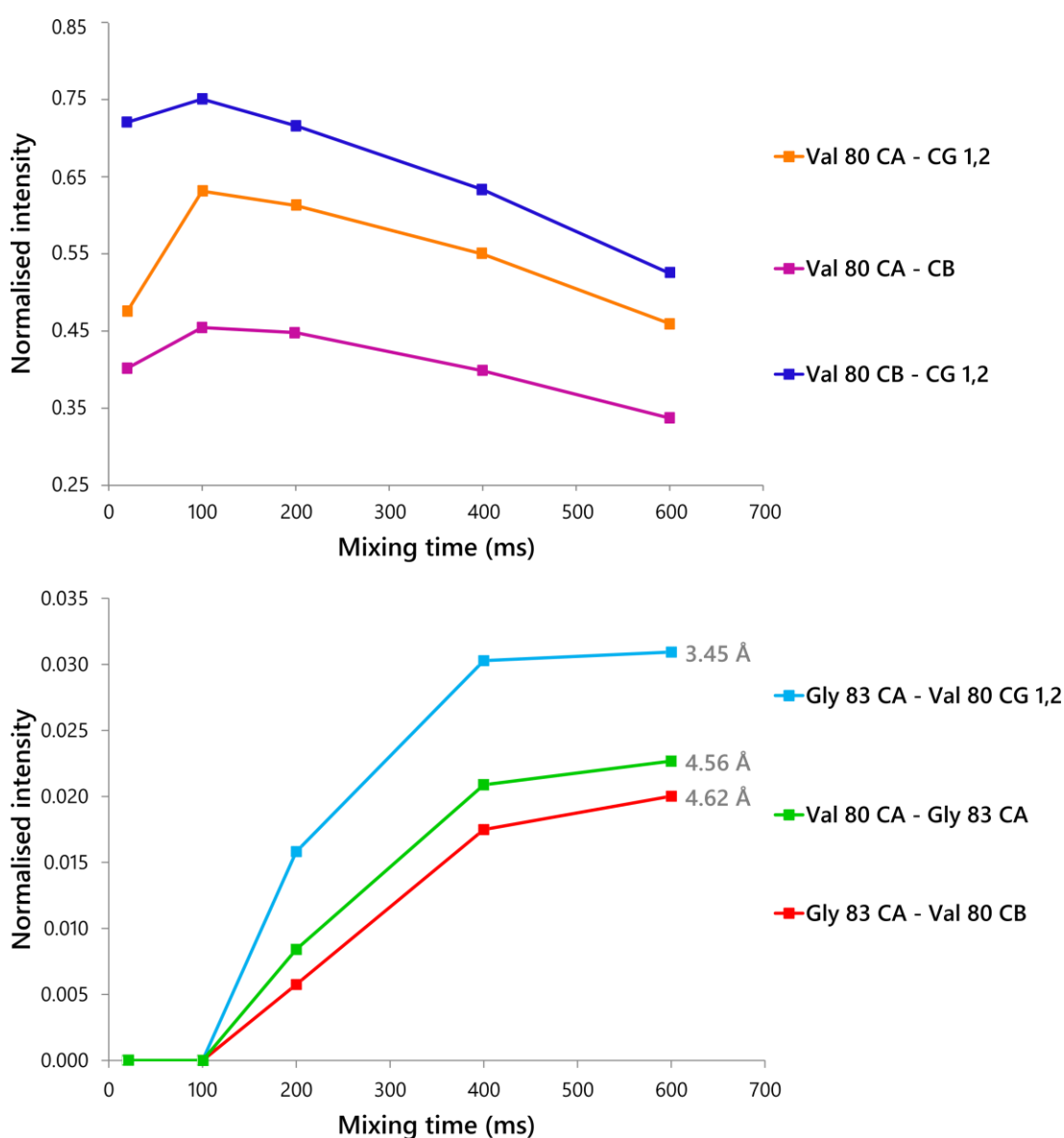


Figure 4.13 Build-up curves for intra-residue and inter-residue GpA cross peaks

Build up curves from 2D ^{13}C - ^{13}C DARR experiments for intra-residue (top panel) and inter-residue cross peaks (bottom panel) at increasing mixing time from 20, 100, 200, 400 to 600 ms. Spectra were recorded with 308 scans, taking 4 and a half days each. Cross peak intensity was normalised to the Gly CA cross peak for each experiment. In the bottom panel the average inter-helical distance between each labelled carbon atom, as obtained from the GpA solution NMR structure is also given alongside each curve.

4.4 ^{15}N GpA ssNMR

In addition to ^{13}C isotopically labelled carbon atoms within selected amino acids in the GpA peptides, which allowed for ^{13}C - ^{13}C homonuclear 2D measurements, the same valine and glycine amino acids contained isotopic ^{15}N labels for amide nitrogen's. Using the ^{15}N isotopic labels within singly labelled GpA peptides, the aim was to confirm the formation of GpA TM homodimers by observing ^{15}N - ^{13}C any dipolar coupling between valine and glycine backbone nitrogen's and carbons in the side chains at the GpA dimer interface.

4.4.1 1D ^{15}N ssNMR of singly labelled GpA sample

One-dimensional (1D) ^{15}N CP MAS NMR spectra were initially acquired for the singly labelled GpA sample reconstituted into DMPC/cholesterol lipid bilayers in order to gain ^{15}N chemical shift information for the labelled valine and glycine amino acids. As with the ^{13}C experiments, all ^{15}N experiments were carried out at low temperature (-15°C) where DMPC lipid is in the gel phase.

The two expected ^{15}N resonances arising from the isotopically labelled valine and glycine residues within the GpA peptide are shown in the 1D ^{15}N CP-MAS NMR spectrum (**Figure 4.14**). There was no observable signal arising due to natural abundance ^{15}N from the choline head group of DMPC phospholipid. The resonances arising from valine and glycine (given in **Table 4.6**) agreed well with the published average chemical shift values for these amino acids as found at the BMRB. Resonances arising from the labelled amino acids were well resolved with distinct chemical shifts (120.70 ppm for valine NH and 104.48 ppm for glycine NH). Line widths for each resonance were relatively broad with peak widths at half height of 3.01 ppm/152.93Hz and 3.1 ppm/157.47 Hz for valine and glycine NH, respectively, after 4096 scans. Due to the low gyromagnetic ratio of ^{15}N nuclei, the signal to noise ratio was poor (4.6 for valine and 4.4 for glycine), requiring a much larger number of scans per experiment in order to observe any signal for labelled nitrogen within the GpA sample.

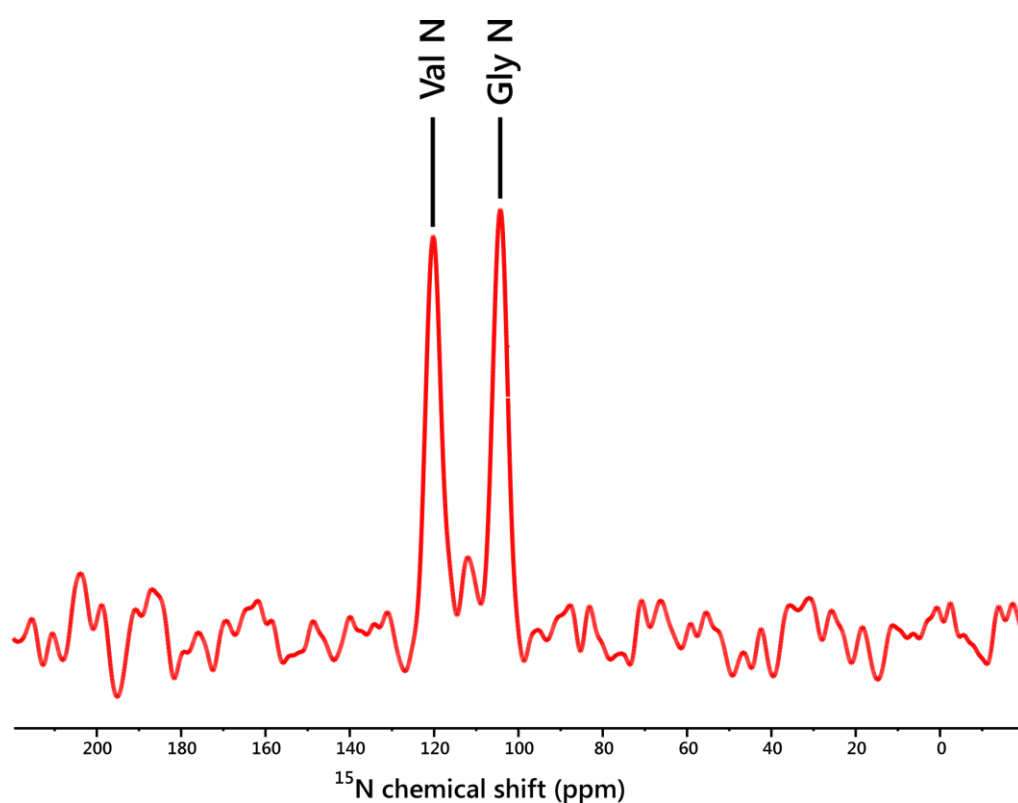


Figure 4.14 1D proton-decoupled ^{15}N CP-MAS spectrum of singly labelled GpA in DMPC liposomes prepared using the co-solubilisation method

The spectrum was recorded at 500 MHz with 12.5 kHz MAS at 258 K (-15 °C). A CP contact time of 1000 μs was used, a 2.5 second recycle delay with 100 kHz proton decoupling during acquisition and 4096 co-added transients recorded. Resonances arising from ^{15}N labelled valine and glycine are labelled accordingly. -1 Hz GM line broadening was applied before Fourier transformation

Table 4.6 ^{15}N chemical shift data for labelled GpA Val 80 and Gly 83

Val 80	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
N	120.70	119.26	1.44

Gly 83	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
N	104.48	107.82	-3.34

4.4.2 2D ^{15}N - ^{13}C ssNMR of singly labelled GpA sample

Triple resonance ^1H - ^{13}C - ^{15}N experiments were setup with the help of Fredrick Romer (Department of Physics, Warwick University)

In order to measure heteronuclear ^{15}N - ^{13}C interactions, a number of ssNMR methods have been developed including the more popular rotational echo double resonance (REDOR) (Gullion and Schaefer 1989) and transferred echo double resonance (TEDOR) (Hing, Vega et al. 1992) for recoupling dipolar interactions between isolated hetero-nuclear spin $\frac{1}{2}$ pairs (e.g. ^{13}C , ^{15}N) in MAS experiments. The use of such experiments in distance measurements has been well established (Hing, Vega et al. 1992; Fyfe and Lewis 2000). The challenge when directly applying such techniques arises when using U- ^{13}C , ^{15}N labelled samples, due to ^{13}C and ^{15}N labelled nuclei experiencing multiple through bond (J -coupling) and through space (dipolar couplings) from neighbouring spins. These stronger one/two bond couplings can result in the complication of magnetisation transfer pathways which can lead to degradation of spectral sensitivity and resolution, as well as interfering with accurate determination of weaker couplings (Fyfe and Lewis 2000; van Rossum, de Groot et al. 2000)).

A number of experiments based upon REDOR and TEDOR have been developed in order to get around the issue of multiple spin-systems when using U- ^{13}C , ^{15}N labelled NMR samples. These include frequency selective REDOR (FSR) (Jaroniec, Tounge et al. 2001) and three-dimensional (3D) z -filtered TEDOR (Fyfe, Mueller et al. 1992; Jaroniec, Filip et al. 2002) The z -filtered TEDOR experiment is a modification of the TEDOR experiment and addresses a number of issues associated with using U- ^{13}C , ^{15}N labelled samples and the problems arising from homonuclear J -couplings that interfere with measurement of weak ^{13}C - ^{15}N dipolar couplings. The use of a z -filter period to suppress anti-phase and multiple quantum coherences, results in 2D ^{15}N - ^{13}C correlation spectra with pure absorption mode peaks at all TEDOR mixing times and provides information about carbon-nitrogen distances.

2D ^{15}N - ^{13}C z -filtered TEDOR correlation spectra at short (6 ms) and long (10 ms) mixing times are shown in **Figure 4.15**. From the 2D correlation spectra obtained, number of cross peaks were observed arising from intra-residue dipolar coupling between ^{15}N and ^{13}C labelled atoms within valine and glycine. For the spectrum acquired with a 6 ms mixing time, the shorter distance coupling between Val80 C α -N (1.49 Å)

and Val80 C β – N (2.46 Å) give rise to relatively intense cross peaks. These cross peaks are no longer observable in the 10 ms mixing time experiment. The Val80 CO – N (2.46 Å) cross peak has a similar intensity at both mixing times, whilst the Val80 C γ 1/2 (3.81/2.90 Å) cross peak intensity increases at the longer mixing time (magnetisation is allowed to travel further). Both the Gly83 CO – N (1.49 Å) and the Gly83 C α – N (2.46 Å) cross peaks are observable at both short and long mixing times, although at the longer mixing time, the Gly83 C α – N cross peak decreases in intensity. This decrease/increase in signal intensities can be seen more clearly in extracted rows from each of the two mixing time experiments, with each ^{15}N row showing correlations to ^{13}C (**Figures 4.16 and 4.17**).

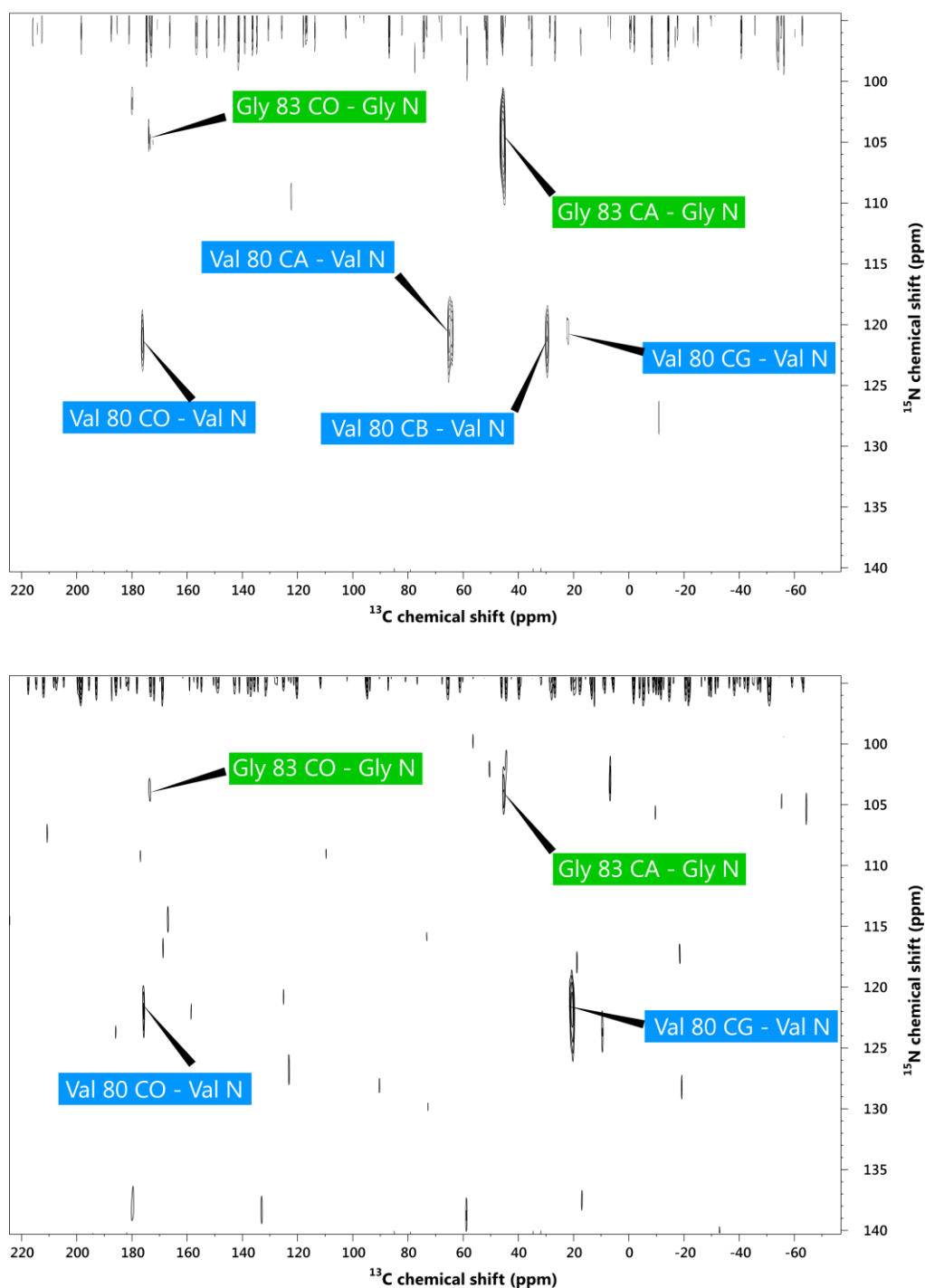


Figure 4.15 2D ^{15}N - ^{13}C TEDOR spectra of singly labelled GpA_V + GpA_G peptides in DMPC/cholesterol liposomes

2D ^{15}N - ^{13}C TEDOR correlation spectra of singly labelled GpA, acquired over 36 hours with a TEDOR mixing time of 6 ms (top panel) and 10 ms (bottom panel). Spectra were recorded at 500 MHz with 10 kHz MAS at 258 K ($-15\text{ }^{\circ}\text{C}$) with 3048 scans. Intra residue cross peaks are labelled according to amino acid spin system, with cross peaks arising from valine 80 labelled in blue and from glycine 83 in green.

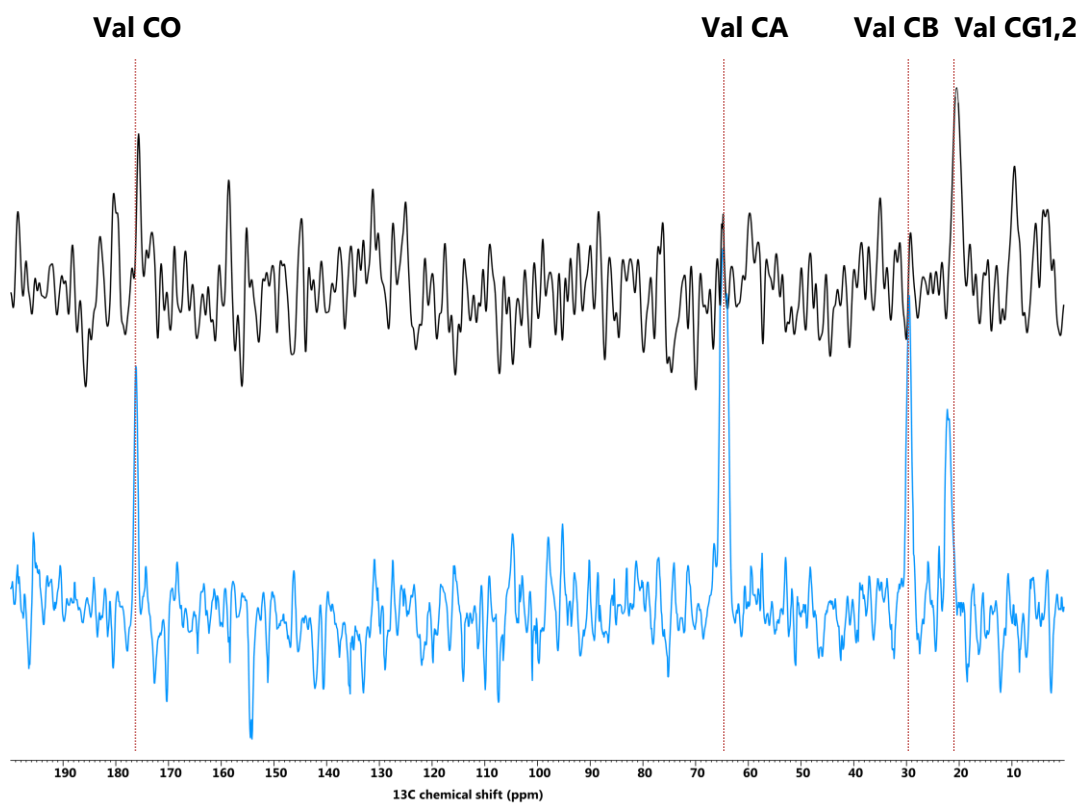


Figure 4.16 1D extracted rows from 2D ^{15}N - ^{13}C TEDOR spectra of valine resonances in singly labelled GpA_V + GpA_G peptides

1D row extracted from 2D ^{15}N - ^{13}C TEDOR correlation spectrum of singly labelled GpA, with a TEDOR mixing time of 10 ms (black spectrum) and 6 ms (blue spectrum). Spectra were recorded at 500 MHz with 10 kHz MAS at 258 K (-15 °C) with 3048 scans. Resonances from ^{13}C labelled valine carbon atoms with correlation to valine ^{15}N are labelled.

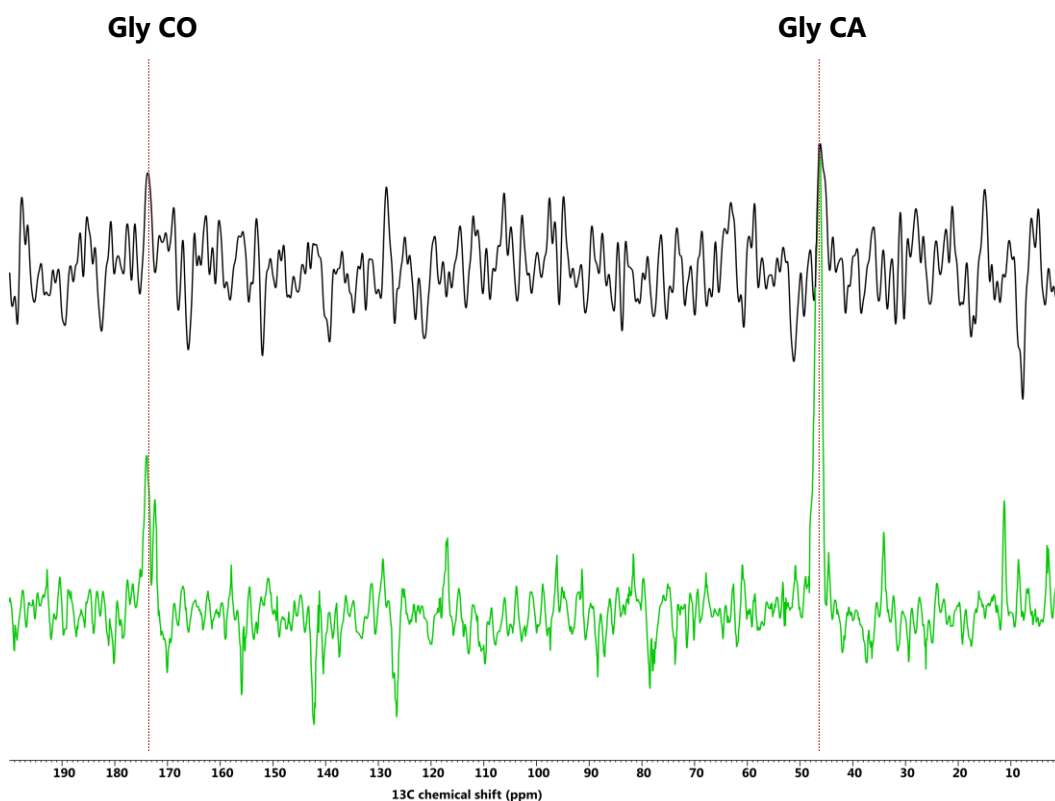


Figure 4.17 1D extracted row from 2D ^{15}N - ^{13}C TEDOR spectra of glycine resonances in singly labelled GpA_V + GpA_G peptides

1D row extracted from 2D ^{15}N - ^{13}C TEDOR correlation spectrum of singly labelled GpA, with a TEDOR mixing time of 10 ms (black spectrum) and 6 ms (green spectrum). Spectra were recorded at 500 MHz with 10 kHz MAS at 258 K (-15 °C) with 3048 co-added transients. Resonances from ^{13}C labelled glycine carbon atoms with correlation to valine ^{15}N are labelled.

Figures 4.16 and 4.17 show extracted rows from the 2D ^{15}N - ^{13}C z -filtered TEDOR experiments for valine 80 and glycine 83 respectively, at short (6 ms) and long (10 ms) TEDOR mixing times. From the 1D extracted rows, it can clearly be seen which carbon resonances decrease with increasing mixing time and which increase as magnetisation is allowed to travel further at longer mixing times (allowing for longer range couplings to be observed). Line widths for each of the carbon resonances were obtained from the extracted rows and are given in **Table 4.7**.

Table 4.7 ^{13}C line widths for GpA Val 80 and Gly 83 extracted from rows of 2D ^{15}N - ^{13}C z-filtered TEDOR experiments

Carbon atom	Line widths (ppm)	
	6 ms	10 ms
Val 80 CO	0.82	0.82
Val 80 CA	1.80	--
Val 80 CB	1.10	--
Val 80 CG 1,2	1.39	1.82
Gly 83 CO	1.37	1.44
Gly 83 CA	1.56	1.65

In both the 2D ^{15}N - ^{13}C z-filtered TEDOR spectra, and extracted 1D rows, it is clear that no long range, inter-residue (i.e. inter-helical) cross peaks between labelled valine and glycine were observed at either mixing time. The shortest inter-helical ^{15}N - ^{13}C distance predicted from the GpA solution structure is between Val80 Cy1 – Gly 83 C α , with a distance of 3.51 Å in the solution NMR GpA structure. The fact that no inter-residue ^{15}N - ^{13}C coupling was observed is likely due to the poor sensitivity of ^{15}N labelled nuclei and the performance of the spectrometer when running in triple resonance mode, which resulted in a marked decrease in spectral sensitivity when switching from double to triple resonance mode. This is not surprising when the signal to noise ratio of the TEDOR spectra is compared to that obtained in the ^{13}C - ^{13}C DARR spectra. The average signal to noise value for rows extracted from 2D ^{13}C - ^{13}C DARR experiments of 33.1 for resonances arising from intra residue couplings and 8.48 for cross peaks arising from inter helical couplings in our original doubly labelled GpA sample. This value is reduced to 13.4 for intra residue cross peaks and 1.68 for inter helical cross peaks for our singly labelled GpA peptides which required twice as many scans per DARR experiment. These values when compared to the much lower signal to noise values of 5.21 for intra residue couplings for ^{15}N - ^{13}C experiments requiring almost ten times as many scans per 2D experiment. As such we were unable to detect the presence of GpA homodimer formation in our hydrated lipid bilayers using ^{15}N - ^{13}C TEDOR measurements. With further work to improve signal to noise ratios in TEDOR measurements, it is hoped that

the same inter-helical interactions between labelled valine and glycine at the GpA dimer interface could be confirmed using the ^{15}N labels contained within our peptides.

4.5 Summary

In this chapter, ssNMR data for the homodimeric GpA peptide reconstituted into hydrated DMPC lipid bilayers containing 5% cholesterol was presented. The addition of cholesterol made our lipid bilayers more biologically relevant in comparison to a single lipid system, as the native environment of the GpA peptide can contain as much as 60% cholesterol (van Meer, Voelker et al. 2008). We showed that GpA peptide could be reconstituted into DMPC liposomes containing cholesterol successfully using a co-solubilisation method, as using the frequently reported detergent removal method (Rigaud, Levy et al. 1998) for reconstitution of peptide resulted in the formation of a minor secondary species giving rise to a second set of chemical shifts. Using secondary shift analysis, the minor secondary species observed in this sample was predicted to be due to peptide incorrectly folded into a β -sheet secondary structure, whilst the major species was correctly folded α -helix. The co-solubilisation method described here resulted in a marked reduction of this secondary species with the majority of peptide in a correctly folded α -helical state.

Following 1D ^{13}C ssNMR analyses of the GpA peptide in DMPC/cholesterol liposomes, 2D ^{13}C - ^{13}C DARR spectra were acquired in order to observe any inter-residue, inter-helical interactions between the two labelled amino acids at the GpA homodimer interface. At short (20 ms) mixing times, we observed short-range 1- and 2-bond coupling interactions between isotopically labelled carbon atoms within the same amino acid. At longer mixing times, inter-residue couplings between the labelled Val80 and Gly83 amino acids were observed. Using the published solution structure of the GpA TM homodimer in detergent micelles as a model, these long range couplings were attributed to inter-helical interactions at the homodimer interface rather than intra-helical couplings. The possibility of measuring intra-helical interactions at long mixing times still existed due to through bond spin-diffusion within the helix in the double labelled sample, and as such an alternative labelling scheme was sought.

Switching from labelling both amino acids (valine and glycine) within a single peptide chain to a scheme in which only one or the other amino acid was labelled (i.e.

singly labelled), the DARR experiments were repeated to unambiguously assign the inter-residue cross peaks. Using singly labelled peptides ruled out any possibility of measuring intra-helical interactions at longer mixing times and would be a benefit when applying similar techniques to proteins with no pre-existing structure (i.e. BPV E5). In the 2D ^{13}C - ^{13}C DARR experiments using singly labelled peptides, the same set of inter-residue cross peaks were observed arising from long-range dipolar coupling between valine and glycine. These cross peaks had a lower intensity, thus requiring longer experimental times to be observed, due to loss of signal from formation of homodimers of similarly labelled peptides. From these results, the inter-residue cross peaks observed were assigned with confidence to inter-helical interactions between Gly and Val suggesting that the GpA TM domain forms helical homodimers in DMPC lipid bilayers containing 5% (w/w) cholesterol. Comparison of the DARR build-up curves to distances derived from the solution structure of GpA also indicates that the homodimers formed in DMPC/cholesterol are very similar if not identical to those formed in DPC detergent micelles and pure DMPC bilayers. DMPC bilayers containing cholesterol represent a more "native-like" and more biologically relevant membrane mimetic than detergent micelles or bilayers composed of pure DMPC. By obtaining these results indicating dimer formation and the interaction of labelled amino acids through dipolar couplings, it would suggest that the GpA homodimer is able to form stable dimers in a number of membrane mimetics; from detergent micelles (MacKenzie, Prestegard et al. 1997) to liposomes containing pure DMPC (Smith, Jonas et al. 1994) and now in liposomes containing cholesterol, that can significantly alter the properties of the lipid membrane.

Finally, preliminary ^{15}N - ^{13}C TEDOR data for the singly labelled GpA peptide sample was presented. Making use of the ^{15}N labels within the peptides, we aimed to again confirm the presence of GpA homodimers via measurement of inter-helical dipolar couplings between ^{15}N and ^{13}C labelled atoms. Whilst successful at obtaining intra-residue cross peaks, we were unable to detect inter-helical couplings (even with longer mixing times). As these spectra appeared to be dominated by intra-residue ^{15}N - ^{13}C couplings, this made the observation of weaker inter-helical couplings a more difficult task, therefore the use of frequency selective REDOR as opposed to z-filtered REDOR may have been more beneficial in order to observe specific hetero nuclear couplings between the two labelled amino acids. In addition the relatively low signal to noise ratio

in these spectra meant that a much larger number of scans was required; thereby increasing the experimental time significantly.

Having been able to successfully apply an optimised sample preparation protocol and 2D ^{13}C - ^{13}C DARR methods to analysis of the GpA TM homodimer, we then moved onto applying the same techniques to the BPV E5 protein as described in the following chapter (**Chapter 5**) in order to gain structural information about the E5 homodimer interface.

4.6 References

- Abdine, A., M. A. Verhoeven, K. H. Park, A. Ghazi, E. Guittet, C. Berrier, C. Van Heijenoort and D. E. Warschawski (2010). "Structural study of the membrane protein MscL using cell-free expression and solid-state NMR." *J Magn Reson* **204**(1): 155-159.
- Crocker, E., A. B. Patel, M. Eilers, S. Jayaraman, E. Getmanova, P. J. Reeves, M. Ziliox, H. G. Khorana, M. Sheves and S. O. Smith (2004). "Dipolar assisted rotational resonance NMR of tryptophan and tyrosine in rhodopsin." *J Biomol NMR* **29**(1): 11-20.
- Fyfe, C. A. and A. R. Lewis (2000). "Investigation of the viability of solid-state NMR distance determinations in multiple spin systems of unknown structure." *J Phys Chem B* **104**(1): 48-55.
- Fyfe, C. A., K. T. Mueller, H. Grondey and K. C. Wongmoon (1992). "Dipolar Dephasing between Quadrupolar and Spin-1/2 Nuclei - Redor and Tedor Nmr Experiments on Vpi-5." *Chem Phys Lett* **199**(1-2): 198-204.
- Gullion, T. and J. Schaefer (1989). "Detection of Weak Heteronuclear Dipolar Coupling by Rotational-Echo Double-Resonance Nuclear-Magnetic-Resonance." *Advances in Magnetic Resonance, Vol 13* **13**: 57-83.
- Higman, V. A., J. Flinders, M. Hiller, S. Jehle, S. Markovic, S. Fiedler, B. J. van Rossum and H. Oshkinat (2009). "Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively C-13-labelled proteins." *J Biomol NMR* **44**(4): 245-260.
- Hing, A. W., S. Vega and J. Schaefer (1992). "Transferred-Echo Double-Resonance Nmr." *J Magn Reson* **96**(1): 205-209.
- Jaroniec, C. P., C. Filip and R. G. Griffin (2002). "3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly C-13, N-15-labeled solids." *J Am Chem Soc* **124**(36): 10728-10742.
- Jaroniec, C. P., B. A. Tounge, J. Herzfeld and R. G. Griffin (2001). "Frequency selective heteronuclear dipolar recoupling in rotating solids: Accurate C-13-N-15 distance measurements in uniformly C-13,N-15-labeled peptides." *J Am Chem Soc* **123**(15): 3507-3519.
- Lee, C. and R. Griffin (1989). "Two-dimensional $^1\text{H}/^{13}\text{C}$ heteronuclear chemical shift correlation spectroscopy of lipid bilayers." *Biophys J* **55**(2): 355-358.
- MacKenzie, K., J. Prestegard and D. Engelman (1997). "A transmembrane helix dimer: structure and implications." *Science* **276**(5309): 131-133.

- Needham, D. and E. Evans (1988). "Structure and Mechanical-Properties of Giant Lipid (DmPC) Vesicle Bilayers from 20-Degrees-C Below to 10-Degrees-C above the Liquid-Crystal Crystalline Phase-Transition at 24-Degrees-C." Biochemistry **27**(21): 8261-8269.
- Poget, S. F. and M. E. Girvin (2007). "Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better." Biochim Biophys Acta **1768**(12): 3098-3106.
- Rigaud, J. L., D. Levy, G. Mosser and O. Lambert (1998). "Detergent removal by non-polar polystyrene beads - Applications to membrane protein reconstitution and two-dimensional crystallization." European Biophysics Journal with Biophysics Letters **27**(4): 305-319.
- Saffman, P. G. and M. Delbruck (1975). "Brownian motion in biological membranes." Proc Natl Acad Sci U S A **72**(8): 3111-3113.
- Smith, S., M. Eilers, D. Song, E. Crocker, W. Ying, M. Groesbeek, G. Metz, M. Ziliox and S. Aimoto (2002). "Implications of threonine hydrogen bonding in the glycoporin A transmembrane helix dimer." Biophys J **82**(5): 2476-2486.
- Smith, S., R. Jonas, M. Braiman and B. Bormann (1994). "Structure and orientation of the transmembrane domain of glycoporin A in lipid bilayers." Biochemistry **33**(20): 6334-6341.
- Smith, S., D. Song, S. Shekar, M. Groesbeek, M. Ziliox and S. Aimoto (2001). "Structure of the transmembrane dimer interface of glycoporin A in membrane bilayers." Biochemistry **40**(22): 6553-6558.
- Smith, S. O., J. Hamilton, A. Salmon and B. J. Bormann (1994). "Rotational resonance NMR determination of intra- and intermolecular distance constraints in dipalmitoylphosphatidylcholine bilayers." Biochemistry **33**(20): 6327-6333.
- Spera, S., M. Ikura and A. Bax (1991). "Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment." J Biomol NMR **1**(2): 155-165.
- Takegoshi, K., S. Nakamura and T. Terao (2001). "C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR." Chem Phys Lett **344**(5-6): 631-637.
- Ulrich, E. L., H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao and J. L. Markley (2008). "BioMagResBank." Nucleic Acids Res **36**(Database issue): D402-408.
- van Meer, G., D. R. Voelker and G. W. Feigenson (2008). "Membrane lipids: where they are and how they behave." Nat Rev Mol Cell Biol **9**(2): 112-124.
- van Rossum, B. J., C. P. de Groot, V. Ladizhansky, S. Vega and H. J. M. de Groot (2000). "A method for measuring heteronuclear (H-1-C-13) distances in high speed MAS NMR." J Am Chem Soc **122**(14): 3465-3472.
- Wishart, D. S. (2011). "Interpreting protein chemical shift data." Prog Nucl Magn Reson Spectrosc **58**(1-2): 62-87.

5

SOLID STATE NMR ANALYSIS OF BPV E5

5.1 Introduction

The Bovine Papillomavirus E5 (BPV E5) protein is currently the smallest known viral transforming oncogenic protein (Venuti, Paolini et al. 2011), which is 44 amino acids in length, and forms a small alpha-helical transmembrane homodimer, similar to GpA (Schlegel, Wade-Glass et al. 1986). No three-dimensional structure for BPV E5 exists to date; however as described in **Chapter 1.6**, techniques such as molecular dynamics simulation, polarised IR, solution NMR (King, Oates et al. 2011) and mutagenesis studies have all been used to predict the structure of the E5 homodimer and its complex with the transmembrane domain of its cellular target, the platelet-derived growth factor- β receptor (PDGF β R). Using the singly U- $^{13}\text{C}/^{15}\text{N}$ labelled peptides detailed in **Chapter 2.3** and **Chapter 3**, we aimed to investigate the BPV E5 homodimer interface and feed any structural information obtained by ssNMR into our molecular model in order to further refine its structure.

This chapter summarises analyses of the E5 protein in lipid bilayers by applying the sample preparation and ssNMR experimental techniques previously optimised using GpA (**Chapter 4**). One and two-dimensional ssNMR data are presented for singly labelled E5 peptides to probe possible through space interactions between isotopically labelled leucine and phenylalanine (BPV E5_{LF}) residues as well as phenylalanine and tyrosine residues (BPV E5_{FY}). The rationale for labelling these residues is given in Chapter 3. One-dimensional ^{13}C CP-MAS spectra of BPV E5_{FY} were acquired and used to optimise signal to noise ratios when collecting 2D ^{13}C - ^{13}C correlation spectra. Experimental parameters such as temperature, cross polarisation contact time, as well as pulse length and power level and their effect on the signal and resolution of 1D ^{13}C spectra is

detailed. In an attempt to improve spectral resolution of the aromatic ring carbons of phenylalanine and tyrosine residues in the E5 peptide, the effects of changes in bilayer composition (specifically with respect to cholesterol content) were explored using ^{31}P and ^1H ssNMR spectra.

5.2 1D ^{13}C ssNMR of BPV E5 labelled at leucine and phenylalanine (BPV E5_{LF})

E5 peptide samples were reconstituted into DMPC/cholesterol lipid vesicles for ssNMR using the co-solubilisation method, as this method was shown for the GpA peptide to produce a more homogenous sample with an increase in correctly folded α -helical protein (**Chapter 3.2**). The rationale behind the amino acids that were isotopically labelled was explained in **Chapter 3.1** and was based upon CHI molecular modelling of the BPV E5 homodimer to identify those amino acids at homodimer interface which were predicted to be in close proximity upon dimerisation. The first BPV E5 ssNMR sample was prepared by mixing equal amounts of peptides labelled at leucine 24 (BPV E5_L) and phenylalanine 28 (BPV E5_F) (**Figure 5.1**) in a similar manner to that of our co-solubilised GpA sample, with a 2.5:1 lipid to protein ratio and similar cholesterol content of 0.08 molar (5% w/w).

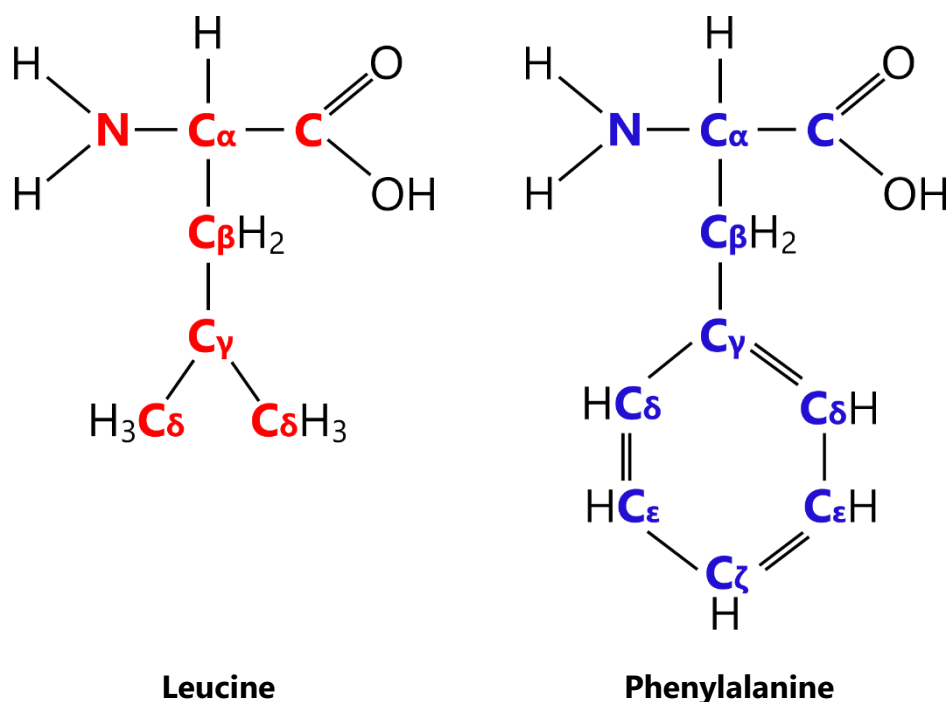


Figure 5.1 Chemical structure of leucine and phenylalanine

Chemical structures of U-¹³C/¹⁵N labelled leucine and phenylalanine are shown, with isotopically labelled carbons and nitrogen's indicated in colour.

One-dimensional ¹³C ssNMR measurements were used to obtain chemical shift information for the BPV E5 peptide in DMPC/cholesterol lipid bilayers for the labelled amino acids leucine 24 and phenylalanine 28. A 1D ¹³C CP-MAS spectrum is shown in **Figure 5.2**. As with our GpA sample, experiments were carried out at sub-zero temperatures (−15 °C).

U-¹³C labelled leucine and phenylalanine within the BPV E5_{LF} sample gave rise to resonances which were assigned using average chemical shift values as found at BMRB. Resonances arising from labelled amino acids matched well with published BRMB values and had good line shape, with an average line width at half height of 1.57 ppm/ 201.64 Hz, and a signal to noise ratio of 23.9. Due to the relatively small difference in chemical shift between a number of carbon atoms (specifically aromatic carbons) and due to the broadness of peaks, a number of resonances were unresolvable due to considerable overlap between resonances, far more than in spectra recorded for GpA. In particular the carbonyl carbons of leucine and phenylalanine (176.95 and 175.39 ppm), leucine C_γ and C_δ 1,2 (26.81 and 24.73/24.09 ppm) and the aromatic carbons of phenylalanine; C_δ 1,2,

C ϵ 1,2 and C ζ (131.32/131.38, 130.46/130.55 and 128.5 ppm respectively) unfortunately resulted in peaks that were unresolvable, and appeared as a single broad peak. The origin for this broadness of the aromatic carbons may be due to the dynamics of the aromatic ring, with internal motions such as ring flips occurring at the ssNMR time scale (micro seconds). The broadness of the aromatic carbon signals in phenylalanine may be increased further as a consequence of the amino acid adopting multiple conformations at the homodimer interface resulting in a much broader signal than when in a single fixed confirmation. Remaining carbon resonances were reasonably well-resolved and were assigned based upon their chemical shift values as summarised in **Table 5.1**.

From the 1D ^{13}C spectrum it was also observed that the phenylalanine C γ carbon had a much lower intensity in comparison to other labelled carbons within phenylalanine. The C γ carbon atom of phenylalanine is a quaternary carbon with no directly bound hydrogen atom from which to transfer magnetisation from in CP experiments, as such the intensity is greatly reduced in comparison. From initial 1D ^{13}C experiments (data not shown) it was also observed that at 10 kHz MAS the spinning side band from DMPC $-(\text{CH}_2)_n-$ (lipid acyl tails) fell in the same region as the phenylalanine ring carbons, as such the spinning speed was increased from 10 to 11 kHz to remove any artefacts from running experiments at this speed. Unlike the GpA sample that was also prepared using the co-solubilisation method, BPV E5 samples showed no evidence in the 1D ^{13}C spectrum of any observable secondary species, with all protein correctly folded in one single environment in agreement with our biophysical characterisation of BPV E5 samples (**Chapter 3.2**).

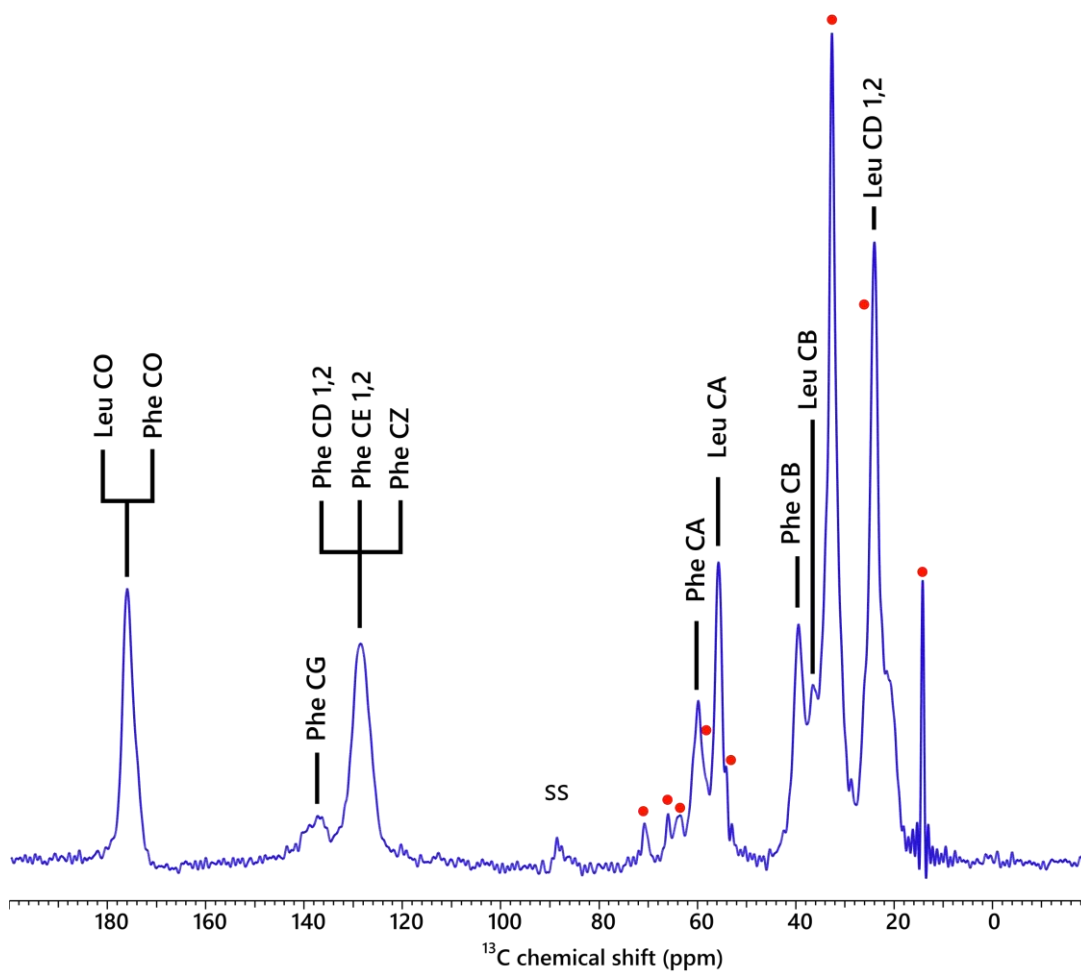


Figure 5.2 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled BPV E5_{LF} in DMPC liposomes

Spectrum recorded at 500 MHz with 11 kHz MAS at 258 K (-15 °C). A CP contact time of 1000 μs was used, a 2.5 second recycle delay with 100 kHz proton decoupling during acquisition for 1024 scans. Resonances arising from U- ^{13}C labelled leucine and phenylalanine are labelled accordingly and those arising from natural abundance ^{13}C from DMPC indicated by •.

Table 5.1 ^{13}C chemical shift data for labelled BPV E5 Leu 24 and Phe 28

Leu 24	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
CO	175.98	174.97	1.01
C α	55.70	53.64	2.06
C β	39.41	40.30	-0.89
C γ	23.94	24.81	-0.87
C δ 1,2	23.94	24.72	-0.78

Phe 28	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
CO	174.77	173.42	1.35
C α	59.88	56.11	3.77
C β	36.58	37.99	-1.41
C γ	136.54	134.86	1.68
C δ 1,2	128.68	128.49	0.19
C ϵ 1,2	128.68	128.49	0.19
C ζ	128.68	128.57	0.11

5.2.1 2D ^{13}C - ^{13}C ssNMR of BPV E5_{LF}

Once 1D ^{13}C spectra were collected and assigned, two-dimensional (2D) ^{13}C - ^{13}C DARR experiments were recorded on the BPV E5_{LF} sample in order to probe any possible through-space interactions between labelled leucine and phenylalanine at the homodimer interface in the singly labelled peptides. As with GpA, 2D ^{13}C - ^{13}C DARR experiments were run at both short and long mixing times (20 and 400 ms) in order to observe both short range and longer range correlations within and between the two labelled amino acids. **Figure 5.3** shows the 2D ^{13}C - ^{13}C DARR spectrum obtained at a 20 ms mixing time. At this short mixing time, in addition to the strong diagonal peaks that were observed, multiple cross peaks off the diagonal arose due to intra-residue correlations within each labelled amino acid. Intra-residue cross peaks were observed between; leucine C α – C β , leucine C α – C δ 1/2, leucine C α – CO, leucine C β – C δ 1/2, leucine C β – CO, leucine C δ 1/2 – CO, phenylalanine C α – C δ / ϵ / ζ , phenylalanine C α – CO, phenylalanine C β – C δ / ϵ / ζ , phenylalanine C β – CO, and phenylalanine C δ / ϵ / ζ – C γ CO.

Unlike GpA samples where labelled carbon atoms on valine and glycine had distinct chemical shifts, the BPV E5_{LF} sample yielded a number of carbon atoms with similar chemical shifts and as a direct result there was considerable overlap between a number of cross peaks, in particular those from the aromatic carbons in phenylalanine. As with the 1D ^{13}C spectrum, resonances from leucine C γ and C δ 1,2 were also indistinguishable in the 2D DARR spectrum and as such no cross peaks between leucine C γ were assigned.

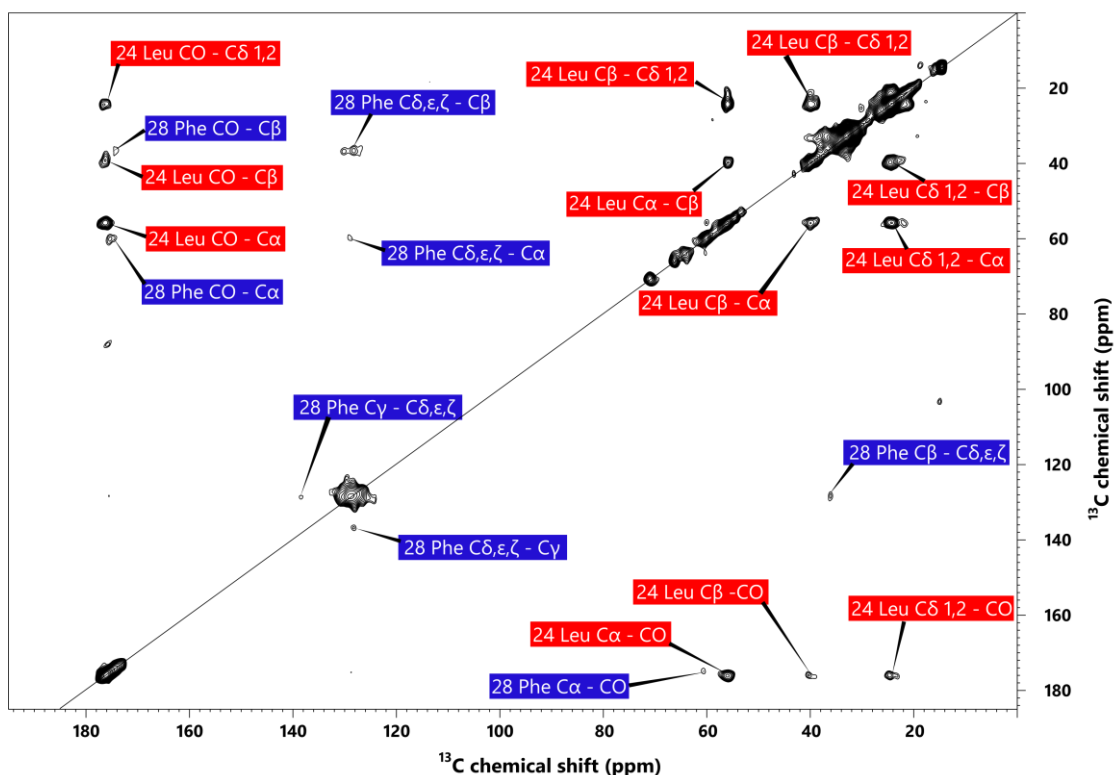


Figure 5.3 20 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5_{LF} in DMPC liposomes.

2D ^{13}C - ^{13}C DARR correlation spectrum of singly labelled BPV E5_{LF}, acquired in 48 hours with a mixing time of 20 ms. Spectrum was recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C) for 160 scans. Cross peaks are labelled according to amino acid spin system, with cross peaks arising from Leucine 24 in red and Phenylalanine 28 in dark blue.

At longer DARR mixing times (400 ms) (**Figure 5.4**), only two additional cross peaks were observed and these are attributable to the intra-residue correlations within phenylalanine of $\text{C}\alpha - \text{C}\beta$ and $\text{C}\delta/\epsilon/\zeta - \text{CO}$. **Figure 5.5** shows an overlay of short and long mixing time spectra. Unlike singly labelled GpA, at long mixing times we observed no inter-helical cross peaks in E5 due to through space dipolar couplings between leucine 24 and phenylalanine 28 and only intra-residue cross peaks were observed.

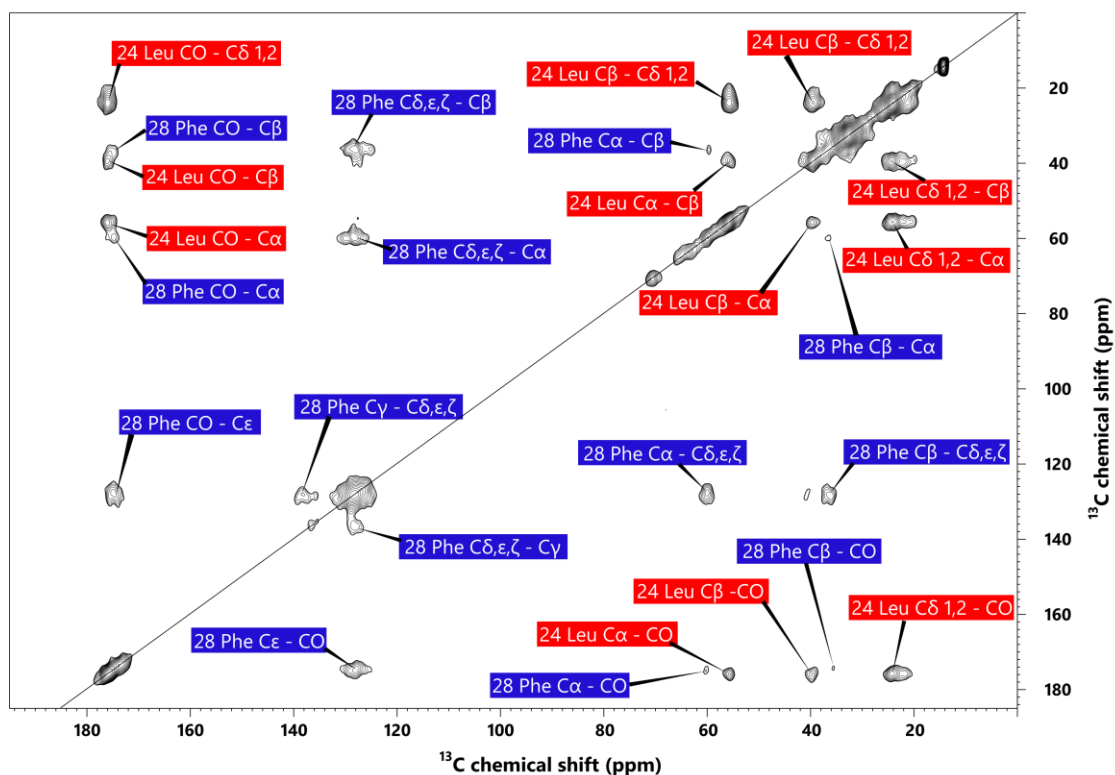


Figure 5.4 400 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5_{LF} in DMPC liposomes.

2D ^{13}C - ^{13}C DARR correlation spectrum of singly labelled BPV E5_{LF}, acquired in 48 hours with a mixing time of 400 ms. Spectrum was recorded at 500 MHz with 11 kHz MAS at 258 K (-15 °C) for 160 scans. Cross peaks are labelled according to amino acid spin system, with cross peaks arising from Leucine 24 in red and Phenylalanine 28 in dark blue.

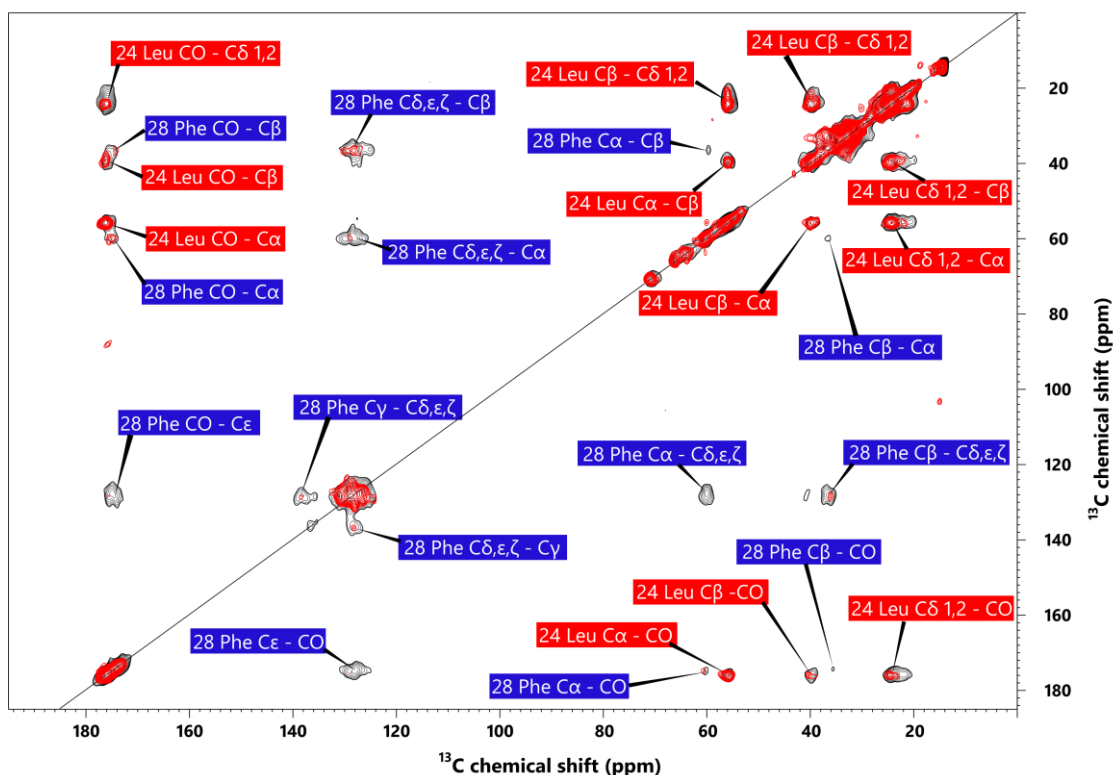


Figure 5.5 Overlay of 20 and 400 ms 2D ^{13}C - ^{13}C DARR spectra of singly labelled BPV E5_{LF} in DMPC liposomes.

Overlay of 2D ^{13}C - ^{13}C DARR correlation spectra of singly labelled BPV E5_{LF}, acquired with a mixing time of 20 ms (red spectrum) and 400 ms (black spectrum). Spectra recorded at 500 MHz with 11 kHz MAS at 258 K (-15 °C). At both short and long mixing times only intra-residue cross peaks were observed.

This was an unexpected result, since the current model of the BPV E5 homodimer suggests that leucine 24 and phenylalanine 28 are in close proximity at the E5 homodimer interface. There are a number of reasons that could explain why we were unable to observe any inter-helical correlation between the two labelled amino acids. The first reason may be simply that the current model for the E5 homodimer is wrong and that leucine and phenylalanine are not in close proximity to one another at the homodimer interface. The second reason may be that the model is wrong, but only slightly. Looking again at the CHI model of the E5 homodimer, which is in very good agreement with mutagenesis studies (Horwitz, Weinstat et al. 1989; Meyer, Xu et al. 1994; Mattoon, Gupta et al. 2001) originally used to construct the current model of the E5 dimer, a number of carbon atoms within leucine and phenylalanine were predicted to be

close together in space (as detailed in **Table 3.1** in **Chapter 3**). The dynamic of the aromatic ring, such as ring flips that can occur may also be a reason for the fact of an observable cross peak as the motional process that occurs interrupts the recoupling of the dipolar interaction. Leucine C δ 1 and phenylalanine C δ 2 gave the shortest predicted inter-helical distance of 4.58 Å, with leucine C δ 1 and phenylalanine C γ yielding the next shortest inter-helical distance of 4.91 Å. If the NMR data obtained using the GpA sample are used as a gauge of what distances are observable using DARR under these conditions, then we know that inter-helical coupling over a distance predicted to be 4.62 Å (between labelled valine and glycine) was observable. As such, inter-helical coupling over a distance of 4.58 Å between labelled leucine and phenylalanine in the very similar E5 sample should have yielded a cross peak at long mixing times. However, at 4.91 Å the chances of observing a cross peak as a result of inter-helical coupling between leucine C δ 1 and phenylalanine C γ is greatly reduced as this distance was starting to approach the limits of those that were detectable in the GpA experiments. Therefore, a very slight rotation of the helices in the E5 homodimer may render the inter-helical coupling unobservable. A third reason for not observing inter-helical coupling is the fact that, in addition to longer range distances having weaker signals, phenylalanine C γ is a quaternary carbon atom with a much lower signal intensity in comparison to other carbon resonances as a result of decreased signal from the CP transfer and as the ^1H - ^{13}C heteronuclear couplings that are reintroduced by the DARR pulse sequence are weaker due to the lack of a directly bound proton. These two strikes make the possibility of observing this inter-helical coupling very low. Finally, due to equipment availability and time constraints the number of scans recorded for E5 samples was much lower than that of singly labelled GpA, and as such this may have been another reason for the lack of inter-helical cross peaks observed.

Having been unable to observe any inter-helical coupling between labelled leucine and phenylalanine in our BPV E5_{LF} ssNMR sample, we then moved onto using our next pairing of isotopically labelled peptides. From the findings observed with our BPV E5_{LF} ssNMR sample, such as low signal intensity for phenylalanine C γ and the broadness of phenylalanine aromatic ring carbons which resulted in a single unresolvable peak, we sought to make some changes to our next sample in an attempt to improve the resolution of the resonances observed and to try and improve signal to

noise ratio values so that when running 2D ^{13}C - ^{13}C experiments any weak inter-helical couplings could be observed.

5.3 1D ^{13}C ssNMR of BPV E5 labelled at tyrosine and phenylalanine (BPV E5_{FY})

A ssNMR sample containing an equimolar mixture of singly labelled E5 peptides, isotopically labelled at phenylalanine 28 (BPV E5_F) or tyrosine 31 (BPV E5_Y) as shown in **Figure 5.6**, was prepared. As before, these residues were selected because they were predicted to be in close contact at the E5 homodimer interface from the CHI molecular model and previous studies.

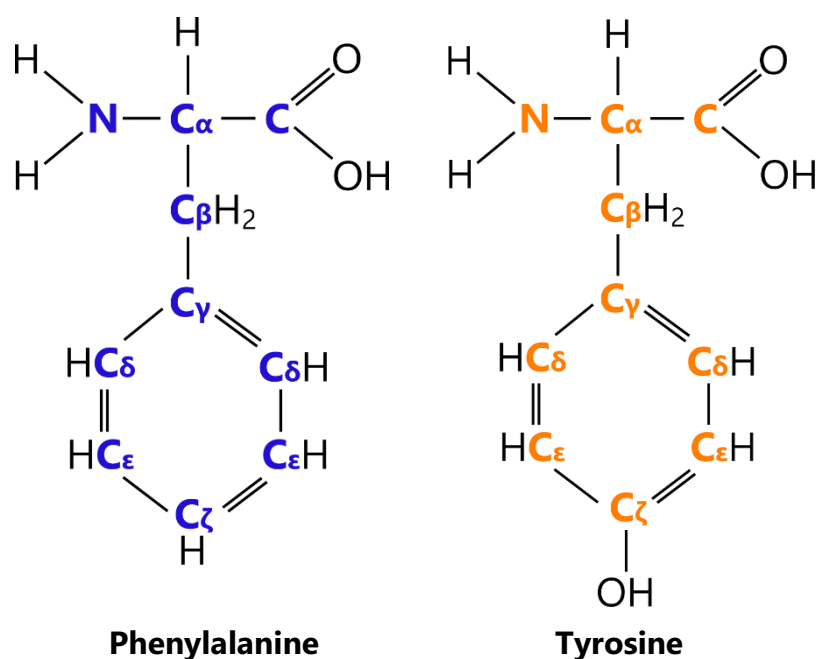


Figure 5.6 Chemical structure of phenylalanine and tyrosine.

Chemical structures of U- ^{13}C / ^{15}N labelled leucine and phenylalanine are shown, with isotopically labelled carbons and nitrogens indicated in colour

Having observed with BPV E5_{LF} that resonances arising from the aromatic carbons of phenylalanine gave rise to a single broad peak (**Figure 5.5**), rather than

being resolvable as separate individual resonances, it was decided to change the sample composition in an attempt to improve the resolution. This was important because in the E5_{FY} sample the shortest inter-helical cross peaks were predicted to be between aromatic carbons. One way to improve resolution is to decrease motion of embedded peptides by increasing the rigidity of the lipid membrane. A recent study (Luo, Cady et al. 2009) showed that using lipid mixtures and increasing the cholesterol content, in order to mimic the native membrane system more closely, can increase the rigidity of the membranes and reduce uniaxial rotational diffusion, allowing for ssNMR experiments to be recorded at physiological temperatures (37 °C). Typically, running ssNMR experiments at low (< 0 °C) temperature is used to help improve spectral resolution by reducing Brownian motions within the lipid bilayer. This reduces internal motions within proteins embedded within the bilayer and freezes uniaxial diffusion, which could otherwise complicate ssNMR spectra due to intermediate timescale line broadening (Luo, Cady et al. 2009). By reducing motions within the protein, multiple side chain conformations that can lead to line broadening are “frozen out” and line widths improved. Although by reducing the temperature, T_1 relaxation times and dipolar couplings are increased, which can lead to increased line width (Abdine, Verhoeven et al. 2010). Furthermore, the cooling of the sample to sub-zero temperatures may also result in line broadening by introducing heterogeneous line broadening as a result of sample inhomogeneity due to the freezing out of multiple side chain conformations. Therefore it would be beneficial to reduce motions within the lipid bilayer whilst also running experiments at physiological temperatures.

By increasing the cholesterol content in ssNMR samples, it was believed that in addition to making the DMPC lipid membrane in ssNMR samples more rigid, it would also make the membranes more biologically relevant. Cholesterol is a major component of mammalian cell membranes, accounting for up to 50 mol% of the membrane lipids of the plasma membrane (van Meer, Voelker et al. 2008) and serves a wide range of functions such as controlling membrane fluidity, reducing passive permeability and increasing mechanical strength of the membrane (Bittman, Clejan et al. 1984; Urbina, Pekerar et al. 1995) . It was hoped that by increasing the cholesterol content of lipid bilayers in samples prepared for ssNMR, membranes would become more rigid, since the DMPC lipid acyl tails become more ordered as cholesterol intercalates between

them, thereby introducing local order within acyl chains on lipid tails and reducing side chain flexibility of the phenylalanine and tyrosine aromatic rings in both labelled amino acids. This would potentially reduce the line width of carbon resonances and allow us to then observe the expected inter-helical couplings as predicted from the CHI molecular model and assign them unambiguously.

Therefore, the cholesterol content of the DMPC lipid bilayers used for the BPV E5_{FY} sample was increased from the 5% w/w cholesterol (0.08:1 molar ratio) used in previous samples to 30% w/w (0.60:1 molar ratio). This ratio is much more in line with the actual molar ratio found at the plasma membrane (1:1) and Golgi apparatus (0.2:1 (van Meer 1998)) to which E5 locates (Burkhardt, Willingham et al. 1989), making these samples more biologically relevant. A 1D ¹³C CP MAS ssNMR spectrum of BPV E5_{FY} sample is shown in **Figure 5.7**. Initial 1D ¹³C experiments were carried out at a more physiologically relevant temperature of 25 °C (as detailed later in this chapter, section **5.3.3**), however only resonances from DMPC lipid were observed at this temperature. Therefore, as with previous samples, 1D ¹³C spectra were recorded at low temperature (–15 °C).

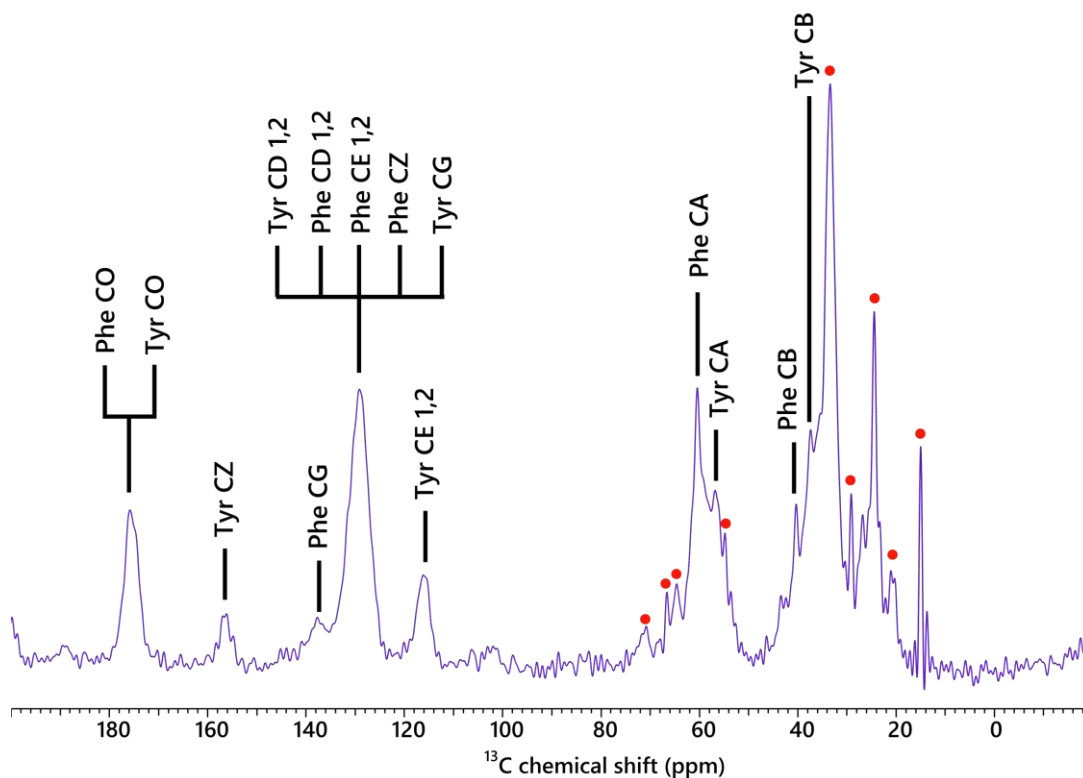


Figure 5.7 1D proton-decoupled ^{13}C CP-MAS spectrum of singly labelled BPV E5_{FY} in DMPC liposomes

Spectrum recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C). A CP contact time of 1000 μs was used, a 2.5 second recycle delay with 80 kHz proton decoupling during acquisition for 1024 scans. Resonances arising from U- ^{13}C labelled phenylalanine and tyrosine are labelled accordingly and those arising from natural abundance ^{13}C from DMPC indicated by •.

Phenylalanine and tyrosine carbon resonances were assigned using BMRB average chemical shifts, with resonances in good agreement with published values. Having increased the cholesterol content of lipid bilayers, an improvement in the resolution of 1D ^{13}C spectra was expected, in particular in the aromatic carbon region of the spectrum. As can be seen in the 1D ^{13}C spectrum obtained (**Figure 5.7**), the additional cholesterol did not appear to improve the spectral resolution of this region. The average signal to noise ratio of 17.4 was lower than that achieved with BPV E5_{FY} (23.9) and the average peak width at half height of 1.73 ppm was slightly larger than that

produced by BPV E5_{LF} at 500 MHz (1.57 ppm). This increase may have been due in part to the broad peak from phenylalanine and tyrosine aromatic carbons which were still unresolvable due to overlap between resonances even with increased cholesterol concentration. Furthermore, the two samples were not directly comparable due to having been recorded at different spectrometer field strengths (500 vs. 600 MHz), where at higher field the effects of chemical shift anisotropy (CSA) are much stronger. The additional resonances from tyrosine C δ 1,2 and C γ over those found in the BPV E5_{LF} sample may also have contributed to the broad aromatic peak having a larger peak width in BPV E5_{FY}, even with a higher cholesterol content.

The larger number of peaks in the BPV E5_{FY} spectrum led to a higher degree of spectral overlap, particularly in carbonyl (175.39 ppm for Phe and 175.33 ppm for Tyr CO) and aromatic regions of the spectrum (Phe C δ 1,2, C ϵ 1,2 and C ζ of 131.32/131.38, 130.46/130.55 and 128.5 ppm respectively vs. Tyr C δ 1,2 = 132.40/132.39 ppm and C γ = 127.34 ppm). This resulted in unresolvable signals that appeared as broad single peaks. Phenylalanine C α (60.14 ppm) appeared as a single peak with a distinguishable shoulder for tyrosine C α (56.66 ppm). Although the majority of resonances arising from labelled phenylalanine and tyrosine overlapped, there were a few resonances which had distinct chemical shifts such as phenylalanine C γ (135.60 ppm), tyrosine C ϵ 1,2 (117.69, 117.77 ppm) and tyrosine C ζ (152.46). These distinct chemical shifts should allow for the predicted inter-helical couplings to be observed. Remaining carbon resonances and were assigned based upon their chemical shift values as summarised in **Table 5.2**.

Table 5.2 ^{13}C chemical shift data for labelled BPV E5 Phe 28 and Tyr 31

Phe 28	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
CO	175.57	173.42	2.15
C α	60.14	56.11	4.03
C β	37.1	37.99	-0.89
C γ	137.92	134.86	3.06
C δ 1,2	128.94	128.49	0.45
C ϵ 1,2	128.94	128.49	0.45
C ζ	128.94	128.57	0.37

Tyr 31	Observed chemical shift (ppm)	BMRB chemical shift TMS (ppm)	Δ
CO	175.57	173.56	2.01
C α	56.66	56.13	0.53
C β	37.18	37.33	-0.15
C γ	128.94	126.51	2.43
C δ 1,2	128.94	130.42	-1.48
C ϵ 1,2	115.72	115.77	-0.05
C ζ	156.18	152.56	3.62

The increased membrane cholesterol content had no obvious effect on the homogeneity of the peptide, as there was no evidence from the 1D ^{13}C spectrum that any secondary species was present in the samples prepared. The spectrum suggested homogeneous protein correctly folded in one single environment, as was seen with the BPV E5_{LF} sample, and that there was no detrimental effect of increasing the cholesterol content of lipid bilayers on protein fold.

5.3.1 Optimisation of experimental parameters in 1D CP experiments

Once initial 1D ^{13}C experiments were recorded, we set about to optimise experimental parameters in order to try and improve signal to noise values and resolution in the regions of our spectra where we expected to observe inter-helical cross peaks in 2D ^{13}C - ^{13}C DARR experiments. Due to the inherent low sensitivity of ssNMR experiments in comparison to its solution counterpart, it is commonplace to optimise

radio frequency (RF) pulses and power level of pulses used, within the experiments within limits of the probe, in order to achieve the optimal amount of signal. In addition to optimising power levels and pulse lengths, the effect of temperature can also be optimised in order to improve signal to noise values and resolution.

5.3.2 Optimisation of experimental parameters: temperature

It has been noted that decreasing the temperature at which ssNMR experiments are conducted can result in an increase in NMR signal leading to much better signal to noise, by slowing molecular motions and improving the efficiency of proton decoupling and cross polarisation (Abdine, Verhoeven et al. 2010) (Hiller, Krabben et al. 2005; Cady, Mishanina et al. 2009) A reduction in temperature is also a commonly used method to improve Boltzmann polarisation to improve signal to noise. Although it has also been shown that by increasing the rigidity of the membranes, experiments can be conducted at higher, more physiological temperatures (Luo, Cady et al. 2009). Increasing the cholesterol content of the lipid membrane used in ssNMR samples prepared should allow for experiments to be run at higher temperatures by increasing order within the membrane.

In order to investigate the effect of temperature on samples prepared using a higher cholesterol content, we ran a series of 1D ^{13}C CP MAS experiments on our BPV E5_{VF} sample which was prepared with an increased cholesterol content of 30% w/w (0.60 molar ratio). 1D ^{13}C experiments were recorded over a range of temperatures (**Figure 5.8**), starting at 25 °C, above the phase transition temperature (T_m) of DMPC lipid, at which the membrane should be in the liquid-crystalline phase (L_α), through to -20 °C, at which membranes should be in the gel phase (L_β). At 25 °C, signal intensity was relatively low with only nuclei from DMPC lipid molecules giving observable, sharp peaks, with average line widths of $\nu_{1/2} = 0.74$ ppm. Decreasing the temperature from 25 °C to 10 °C showed an improvement in the magnitude of the NMR signal obtained, with sharp peaks from lipid ($\nu_{1/2} = 0.86$ ppm) and the appearance of broad resonances from phenylalanine and tyrosine carbonyl and aromatic ring carbons. At this temperature, signals from aromatic ring carbons were unresolvable with a single broad peak at around 130 ppm.

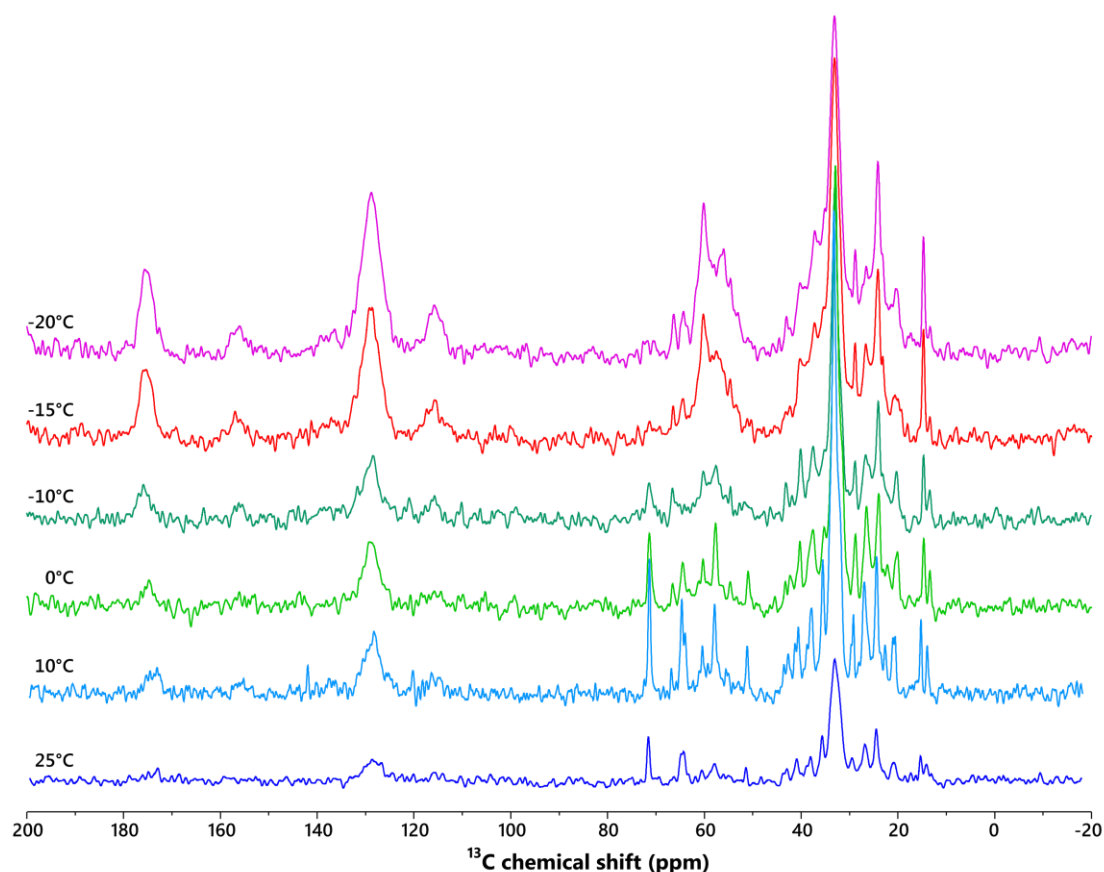


Figure 5.8 1D proton-decoupled ^{13}C CP-MAS spectra of singly labelled BPV E5_{FY} at decreasing temperature

Spectra recorded at 600 MHz with 11 kHz MAS. A 2.5 second recycle delay and a CP contact time of 1000 μs was used with 80 kHz proton decoupling during acquisition each for 512 scans. The sample cooling gas temperature was decreased from 25 to $-20\text{ }^{\circ}\text{C}$ between each experiment as indicated alongside individual 1D spectra.

Cooling the sample further to $0\text{ }^{\circ}\text{C}$, showed no significant improvement in signal-to noise, line widths, or resolution observed. Recording spectra at sub-zero temperatures from $-10\text{ }^{\circ}\text{C}$ - $-20\text{ }^{\circ}\text{C}$ resulted in steady improvement in signal to noise. However, at $-10\text{ }^{\circ}\text{C}$ the resolution of the 1D spectrum was reduced ($\nu_{1/2} = 1.31\text{ ppm}$ on average) and overlap was more pronounced. Decreasing the temperature further to $-15\text{ }^{\circ}\text{C}$ saw a further decrease in resolution (and increase in signal-to noise), in particular in the aliphatic region of the spectrum where previously it was possible to differentiate between individual resonances. At $-15\text{ }^{\circ}\text{C}$ there was also a decrease in the number of resonances arising from DMPC lipid in the aliphatic region, likely due to the reduced

internal motions within the lipid molecules causing these resonances to broaden out to the point at which they become unobservable. At $-20\text{ }^{\circ}\text{C}$ there was no significant change in either signal-to-noise or resolution compared to the spectrum acquired at $-15\text{ }^{\circ}\text{C}$ and, as such, all further experiments were recorded at $-15\text{ }^{\circ}\text{C}$. **Figure 5.9** summarises the effect of sample cooling on average peak width at half height of ^{13}C resonances observed. Although the effect on the average peak width appears to be dominated by the width of lipid resonances rather than of those arising from protein signals.

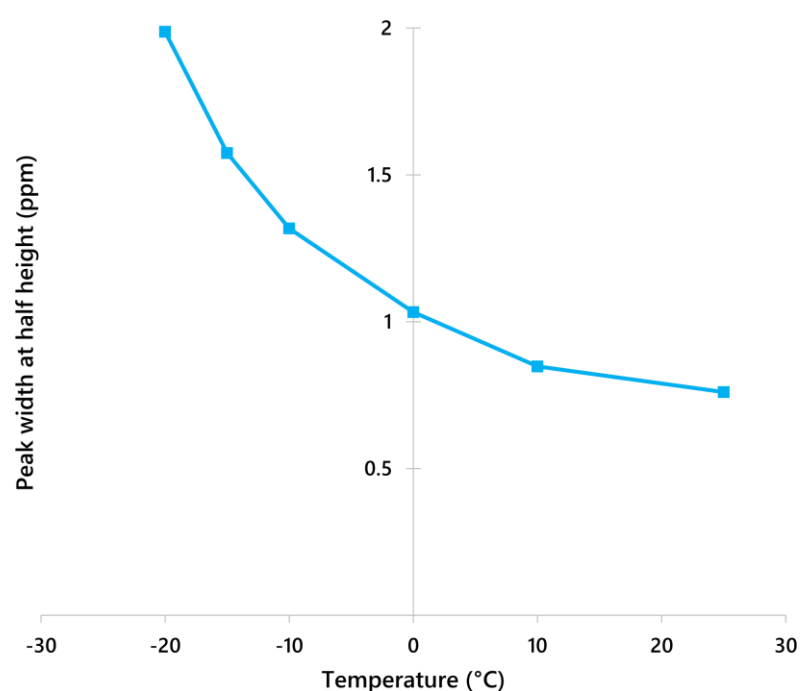


Figure 5.9 Average peak width of ^{13}C BPV E5_{FY} resonances as a function of temperature

The average peak width at half height of ^{13}C resonances arising from labelled phenylalanine and tyrosine in addition to those arising from natural abundance DMPC, were measured and plotted as a function of the experimental temperature at which spectra were recorded. Spectra were recorded at 25 - $-20\text{ }^{\circ}\text{C}$.

The results obtained from the 1D ^{13}C spectra recorded suggest increasing the cholesterol content of the membranes did not facilitate use of higher more physiological temperatures for ssNMR measurements, with higher temperatures only yielding signals for natural abundance nuclei in DMPC lipid molecules although with sharp well resolved peaks. At higher temperatures lipid molecules undergo fast Brownian motions, with fast

uniaxial rotational diffusion resulting in sharpening of peaks. To understand why the addition of increased cholesterol resulted in the loss of protein signal at higher temperature an analysis of the effect of increasing cholesterol content on DMPC lipid membrane composition was conducted, as described later in this chapter (section 5.4).

5.3.3 Optimisation of experimental parameters: contact time

For ^{13}C ssNMR experiments, cross polarisation (CP) was used in order to transfer magnetisation from highly abundant protons with a high gyromagnetic ratio to less abundant carbon atoms which have a lower gyromagnetic ratio, in order to improve sensitivity. The experimental parameter in the pulse program that relates to the transfer of this magnetisation is the so called contact time. The build-up of magnetisation for each individual carbon is dependent on the proton dipolar coupling network to which it belongs and the degree of protonation of each carbon as well as molecular kinetics such as methyl group rotation or aromatic ring flips (T_{HC}) (Kolodziejcki and Klinowski 2002). The larger the number of protons, the stronger the dipolar coupling and therefore the faster the cross polarisation transfer. Molecular motions that occur on the microsecond time scale may average the dipolar coupling used to transfer magnetisation resulting in a decay in build-up intensity. For example protonated carbons, with reduced mobility, build-up at a faster rate, methyl groups which have higher rotational mobility and build up more slowly. Additionally, at longer CP contact times, more magnetisation is allowed to be transferred to the carbon atoms, therefore improving sensitivity of the experiment, up until a point at which the protons begin to relax as a function of $T_{1\rho}$. As a result different peaks will cross polarise and build up at different rates. Therefore when choosing the optimal CP contact time, the entire spectrum as a whole has to be considered and a compromised contact time chosen based upon the overall intensity of resonances of interest from labelled amino acids within the sample. In order to optimise signal to noise, a range of 1D ^{13}C CP MAS spectra of BPV E5_{FY} were recorded with increasing CP contact times, from 100 to 1500 μs (Figure 5.10).

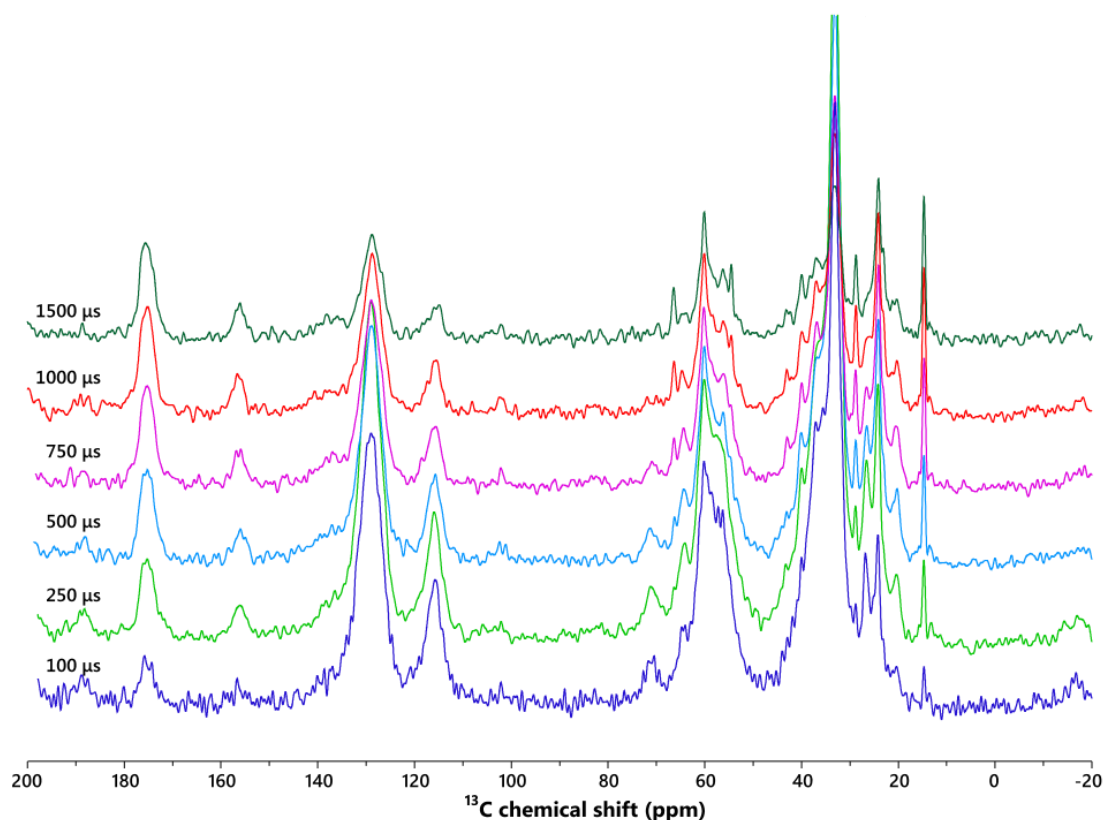


Figure 5.10 1D proton-decoupled ^{13}C CP-MAS spectra of singly labelled BPV E5_{FV} at increasing CP contact time

Spectra recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C). A 2.5 μsec ^1H 90° with a 2.5 second recycle delay was used with 100 kHz SPINAL-64 proton decoupling during acquisition for 512 scans. The cross polarisation (CP) contact time was increased from 100 to 1500 μs between each experiment as indicated alongside 1D spectra.

As can be seen from the 1D ^{13}C spectra obtained in **Figure 5.10** and the graph of plotted signal intensities at increasing CP contact time (**Figure 5.11**) using shorter contact times, resulted in improved signal for phenylalanine ($\text{C}\delta$ 1,2/ $\text{C}\epsilon$ 1,2/ $\text{C}\zeta$) and tyrosine ($\text{C}\delta$ 1,2) aromatic ring carbons. As the contact time was increased, from 100 to 250 μs the intensity of these resonances increased, but beyond 250 μs the intensity of these aromatic carbon resonances decreased at a fast rate as a function of increasing contact time. A similar trend was also seen for tyrosine $\text{C}\epsilon$ 1,2 carbon atoms, although the initial starting intensity was much lower and the rate at which the signal intensity decreased was much slower. For the resonance relating to phenylalanine and tyrosine CO carbonyl carbons, which have no directly-bound ^1H atom, the opposite trend was observed. At lower contact times the carbonyl peaks had a much lower intensity, but as

the CP contact time was increased the signal intensity increased up until 1000 μs , at which point the signal intensity remained the same upon further increase to 1500 μs . For tyrosine C ζ , signal was only observed at contact times of 750 μs and above, before which no distinguishable peak was detected. Increasing the contact time resulted in an increase in the signal intensity for this peak between 750 and 1000 μs after which there was little increase in signal.

A CP contact time of 750 μs was selected for use in all further 1D and 2D ^{13}C experiments, so as to ensure that all inter-helical cross peaks were observable between the predicted interacting atoms of each labelled amino acid.

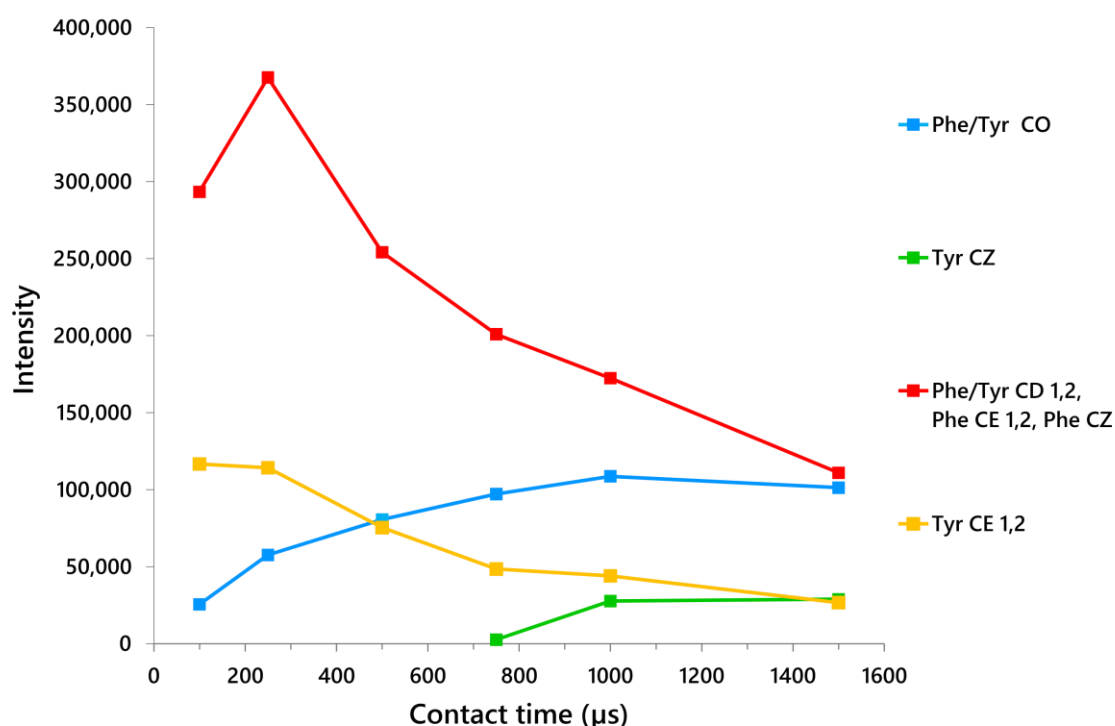


Figure 5.11 Graph of 1D ^{13}C resonance signal intensity at increasing cross polarisation (CP) contact times

Absolute intensity of assigned phenylalanine and tyrosine aromatic ring and carbonyl carbon resonances, plotted against increasing cross polarisation (CP) contact time. Resonance intensities were obtained from 1D ^{13}C spectra (as shown in **Figure 5.10**) recorded at increasing contact time.

5.3.4 2D ^{13}C - ^{13}C ssNMR of BPV E5 labelled at phenylalanine & tyrosine (BPV E5_{FY})

Following optimisation of ssNMR experimental parameters to obtain optimal signal intensity, 2D ^{13}C - ^{13}C DARR correlation spectra were recorded at short and long mixing times (20 and 400 ms) to probe any possible inter-helical, through-space interactions between the two labelled amino acids in the BPV E5_{FY} ssNMR sample.

A 2D ^{13}C - ^{13}C DARR spectrum obtained using a 50 ms mixing time is shown in **Figure 5.12**, and a number of off diagonal cross peaks were observed. Although experimental parameters were optimised in order to achieve optimal signal to noise and to increase signal in the region of interest in the ^{13}C spectrum, in moving to higher field we observed a decrease in signal to noise compared to spectra recorded at 500 MHz. Therefore, the intensity of these cross peaks was much lower than that observed in our previous BPV E5_{LF} sample at lower field (500 MHz). As observed in the 1D ^{13}C spectrum (**Figure 5.7**), there was considerable overlap between signals obtained for phenylalanine and tyrosine, making assignment of individual cross peaks difficult. From the 2D spectrum recorded at short mixing time, intra-residue cross peaks were observed between; phenylalanine C α – C β , C α – (C δ 1/2, C ϵ 1/2, C ζ), C α – CO, C β – (C δ 1/2, C ϵ 1/2, C ζ), C β – CO, and (C δ 1/2, C ϵ 1/2, C ζ) – CO, no intra-residue cross peaks to phenylalanine C γ were observed. For labelled tyrosine; C α – C β , C α – (C δ 1/2, C γ), C α – CO, C β – (C δ 1/2, C γ), C β – CO, (C δ 1/2, C γ) – C ϵ , (C δ 1/2, C γ) – CO intra-residue cross peaks were observed.

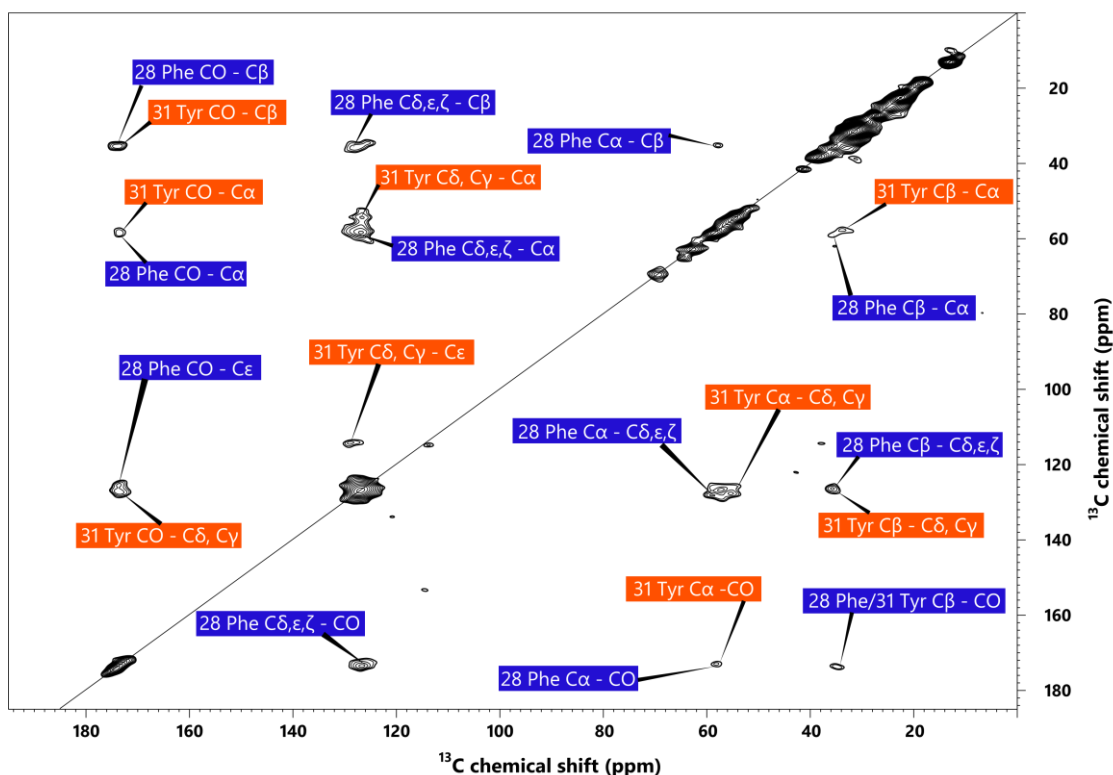


Figure 5.12 50 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5_{YF} in DMPC liposomes

2D ^{13}C - ^{13}C DARR correlation spectrum of singly labelled BPV E5_{YF}, acquired in 48 hours with a mixing time of 50 ms. Spectrum was recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C) for 148 scans. Cross peaks are labelled according to amino acid spin system, with cross peaks arising from phenylalanine 28 in dark blue and tyrosine 31 in orange.

Increasing the DARR mixing time from 50 ms to 400 ms (**Figure 5.15**) yielded only two additional intra-residue cross peaks, namely phenylalanine C γ – C α and tyrosine (C δ 1/2, C γ) to C ζ . As with our BPV E5_{LF} sample, no inter-helical cross peaks were observed at long mixing times between labelled phenylalanine and tyrosine, with all cross peaks observed arising from intra-residue correlations between carbon atoms within each of the two labelled amino acids.

From the CHI molecular model of the BPV E5 homodimer, it was expected that side chains of phenylalanine 28 and tyrosine 31 would be in close proximity to each other at the homodimer interface, with the coupling between phenylalanine C ϵ 1 and tyrosine C ϵ 1 predicted to have the shortest average inter-helical distance of 3.64 Å. Additional inter-helical couplings between phenylalanine C δ 1, C ϵ 1 and C ζ to tyrosine

C δ 1, C ϵ 1 and C ζ , as detailed in **Table 3.1**, (**Chapter 3**) were all predicted to have inter-helical couplings with inter-atomic distances shorter than 4.5 Å. Based upon the results obtained with singly labelled GpA, coupling over these distances should be detectable at long mixing times. Only the coupling between phenylalanine C ζ and tyrosine C ζ was predicted to have an inter-helical distance (4.81 Å) near the limits of the observable range.

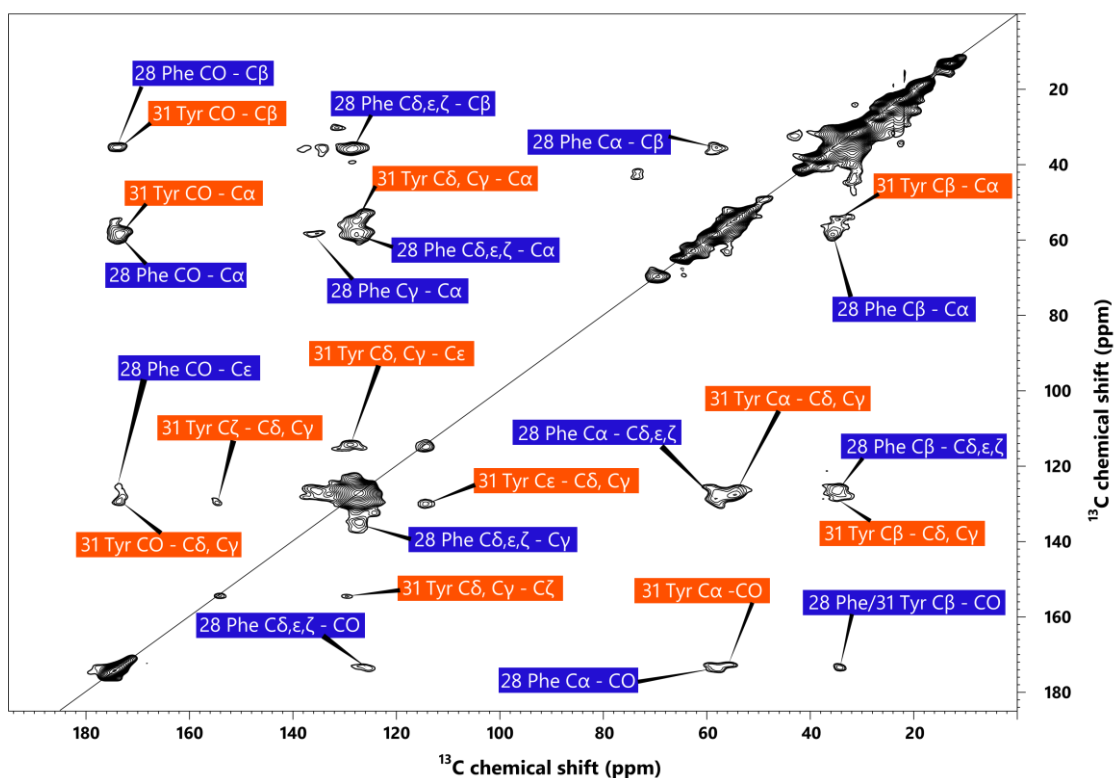


Figure 5.13 400 ms 2D ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5_{VF} in DMPC liposomes

2D ^{13}C - ^{13}C DARR correlation spectrum of singly labelled BPV E5_{VF}, acquired in 48 hours with a mixing time of 400 ms. Spectrum was recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C) for 512 scans. Cross peaks are labelled according to amino acid spin system, with cross peaks arising from phenylalanine 28 in dark blue and tyrosine 31 in orange.

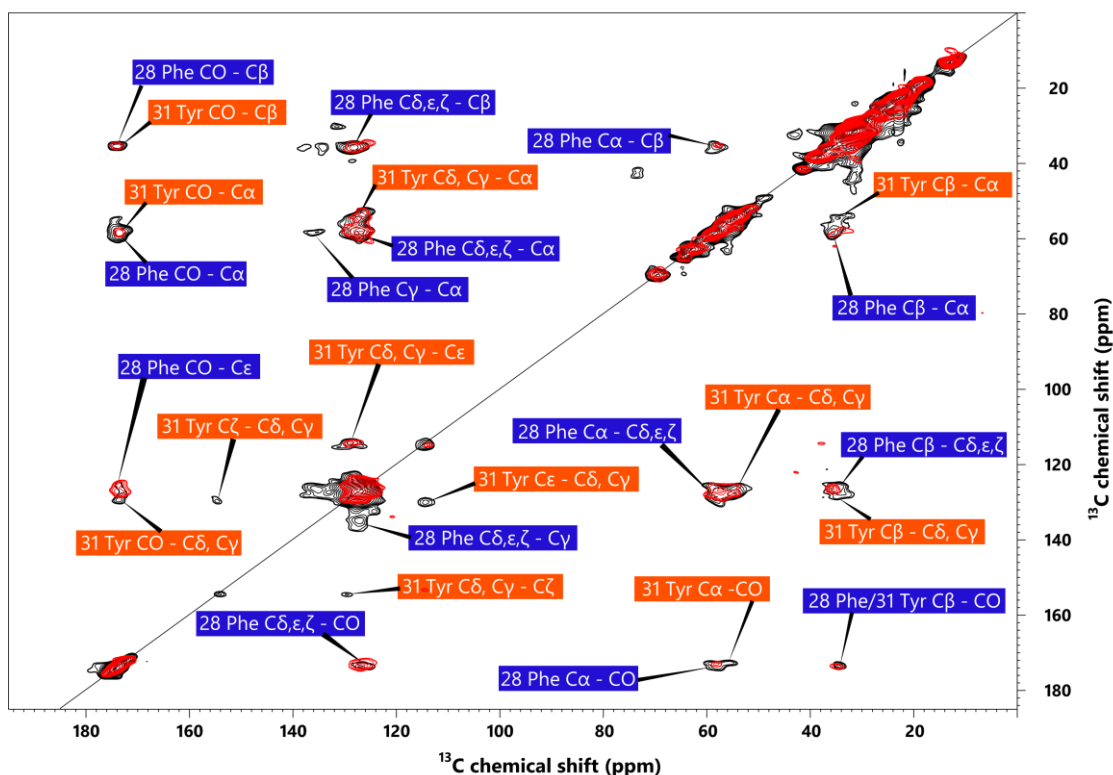


Figure 5.14 50 vs 400 ms $2D$ ^{13}C - ^{13}C DARR spectrum of singly labelled BPV E5_{YF} in DMPC liposomes

$2D$ ^{13}C - ^{13}C DARR correlation spectrum of singly labelled BPV E5_{YF}, with a mixing time of 50 ms (red) and 400 ms (black). Spectra was recorded at 600 MHz with 11 kHz MAS at 258 K (-15 °C). Cross peaks are labelled according to amino acid spin system, with cross peaks arising from phenylalanine 28 in dark blue and tyrosine 31 in orange.

As seen from both 1D and $2D$ ^{13}C - ^{13}C experiments recorded on BPV E5_{YF}, the aromatic region of the ^{13}C spectrum was unresolvable and appeared as a single broad peak, consisting of resonances from both phenylalanine and tyrosine carbon signals. This greatly hampered any chances of being able to identify any inter-helical cross peaks, due to spectral overlap of resonances, making potential intra and inter-helical cross peaks extremely difficult to differentiate between. Although the Cε and Cζ carbon atoms of tyrosine had distinct chemical shifts (115.7 and 156.1 ppm respectively) that would theoretically allow for any inter-helical couplings between phenylalanine and tyrosine to be observed, the similarity in chemical shift between phenylalanine Cδ 1/2 and tyrosine Cδ 1/2 made it impossible to differentiate between a cross peak arising from couplings between phenylalanine Cδ 1/2 to tyrosine Cε and tyrosine Cδ 1/2 to Cε.

Therefore it was not possible to confidently state that any observed cross peaks arising between the two were solely due to an inter-helical coupling rather than an intra-residue coupling, even when using singly labelled peptides. As such, since the 2D ^{13}C - ^{13}C DARR spectra of BPV E5_{FY} at both short and long mixing times yielded only intra-residue couplings and no cross peaks arising from possible inter-helical couplings, we were unable to find any evidence of inter-helical interactions through dipolar couplings between labelled phenylalanine 28 and tyrosine 31 at the BPV E5 homodimer interface.

5.4 ssNMR characterisation of DMPC/cholesterol lipid membranes

In order to better understand the effect of increasing the cholesterol content of the BPV E5_{FY} sample prepared for ssNMR, specifically why this increase did not yield an improvement in spectral resolution at high or low experimental temperatures, ssNMR samples of the DMPC liposomes were prepared with increasing molar ratios of cholesterol, ranging from no cholesterol to 0.60:1 cholesterol:DMPC (as used in our BPV E5_{FY} sample), for analysis by ^{31}P and ^1H ssNMR.

5.4.1 Characterisation of lipid membranes by static ^{31}P ssNMR

^{31}P is commonly used in NMR studies of lipid membranes due to the high natural abundance of phosphate nuclei, since the majority of phospholipids carry a phosphate moiety in their polar head group. In addition ^{31}P is a good spin $\frac{1}{2}$ NMR nucleus, with 100% natural abundance and a high gyromagnetic ratio making it ideal for studying lipid membranes without having to attach additional non-native reporter groups. Static ^{31}P NMR spectra are generally recorded in order to gain information about the morphology of the lipid bilayer and can be used to determine whether lamellar bilayers are correctly formed or whether non-lamellar, hexagonal (H_{II}) cylindrical structures or isotropic micellar lipid phases have been formed (Cullis and de Kruijff 1979). As such, ^{31}P ssNMR is a great tool for confirmation of the correct formation of lipid bilayers in the presence or absence of cholesterol/peptides, as the phase of the

lipid structures formed can easily be determined using simple two-pulse/decoupling experiments, requiring only a small number of scans and short experimental times.

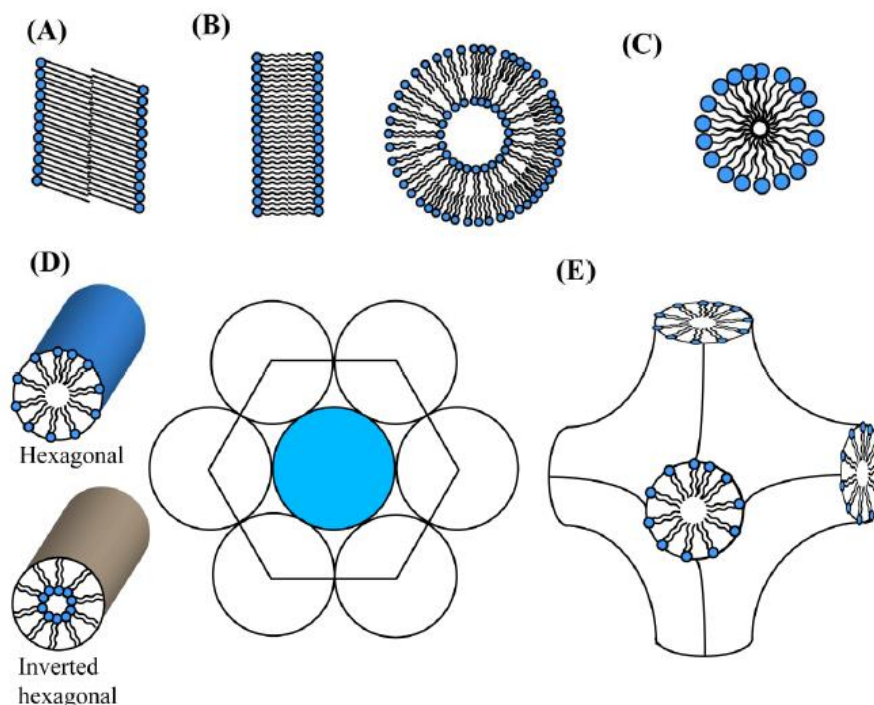


Figure 5.15 Lipid polymorph and phase behaviour

Lipid morphologies that can occur include; A) gel phase B) fluid lamellar phase C) micelle phase D) hexagonal phase E) cubic phase

A number of DMPC lipid samples containing a range of cholesterol concentrations were prepared for analysis by ^{31}P and ^1H ssNMR in a similar manner to GpA and BPV peptide-containing samples. DMPC liposomes were prepared with; no cholesterol (DMPC only), 5% w/w (0.08 molar) cholesterol as used with GpA and BPV E5_{LF} ssNMR samples, 15 w/w (0.30 molar cholesterol) and 30% w/w (0.60 molar) cholesterol as used in our BPV E5_{FY} sample. As shown in **Figure 5.16**, static wide line ^{31}P ssNMR spectra were recorded at a range of temperatures, starting at 25 °C, above the phase transition temperature (T_m) for DMPC, through to -20 °C, which was near the typical temperature at which previous ssNMR experiments on GpA and BPV E5 were recorded.

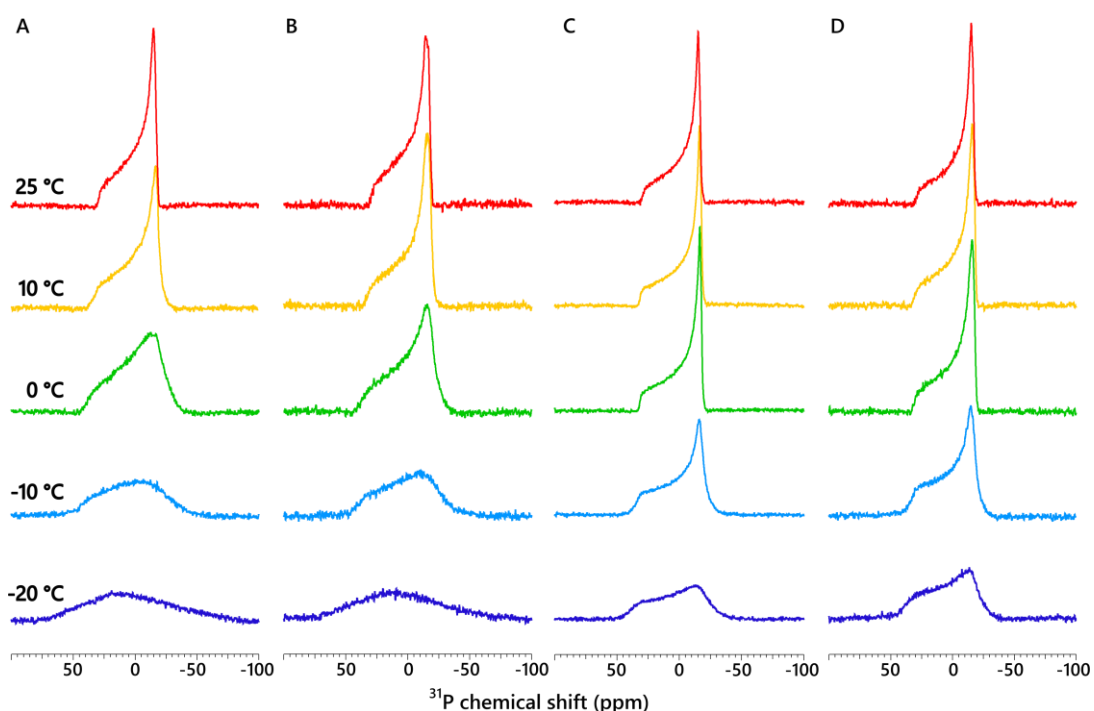


Figure 5.16 1D proton-decoupled static ^{31}P spectra of DMPC liposomes with increasing cholesterol concentration at varying temperature

Static wide line ^{31}P spectra of DMPC liposomes composed of A) pure DMPC B) 0.08 C) 0.30 and D) 0.60 :1 molar ratios of cholesterol were recorded at 600 MHz at temperatures ranging from 25 to $-20\text{ }^{\circ}\text{C}$. Spectra were recorded using a Hahn Echo pulse sequence with a ^1H 90° pulse length of $4\ \mu\text{s}$, an echo delay of $50\ \mu\text{s}$, 80 kHz SPINAL64 proton decoupling and a 5 second recycle delay with 80 kHz proton decoupling during acquisition for 256 co-added transients.

From static ^{31}P experiments recorded (**Figure 5.16**) we saw typical wide line spectra, which are dominated by the large chemical shift anisotropy (CSA) range that showed characteristic broad axially symmetric ^{31}P line shape arising from the phosphate head group of DMPC, with a high field peak and a low field shoulder (width of 50 ppm), indicative of lamellar bilayer structure, with rapid axial rotation in the bilayer resulting in such patterns. As such all DMPC lipid samples prepared were observed to be correctly forming lamellar lipid bilayers, without cholesterol and at both low and high cholesterol concentrations. At $25\text{ }^{\circ}\text{C}$, above the phase transition temperature for DMPC, all lipid samples prepared showed similar axially symmetrical ^{31}P line shapes, with a sharp high field component suggesting that at this temperature all DMPC samples were in the liquid-crystalline phase (L_{α}). In the liquid-crystalline phase there is low conformational

order within the hydrophobic lipid carbon chains and low translational order within the membrane and as such this phase is also referred to as the liquid disordered phase (L_d). As the temperature of the cooling gas was decreased in order to cool the sample below the phase transition temperature, the line shape of the spectra obtained began to change with an increase in broadness of the CSA width, indicative of changes within the polar lipid head groups. The change in line shape at 10 °C however was only observed in pure DMPC and 0.08:1 cholesterol:DMPC samples. In samples with higher cholesterol content no observable change was recorded upon decreasing the temperature to 10 °C. At 0 °C the line shape changed considerably with a loss of the sharp uniaxial peak, becoming much broader and lower in intensity, representing a change in phase from the liquid-crystalline phase to a more ordered gel phase (L_β) in which the lipid hydrocarbon chains are ordered in an all-trans conformation, with translational order preventing diffusion within the bilayer. This phase is also referred to as the solid ordered phase (S_0). At -10 °C, pure DMPC and 0.08:1 cholesterol:DMPC samples showed a distinct loss of line shape with increased broadening of CSA line width and decrease in overall signal intensity. It was at this temperature (-10 °C) that changes in the spectra for 0.3:1 and 0.6:1 cholesterol:DMPC samples were first observed. These spectra looked similar to those of pure DMPC and 0.08:1 cholesterol:DMPC recorded at 0 °C, although the spectra still had a sharp asymmetrical component to the ^{31}P line shape. Decreasing the cooling gas temperature further to -20 °C saw a total broadening out of ^{31}P signal for pure DMPC and 0.08:1 cholesterol:DMPC samples. For samples containing 0.3:1 and 0.6:1 cholesterol:DMPC, the signal although broadened out, still maintained some characteristic features of lamellar bilayer spectra in the L_β phase, similar to those recorded at -10 °C for DMPC only and DMPC + 0.08 molar cholesterol. The results from the ^{31}P spectra recorded suggest that liposomes prepared with no or little (0.08 molar) cholesterol undergo a change in lipid phase from a disordered liquid-crystalline (L_α) to a more ordered gel phase (L_β) between 25 and 10 °C. For samples where the membrane was composed of higher cholesterol content, such as the DMPC + 0.30 and 0.60, there appeared to be no distinction in lipid phase to a more ordered gel phase, as from the ^{31}P spectra obtained the results suggest that there is still axial rotation of the phospholipid head groups, resulting in the presence of a sharp asymmetric peak that is still visible at low temperature.

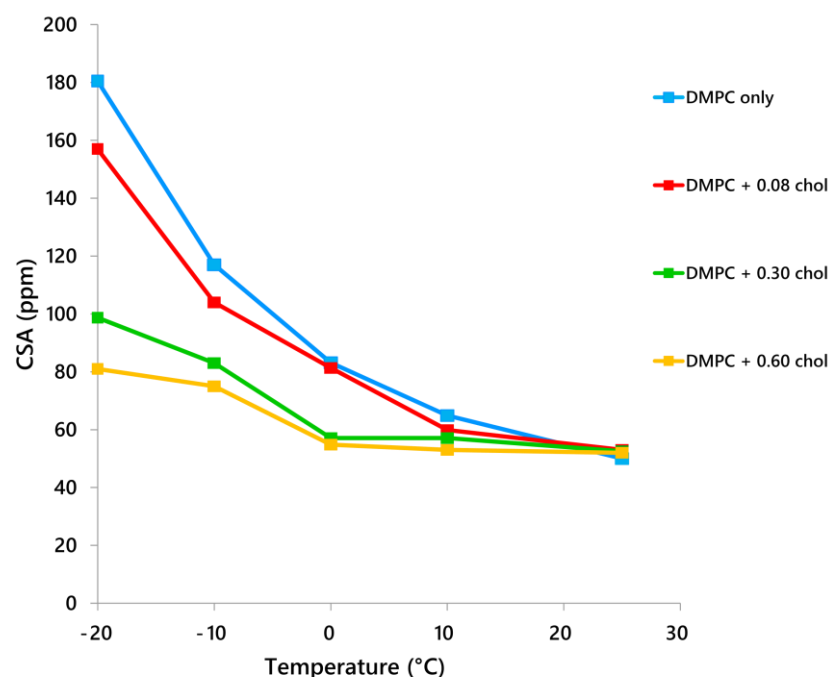


Figure 5.17 Variation of ^{31}P chemical shift anisotropy of DMPC/cholesterol liposomes with temperature

Chemical shift anisotropy component (in ppm) of wide line ^{31}P static NMR spectra of DMPC liposomes with varying cholesterol content as a function of temperature. CSA (measured as a function of the width of the powder pattern) from individual ^{31}P spectra plotted against sample cooling gas temperature.

These results compare well with previously published studies (in the 0–25 °C range) that have concentrated on the effects of cholesterol on phosphocholine bilayers. The ordering effects of cholesterol on the lipid acyl chains at higher temperatures has been previously noted as has the increase in membrane fluidity at higher cholesterol concentrations at low temperatures due to the lack of an gel phase (L_{β}) (Vist and Davis 1990; de Meyer, Benjamini et al. 2010), although in this present study much lower temperatures were used in order to obtain a more accurate representation of the membrane fluidity at experimental temperatures.

5.4.2 Characterisation of lipid membranes by ^1H MAS ssNMR

In order to investigate the effect of cholesterol on the lipid hydrocarbon acyl chains (as opposed to the phospholipid head groups), the same samples that were analysed using ^{31}P NMR were also analysed using ^1H MAS ssNMR (**Figure 5.18**). As with ^{31}P experiments, ^1H spectra were recorded at a range of temperatures from 25 °C - -20 °C.

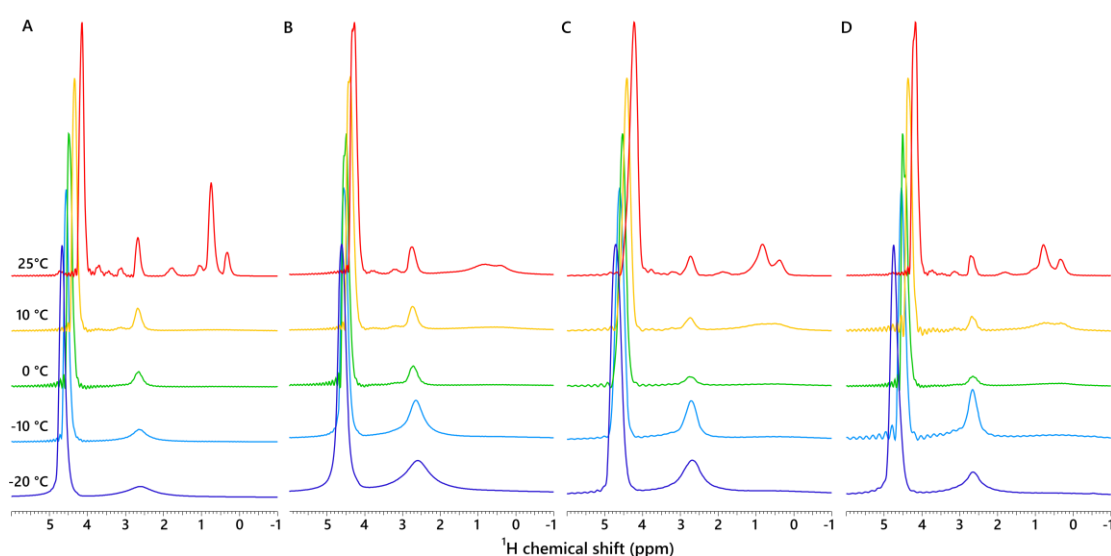


Figure 5.18 1D ^1H spectra of DMPC liposomes with increasing cholesterol concentration at varying temperature

^1H spectra of DMPC liposomes composed of A) DMPC only, and liposomes with the addition of B) 0.08 C) 0.30 and D) 0.60 molar equivalent cholesterol. Spectra were recorded at 600 MHz at temperatures ranging from 25 to -20 °C. A proton one pulse program was used with a 90° pulse length of 2.5 μsec , a 3 second recycle delay and individual spectra acquired for 128 scans each.

^1H MAS NMR can be used to study membrane lipids due to the fast intrinsic reorientation of lipids, which allows restoration of resolution at low spinning speeds (Oldfield, Bowers et al. 1987). A number of proton resonances arising from the hydrophobic hydrocarbon tails and the polar head group were observed, and were assigned (as listed in **Table 5.3**) based upon previously reported ^1H chemical shifts (Schuh, Banerjee et al. 1982; Nomura, Lintuluoto et al. 2011).

Table 5.3 ^1H chemical shift data for DMPC natural abundance lipid resonances

DMPC ^1H chemical shifts (ppm)	
H2	2.39
H3	1.61
H4-13	1.34
H14	0.94
HG1	4.06
HG2	5.35
HG3	3.93
Hα	4.36
Hβ	3.75
Hγ	3.27

From the spectra obtained (**Figure 5.18**) it was again observed that the cholesterol concentration in DMPC lipid membranes affects the membrane structure as a function of temperature. At 25 °C the pure DMPC spectra produced the most and sharpest resonances, with both signals from hydrocarbon tails and the polar head group being observable. The sharp and high intensity peak from the hydrocarbon lipid tails indicated rapid internal motion from the lipid chains when in the $L\alpha$ phase. As the cholesterol composition of the membranes was increased to 0.08:1, the hydrocarbon chain resonances appeared to broaden considerably and decrease in signal intensity, indicating a reduction in motion of the lipid tails within the membrane, whilst resonances from the polar head group remained relatively sharp at this temperature. The same was also observed for liposomes composed of 0.30 and 0.60 molar cholesterol. As the temperature was decreased below the phase transition temperature to 10 °C, the changes to the ^1H spectra recorded were more pronounced. At 10 °C in the spectra recorded for 0.08:1 cholesterol:DMPC, all resonances apart from signals arising from the polar head group broadened out to a point where they were no longer observable, whilst a small broad hump was still visible in the 0.30 and 0.60 molar cholesterol samples. This would indicate that at this temperature the hydrocarbon lipid chains were immobilised whilst the polar head group still showed signs of motion. Unlike the lipid tails, the signal from the head group (2.39 ppm) was still present. Upon a

further decrease in temperature to 0 °C, all resonances from lipid hydrocarbon tails were broadened out and non-observable for all samples. At this temperature the spectra for all samples looked almost identical, with the only resonance observable being that from the polar head group. Upon reducing the cooling gas temperature further to sub-zero temperatures, it was observed that for samples containing cholesterol the resonance arising from the polar head group was still present and appeared as broad peak with reasonable signal intensity, with the resonance in the 0.60 molar cholesterol sample appearing moderately sharp in comparison to other samples. This would suggest that at sub-zero temperatures in the absence of cholesterol, the DMPC lipid molecules in the gel-phase (L_{β}) showed little internal movement of either hydrocarbon chains or the polar head group. With the addition of cholesterol to the lipid bilayers, at sub-zero temperatures, membranes still appeared to have little internal motion and disorder within hydrocarbon tails but, unlike membranes containing pure DMPC, there was still motion within the polar head groups of the bilayer.

These results in combination with the results obtained by ^{31}P analysis of our liposome preparations would indicate that membranes composed solely of DMPC, go through typical phase transitions of L_{α} with both lipid chain and translational disorder above the phase transition temperature. Below the phase transition temperature, both ^{31}P and ^1H spectra indicated reduced motion in the L_{β} phase by line broadening of resonances as the lipid molecules had local order within the lipid chains and phosphate head groups. For liposomes containing cholesterol, the results from ^{31}P spectra indicate that above the T_m , the phosphate head groups behaved similar to membranes lacking cholesterol. Below this temperature however, membranes containing cholesterol showed increased disorder and internal motion of the phosphate head group. From the ^1H recorded, samples prepared with cholesterol above the T_m showed an increase in lipid chain order with reduced motion in the L_{α} phase. Upon cooling, these cholesterol-containing membranes showed similar hydrocarbon chain ordering as membranes lacking cholesterol but, as seen with ^{31}P experiments, ^1H signals from the head group showed disorder and motion.

These results are indicative of liposome samples prepared using cholesterol as being in the liquid-ordered phase (L_o). In the liquid-ordered phase, hydrophobic lipid acyl chains are ordered as in the L_{β} phase, as cholesterol molecules intercalate between

the acyl chains interacting with the hydrocarbon tails resulting in reduced internal motions and disorder between chains. Although there is high conformational order within the lipid acyl chains as in the L_{β} phase (shown in the ^1H spectra), in the L_o phase there is still high translational disorder allowing for lateral movement within the bilayer (as in the L_{α} phase, depicted in **Figure 5.18** and seen from ^{31}P spectra). The addition of cholesterol to lipid bilayers has been shown to expand the cross sectional head group area of lipid molecules, thereby increasing the head group mobility (Shepherd and Buldt 1979). ^{31}P spectra given in **Figure 5.16** show that for cholesterol containing membranes below the T_m , the lamellar bilayers appeared to have increased disorder within the head groups, as increasing cholesterol into lipid membranes below the T_m reduces the amount of L_{β} phase and increased the amount in L_{α} and L_o phase. It has previously been reported that incorporation of cholesterol into sphingomyelin (SM) lipids gradually disrupts bilayers in the L_{β} phase, and can even eliminate the L_{β} phase when the molar concentration of cholesterol was above 15% (Lund-Katz, Laboda et al. 1988) with the L_o transition temperature estimated to be completed below 10 – 15 °C for DMPC bilayers with 30% molar cholesterol (McMullen, Lewis et al. 1993).

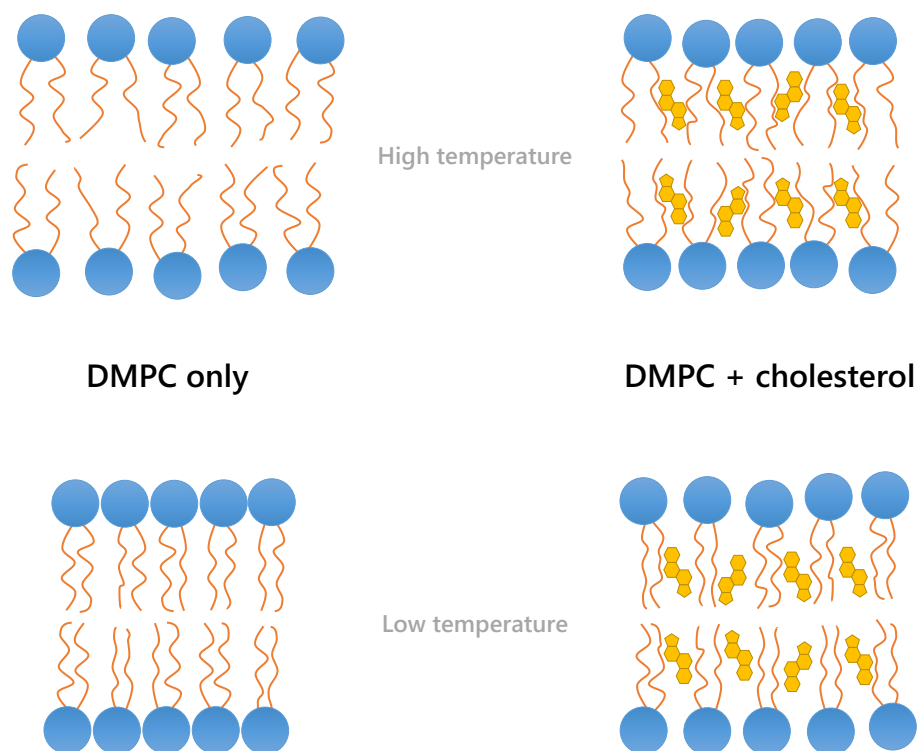


Figure 5.19 Cartoon of cholesterol intercalating between lipid molecules at high and low temperatures

Diagrammatic representation of the effect of cholesterol on DMPC lipid bilayers at high and low temperatures. The left side of the figure shows DMPC bilayers only whilst on the right the effect of the addition of cholesterol on lamellar lipid bilayers. At temperatures above the DMPC phase transition T_m , DMPC only bilayers exist in the liquid crystalline (L_α) phase, where there is lipid chain and translation disorder within the membrane. Below the T_m for DMPC only, lamellar bilayers exist in the gel-phase (L_β) where lipid acyl chains are ordered and there is little translation diffusion within the bilayer. With the addition of cholesterol, above the T_m DMPC bilayers exist in the liquid ordered phase (L_α) where lipid acyl chains are ordered but there translation diffusion within the bilayer still occurs. Below the T_m the addition of cholesterol prevents the phosphocholine head groups from packing together whilst acyl chains are still ordered as cholesterol intercalates between them.

Therefore increasing the cholesterol content of DMPC membranes in the BPV E5_{FY} sample and acquiring spectra at $-15\text{ }^\circ\text{C}$ had a detrimental effect on bilayer rigidity. In our hands, higher experimental temperatures produced poor spectra containing only resonances from DMPC lipid, and peptide resonances were only observed at low temperature. At low temperature, rather than making the bilayers more rigid, the addition of cholesterol had the opposite effect, yielding bilayers in an L_α phase rather than the L_β phase, and this increased motion may explain why aromatic peaks were

broadened, the signal to noise ratio was lower, and no inter-helical cross peaks were observed between labelled phenylalanine and tyrosine.

5.5 Summary

In this chapter, ^{13}C ssNMR for the viral oncogenic membrane protein BPV E5 that had been reconstituted into DMPC lipid bilayers with the addition of cholesterol at varying compositions, as well as the effect of the addition of cholesterol on lipid bilayers as analysed by ^{31}P and ^1H ssNMR was presented. Although the BPV E5 protein has been well characterised using a number of biochemical and biophysical techniques in order to gain information about the method of interaction with its cellular target PDGF β R, there currently exists no 3D structural model. Although a number of studies have identified key amino acid residues within the sequence of E5 that are required for correct homodimerisation and interaction with PDGF β R, such as glutamine 17 and aspartate 33 cysteine 37 and 39 in the C-terminal portion that form disulphide bonds between two helices for homodimerisation. Only molecular models produced by computational analysis of BPV E5 exist, although attempts have been made to study E5 by solution NMR (Windisch, Hoffmann et al. 2010; King, Oates et al. 2011)

Having produced an energy minimised predicted structure of the BPV E5 homodimer using CHI, that was in close agreement with the current understanding of the dimer from many other studies, this model was then used to identify suitable amino acids to isotopically U- $^{13}\text{C}/^{15}\text{N}$ label for ssNMR analysis. Amino acids that had short inter-helical but long ($> 5 \text{ \AA}$) intra-helical distances were identified and their suitability for labelling explored. This presented a challenge due to the high leucine content of the BPV E5 peptide, with a large number of leucine side chains predicted to be at the dimer interface, reducing the number of labelling sites available to just a few, as multiple leucines could not be labelled due to spectral overlap. Glutamine 17 that has been shown to play a key role in dimerisation was not a viable site for labelling as economically the cost of obtaining U- $^{13}\text{C}/^{15}\text{N}$ labelled glutamine for incorporation into peptides was too high. As such three amino acids at the dimer interface that fit the criteria were identified as leucine 24, phenylalanine 28 and tyrosine 31. Using these three

amino acids the interaction between leucine and phenylalanine and phenylalanine and tyrosine couplings could be probed.

Using isotopically labelled leucine and phenylalanine a BPV E5_{LF} ssNMR sample was prepared in a similar manner to singly labelled GpA. Using 1D ¹³C analysis of BPV E5 in DMPC/cholesterol liposomes, resonances and chemical shift values of labelled amino acids were assigned. Following which 2D ¹³C - ¹³C DARR spectra were acquired in order to observe any inter-residue, inter-helical interactions between the two labelled amino acids at the BPV E5 homodimer interface. At short (20 ms) mixing times, only intra-residue cross peaks within leucine and phenylalanine were observed. Increasing the mixing time to 400 ms, the only additional couplings observed were intra-residue couplings between phenylalanine carbon atoms that were not observed at short mixing time or had higher signal intensity. For BPV E5_{LF} no inter-helical couplings at long mixing time between labelled leucine and phenylalanine were observed, suggesting that either they are not close together in space, as they were predicted to be from the CHI molecular model of the dimer interface or due to experimental reasons we were unable to observe any such inter-helical coupling. Whereas for singly labelled GpA, we were able to run 2D ¹³C-¹³C DARR experiments for longer times, due to spectrometer availability, the number of scans recorded in 2D experiments for BPV E5_{LF} were not as high. In addition all previous experiments on GpA had been recorded at 500 MHz, for BPV E5 due to equipment down-time we had to switch to recording spectra at 600 MHz, which appeared to give lower signal to noise ratios than at lower field. 2D ¹³C-¹³C DARR spectra were recorded for twice the number of scans at 600 MHz but again due to low signal to noise no improvement in data was seen and no additional inter-helical cross peaks were observed.

With BPV E5_{LF} it was also observed from both 1D and 2D ¹³C spectra that resonances arising from phenylalanine aromatic ring carbons were unresolvable appearing as an unassignable single broad peak in the spectra recorded. Knowing that the next sample to be prepared contained phenylalanine and tyrosine, both of which had aromatic ring carbons in the structure of their side chains and knowing that the inter-helical interactions predicted to occur between the two amino acids was between the two aromatic rings, a method to improve the resolution of these resonances was sought.

As such in the next sample prepared using labelled phenylalanine and tyrosine (BPV E5_{FY}) in an attempt to improve the resolution of ¹³C spectra recorded and to make the lipid bilayers in the samples prepared more biologically relevant, the cholesterol composition of samples prepared for ssNMR analysis was increased greatly from 5% w/w (0.08 mol ratio) to 30% w/w (0.60 mol ratio) which was a compromise value between that of the cholesterol level found in the Golgi apparatus and plasma membrane to which BPV E5 locates *in vitro*. Increasing the cholesterol composition of the bilayer was thought to rigidify the membrane and allow for spectra to be recorded at more physiological temperatures. From 1D ¹³C experiments it was seen that running at higher temperatures saw no improvement in spectra resolution as at this temperature only resonances for DMPC lipid were observed, decreasing the temperature saw an improvement in signal but as with BPV E5_{LF}, resonances arising from the aromatic ring carbons were still unresolvable. Having optimised experimental parameters such as temperature and CP contact time, short and long mixing time 2D ¹³C - ¹³C DARR experiments were recorded on BPV E5_{FY}. Again for this sample at long mixing times only intra-residue cross peaks were observed, with no inter-helical cross peaks between phenylalanine and tyrosine. After recording BPV E5_{FY} spectra it was concluded that using both phenylalanine and tyrosine in one sample was a poor choice of pairings, not only due to the aromatic ring carbons giving rise to broad peaks, but also due to very similar chemical shift values between phenylalanine and tyrosine, this pairing of amino acids would not give rise to cross peaks that could be fully attributed to inter-helical couplings without ruling out the possibility of observing intra-residue couplings due to chemical shift values of phenylalanine C δ 1/2 and tyrosine C δ 1/2 being so similar.

The effect of increasing cholesterol content of DMPC lipid membrane was investigated further using ³¹P and ¹H ssNMR in order to understand why no improvement in spectral resolution was observed with BPV E5_{FY}. From the results obtained using ³¹P and ¹H spectra obtained it was observed from static ³¹P spectra that the addition of cholesterol to DMPC liposomes had no detrimental effect on the correct formation of lamellar bilayers. At high temperature above the T_m the addition of cholesterol to bilayers resulted in membranes being in the liquid ordered phase (L_o) whereas at low temperature below the T_m the addition of cholesterol resulted in loss of the more ordered gel-phase (L _{β}) as ¹H data indicated the reduced ordering of lipid acyl

chains and higher mobility of protons in the inter-facial and polar head group regions of DMPC molecules. The data presented here for DMPC bilayers in the presence of cholesterol can be used to gain a better understanding of the effect of cholesterol on DMPC membranes at sub-zero temperatures as although well studied, typically literature surrounding the effect of cholesterol on lamellar bilayers concerns work conducted above 0 ° C.

5.6 References

- Abdine, A., M. A. Verhoeven, K. H. Park, A. Ghazi, E. Guittet, C. Berrier, C. Van Heijenoort and D. E. Warschawski (2010). "Structural study of the membrane protein MscL using cell-free expression and solid-state NMR." *J Magn Reson* **204**(1): 155-159.
- Bittman, R., S. Clejan, S. Lund-Katz and M. C. Phillips (1984). "Influence of cholesterol on bilayers of ester- and ether-linked phospholipids. Permeability and ¹³C-nuclear magnetic resonance measurements." *Biochim Biophys Acta* **772**(2): 117-126.
- Burkhardt, A., M. Willingham, C. Gay, K. Jeang and R. Schlegel (1989). "The E5 oncoprotein of bovine papillomavirus is oriented asymmetrically in Golgi and plasma membranes." *Virology* **170**(1): 334-339.
- Cady, S. D., T. V. Mishanina and M. Hong (2009). "Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding." *J Mol Biol* **385**(4): 1127-1141.
- Cullis, P. R. and B. de Kruijff (1979). "Lipid polymorphism and the functional roles of lipids in biological membranes." *Biochim Biophys Acta* **559**(4): 399-420.
- Hiller, M., L. Krabben, K. Vinothkumar, F. Castellani, B. van Rossum, W. Kühlbrandt and H. Oschkinat (2005). "Solid-state magic-angle spinning NMR of outer-membrane protein G from Escherichia coli." *Chembiochem* **6**(9): 1679-1684.
- Horwitz, B. H., D. L. Weinstat and D. DiMaio (1989). "Transforming activity of a 16-amino-acid segment of the bovine papillomavirus E5 protein linked to random sequences of hydrophobic amino acids." *J Virol* **63**(11): 4515-4519.
- King, G., J. Oates, D. Patel, H. A. van den Berg and A. M. Dixon (2011). "Towards a structural understanding of the smallest known oncoprotein: investigation of the bovine papillomavirus E5 protein using solution-state NMR." *Biochim Biophys Acta* **1808**(6): 1493-1501.
- Kolodziejewski, W. and J. Klinowski (2002). "Kinetics of cross-polarization in solid-state NMR: a guide for chemists." *Chem Rev* **102**(3): 613-628.
- Lund-Katz, S., H. M. Laboda, L. R. McLean and M. C. Phillips (1988). "Influence of molecular packing and phospholipid type on rates of cholesterol exchange." *Biochemistry* **27**(9): 3416-3423.
- Luo, W., S. D. Cady and M. Hong (2009). "Immobilization of the influenza A M2 transmembrane peptide in virus envelope-mimetic lipid membranes: a solid-state NMR investigation." *Biochemistry* **48**(27): 6361-6368.

- Mattoon, D., K. Gupta, J. Doyon, P. Loll and D. DiMaio (2001). "Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein." *Oncogene* **20**(29): 3824-3834.
- McMullen, T. P., R. N. Lewis and R. N. McElhaney (1993). "Differential scanning calorimetric study of the effect of cholesterol on the thermotropic phase behavior of a homologous series of linear saturated phosphatidylcholines." *Biochemistry* **32**(2): 516-522.
- Meyer, A. N., Y. F. Xu, M. K. Webster, A. E. Smith and D. J. Donoghue (1994). "Cellular transformation by a transmembrane peptide: structural requirements for the bovine papillomavirus E5 oncoprotein." *Proc Natl Acad Sci U S A* **91**(11): 4634-4638.
- Nomura, K., M. Lintuluoto and K. Morigaki (2011). "Hydration and temperature dependence of ¹³C and ¹H NMR spectra of the DMPC phospholipid membrane and complete resonance assignment of its crystalline state." *J Phys Chem B* **115**(50): 14991-15001.
- Oldfield, E., J. L. Bowers and J. Forbes (1987). "High-resolution proton and carbon-13 NMR of membranes: why sonicate?" *Biochemistry* **26**(22): 6919-6923.
- Schlegel, R., M. Wade-Glass, M. Rabson and Y. Yang (1986). "The E5 transforming gene of bovine papillomavirus encodes a small, hydrophobic polypeptide." *Science* **233**(4762): 464-467.
- Schuh, J. R., U. Banerjee, L. Muller and S. I. Chan (1982). "The phospholipid packing arrangement in small bilayer vesicles as revealed by proton magnetic resonance studies at 500 MHz." *Biochim Biophys Acta* **687**(2): 219-225.
- Shepherd, J. C. and G. Buldt (1979). "The influence of cholesterol on head group mobility in phospholipid membranes." *Biochim Biophys Acta* **558**(1): 41-47.
- Urbina, J. A., S. Pekerar, H. B. Le, J. Patterson, B. Montez and E. Oldfield (1995). "Molecular order and dynamics of phosphatidylcholine bilayer membranes in the presence of cholesterol, ergosterol and lanosterol: a comparative study using ²H-, ¹³C- and ³¹P-NMR spectroscopy." *Biochim Biophys Acta* **1238**(2): 163-176.
- van Meer, G. (1998). "Lipids of the Golgi membrane." *Trends Cell Biol* **8**(1): 29-33.
- van Meer, G., D. R. Voelker and G. W. Feigenson (2008). "Membrane lipids: where they are and how they behave." *Nat Rev Mol Cell Biol* **9**(2): 112-124.
- Venuti, A., F. Paolini, L. Nasir, A. Corteggio, S. Roperto, M. S. Campo and G. Borzacchiello (2011). "Papillomavirus E5: the smallest oncoprotein with many functions." *Mol Cancer* **10**: 140.
- Vist, M. R. and J. H. Davis (1990). "Phase equilibria of cholesterol/dipalmitoylphosphatidylcholine mixtures: ²H nuclear magnetic resonance and differential scanning calorimetry." *Biochemistry* **29**(2): 451-464.
- Windisch, D., S. Hoffmann, S. Afonin, S. Vollmer, S. Benamira, B. Langer, J. Burck, C. Muhle-Goll and A. S. Ulrich (2010). "Structural role of the conserved cysteines in the dimerization of the viral transmembrane oncoprotein E5." *Biophys J* **99**(6): 1764-1772.

6

SOLUTION NMR ANALYSIS OF BPV E5

6.1 Introduction

In this chapter, solution-state NMR analyses of the ^{15}N isotopically labelled BPV E5_{V2} peptide are presented. The first section of this final results chapter details the use of both one-dimensional ^1H and two-dimensional $^{15}\text{N} - ^1\text{H}$ heteronuclear solution NMR experiments to study the BPV E5_{V2} peptide in both organic solvent (TFE) and reconstituted into DMPC/DHPC bicelles. The BPV E5_{V2} was specifically designed for solution NMR studies, with multiple N^{15} labels at key sites at the homodimer interface and control sites away from the interface. For bicelle samples, the effect of changing the short:long chain lipid ratio on the spectra obtained was investigated using ^1H , ^{15}N and ^{31}P NMR. ^{15}N -edited three-dimensional NOESY and TOSCY experiments were also acquired in an attempt to aid the assignment of resonances observed from BPV E5_{V2} peptide reconstituted in bicelles. Bicelle samples prepared for solution NMR analysis were also analysed by dynamic light scattering, to better understand how the size of the peptide/bicelle complex affected the signals obtained in the NMR spectra. Finally, preliminary two- and three dimensional solution NMR results for bicelle-solubilised $^{13}\text{C}/^{15}\text{N}$ leucine and phenylalanine labelled E5 peptides (originally designed for ssNMR studies) are presented.

Whilst ssNMR is a powerful technique when applied to the study of membrane proteins in a lipid environment, allowing for the analysis of membrane protein structure in a "native-like" environment, it would be impossible to study membrane proteins embedded in liposomes by solution NMR due to the extremely large size of the protein liposome complex (Watts and Spooner 1991). As solution NMR techniques have fast tumbling requirements in order to average out the CSA effects that cause line

broadening in the spectra obtained, liposomes tumble much too slowly on the NMR timescale (Watts and Spooner 1991). Therefore suitable membrane mimetic systems that allow for solubilisation of hydrophobic membrane proteins, while also satisfying the fast tumbling requirement of solution NMR and maintaining correct biophysical function have been sought.

First introduced by Sanders et al. (Sanders and Prestegard 1990; Sanders and Schwonek 1992), bicelles composed of DMPC as the long chain phospholipid component and 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) as detergent are the most commonly used and most characterised bicellar systems used for the study of single and multi-span membrane proteins by solution NMR (De Angelis, Howell et al. 2006; Bocharov, Pustovalova et al. 2007; Bocharov, Mineev et al. 2008). These binary bilayered mixed micelles can also be doped with phospholipids with different polar head groups such as 1,2-dihexanoyl-*sn*-glycero-3-phosphoglycerol (DMPG), which carries a negative charge on the head group that can be used to alter the charge properties of DMPC/DHPC bicelles as well as improve the stability of bicelles (Struppe, Komives et al. 1998; Struppe, Whiles et al. 2000). This modification allows samples to be used for longer periods of time and for more complex 3D NMR experiments to be recorded.

By varying the ratio of long to short chain phospholipids (referred to as the q ratio) (**Equation 9**), two forms of bicelle can be prepared. At temperatures above the T_m of DMPC (25 °C), q ratios above 2.5 and at lipid concentrations of 3-60% w/v, large bicelles that spontaneously align with the magnetic field (B_0) are formed (Sanders and Schwonek 1992). These slower tumbling, magnetically alignable bicelles can be used to orient embedded membrane proteins in solution (De Angelis, Nevzorov et al. 2004; Park, Prytulla et al. 2006; Triba, Zoonens et al. 2006) and can also be used to measure residue dipolar couplings (RDCs) that can aid protein structure determination through the measurement of angles between two atoms (Bax and Tjandra 1997). Larger bicelles have also been used for ssNMR analysis of embedded membrane proteins (De Angelis, Howell et al. 2006) making it possible to collect both solution and ssNMR data in similar membrane mimetic environments.

$$q = \frac{\text{Total Molarity of DMPC}}{\text{Total Molarity of DHPC} - 0.015 \text{ M}} \quad (9)$$

Whilst larger alignable bicelle lipid aggregates can be formed by using higher q ratios, using q ratios less than 0.5 results in the production of small discoid isotropic bicelles that undergo fast tumbling molecular motion in the magnetic field, which allows for the structure determination of membrane proteins embedded within the bicelle by solution NMR techniques. Although smaller in diameter (~8 – 10 nm) (Vold, Prosser et al. 1997), isotropic bicelles have been shown to have a similar structure to that of larger magnetically alignable bicelles, consisting of a central planar bilayer region capped by short chain phospholipids (Andersson and Maler 2005), whilst maintaining a similar bilayer thickness of ~ 4 nm. Though larger than detergent micelles, the smaller overall diameter of isotropic bicelles in comparison to larger alignable bicelles allows them to tumble relatively fast on the NMR time scale and can be used to obtain high resolution NMR spectra (Vold, Prosser et al. 1997; Luchette, Vetman et al. 2001). Although the spectra obtained in lipid bicelles are generally broader than those recorded in detergent micelles, the larger diameter of isotropic bicelles offers an advantage over detergent micelles as the increased diameter leads to a reduction in curvature stress that can affect protein structure and in some cases the functionality of the embedded membrane protein. An example of membrane protein functionality being retained when reconstituted in bicelles can be found with the enzyme diacylglycerol kinase. Diacylglycerol kinase has been shown to retain correct enzymatic activity when embedded in isotropic bicelles whereas enzymatic activity is lost when reconstituted into detergent micelles (Sanders and Landis 1995).

Small isotropic bicelles have been used to elucidate the structure of a number of small membrane proteins or TM domains, such as the dimeric TM domain of the growth factor receptor ErbB2 (Bocharov, Mineev et al. 2008) and the apoptotic protein BNip3 (Bocharov, Pustovalova et al. 2007). As such, fast tumbling isotropic bicelles present a more attractive membrane mimetic system than that of the more commonly used detergent micelle system.

6.2 Solution NMR analysis of BPV E5_{V2}

By preparing DMPC/DHPC lipid bicelle samples, it was hoped that solution NMR spectrometry could be used to study the BPV E5 homodimer in a “native-like” lipid environment, more representative than that of the detergent micelle environment (King, Oates et al. 2011) or the organic solvent trifluoroethanol (TFE) (Windisch, Hoffmann et al. 2010; King, Oates et al. 2011) in which previous attempts to characterise the BPV E5 protein have been presented. For solution NMR analysis of BPV E5, the sequence of the synthetic peptide incorporating multiple ¹⁵N isotopic labels designed by Dr Gavin King (*previously of the Dixon Group, Department of Chemistry, the University of Warwick*) for the study of BPV E5 in TFE and in SDS micelles, was used and is shown in **Figure 6.1**. Due to the spectral overlap that can occur due to the broadness of resonances observed in the solid state, peptides that were designed for ssNMR studies only incorporated a single ¹³C/¹⁵N labelled amino within each peptide (see **Table 2.2** in **Chapter 2.3**). For solution NMR, where experimental line widths obtained are much narrower, multiple amino acid sites containing ¹⁵N labels were appropriate.



Figure 6.1 Peptide sequence of ¹⁵N labelled BPV E5_{V2} peptide

Transmembrane region of synthesised BPV E5 peptide with ¹⁵N isotopically labelled amino acids indicated in blue. Additional end terminal lysine residues were incorporated to improve peptide solubility.

The BPV E5_{V2} peptide was synthesised with ¹⁵N labelled residues incorporated at multiple sites within the peptide at Gly 11, Val 13, Ala 14, Leu 19, Leu 24 and Val 30. A molecular model of the predicted E5 transmembrane domain, produced using the CHI software, with all ¹⁵N labelled sites highlighted is shown in **Figure 6.2**. In this model, Val 13, Ala 14 and Leu 24 are predicted to be at the homodimer interface while Gly 11, Leu 19 and Val 30 are predicted to be outside of the homodimer interface. This model of the E5 homodimer is in agreement with current accepted models of the E5 homodimer

determined using mutagenesis *in vivo* (Mattoon, Gupta et al. 2001). Therefore, these amino acid residues within the BPV E5 peptide were chosen as key sites to label, as sites predicted to interact and as non-interacting controls.

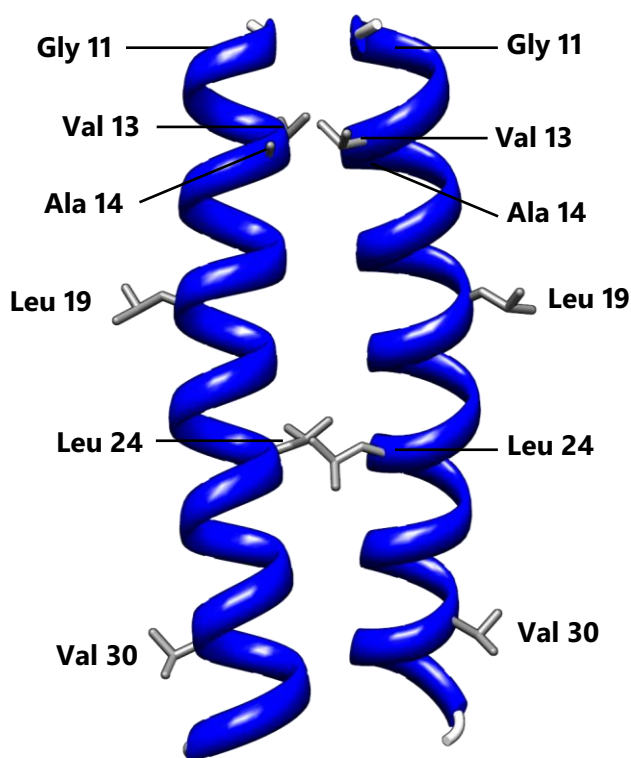


Figure 6.2 Molecular model of BPV E5_{V2} homodimer with ¹⁵N labelled amino acids indicated

Predicted energy minimised molecular model of the BPV E5 homodimer generated using the CHI molecular modelling software. Amino acids predicted to be at the homodimer interface (Mattoon, Gupta et al. 2001) and those predicted to point outwards away from the interfacial region are shown in stick form.

6.2.1 Solution NMR analysis of BPV E5_{V2} in TFE

In preparation for the characterisation of BPV E5_{V2} peptide in bicelles, the peptide was first analysed by solubilisation in the organic solvent trifluoroethanol (TFE) as it was hoped that the spectra recorded in TFE would help with the assignment of resonances observed when reconstituted into DMPC/DHPC lipid bicelles and would also

allow us to compare the spectra obtained with that of previously published data recorded in TFE and SDS micelles (King, Oates et al. 2011). Initial water-suppressed 1D ^1H NMR experiments (Hwang and Shaka 1995) were recorded on the BPV E5_{v2} peptide solubilised in 80% deuterated TFE (TFE-d₃) and 20% H₂O, as shown in **Figure 6.3**. A number of well resolved resonances with good signal to noise were observed, with those in the amide and aliphatic region of the spectrum arising from BPV E5 peptide.

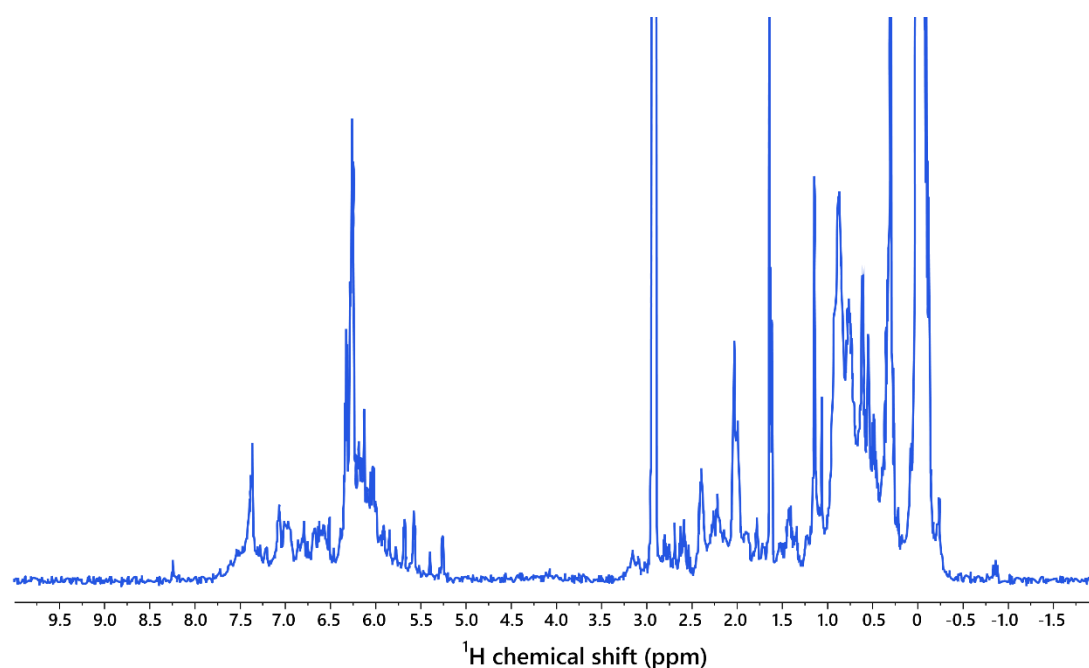


Figure 6.3 1D ^1H spectrum of BPV E5_{v2} in deuterated TFE

1D ^1H spectrum of BPV E5_{v2} peptide in 80:20 TFE-d₃:H₂O. Spectrum recorded on 700 MHz spectrometer (Bruker, Avance II) at 40 °C (313 K) with a 7.71 μsec ^1H 90° pulse length for 16 co-added transients.

Following 1D analysis, a 2D $^{15}\text{N} - ^1\text{H}$ heteronuclear single quantum coherence (HSQC) (Schleucher, Schwendinger et al. 1994) spectrum was recorded. In the HSQC experiment, magnetisation is transferred to a second nucleus, which can either be ^{15}N or ^{13}C via an INEPT transfer (Insensitive Nuclei Enhanced by Polarisation Transfer). Following a t_1 delay, magnetisation is transferred back to the proton and the NMR signal recorded. In a HSQC experiment, the t_1 delay is incremented whilst a series of experiments is recorded in order to produce a 2D spectrum where the ^1H chemical shift is recorded in the direct dimension and the ^{15}N or ^{13}C chemical shift recorded in the

indirect dimension. In a ^{15}N - ^1H HSQC spectrum each amide proton attached to a ^{15}N labelled nitrogen in the peptide bond gives rise to an observable cross peak (with the exception of proline which has no amide proton). The spectrum obtained from a ^{15}N HSQC experiment is often referred to as a "fingerprint" of a protein, as each protein will give rise to a unique pattern of resonances, with a well dispersed spectrum indicative of a folded protein, whereas unfolded proteins form a narrow line down the centre of the spectrum. As such, a 2D ^{15}N - ^1H HSQC spectrum of the BPV E5_{V2} peptide was recorded in 80:20 TFE-d₃:H₂O (**Figure 6.4**) allowing for ^{15}N labelled amide groups within the BPV E5_{V2} peptide to be observed.

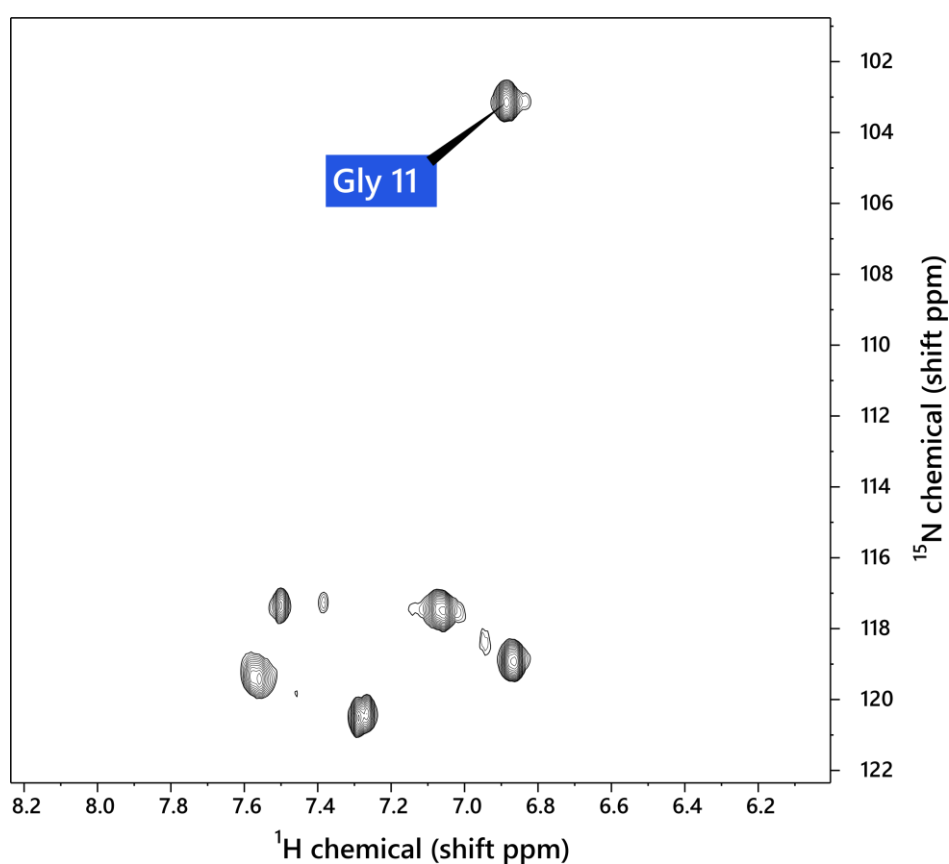


Figure 6.4 2D ^{15}N - ^1H HSQC spectrum of BPV E5_{V2} in deuterated TFE

2D ^{15}N - ^1H HSQC spectrum of BPV E5_{V2} peptide in 80:20 TFE-d₃:H₂O. Spectrum recorded on a 700 MHz spectrometer (Bruker, Avance II) at 40 °C (313 K) with a 7.71 μsec ^1H 90° pulse length with 4K planes in F₂ and 64 planes in F₁ for 16 co-added transients.

A number of well dispersed resonances were observed in the $^{15}\text{N} - ^1\text{H}$ HSQC spectrum obtained, with the number of resonances matching that of the expected number of ^{15}N labels in BPV E5_{v2} peptide. Six major resonances were observed, indicating that the incorporation of ^{15}N labels into the BPV E5_{v2} peptide was successful, and that the majority of peptides were forming a single species in TFE with an alpha-helical secondary structure. The peptide structure in TFE was predicted to be a monomeric helix as TFE promotes the formation of hydrogen bonding and secondary structure but also disrupts oligomerisation via helix-helix interactions (Luo and Baldwin 1997). The presence of a number of minor resonances, of much lower signal intensity, was also observed suggesting the possible formation of a minor species in a second chemical environment. The $^{15}\text{N} - ^1\text{H}$ HSQC spectrum recorded matched well with that of BPV E5_{v2} peptide previously reported in 80:20 TFE-d₃:H₂O (King, Oates et al. 2011). The only resonance observed that was tentatively assignable based upon previously published chemical shift data was glycine 11 due to the relatively low ^{15}N chemical shift (103.27 ppm) in comparison to the other labelled amino acids within the peptide.

6.3 Solution NMR analysis of BPV E5_{v2} in bicelles

Having recorded 1D ^1H and 2D $^{15}\text{N} - ^1\text{H}$ spectra of BPV E5_{v2} in TFE, the BPV E5_{v2} peptide was reconstituted into DMPC/DHPC bicelles that represented a more native-like environment than that of TFE. DMPC/DHPC bicelle samples were prepared using BPV E5_{v2} (as described in more detail in **Chapter 2.10.1**). As is the case with ssNMR membrane protein samples, no single sample preparation is universally applicable to the production of successfully reconstituted protein into bicelles that can be used to obtain good, high resolution NMR spectra. In order to find the best sample preparation method a number of bicelle samples were prepared for analysis by solution NMR. When preparing bicelle samples, a number of important variables were taken into account such as the lipid to protein ratio, the percentage total amphiphile (the total percentage weight of DMPC + DHPC in solution) and q . As described earlier in this chapter, the q is the ratio of long to short chain lipid in the sample, where typically DMPC is used as the long chain lipid and DHPC the short chain lipid. As DHPC can be classified as a detergent, it has an associated critical micelle concentration (CMC). The CMC value is the

concentration above which a detergent forms micelles, and in the case of DHPC the CMC value is quite large (15 mM). As such, when preparing bicelle samples with different q values, it is important to take the CMC of DHPC into account in order to account for the free detergent population below the CMC within the solution that is not in bicelle form. This is accomplished by subtracting 15 mM DHPC when calculating q , as this free detergent population also has an effect on the total amphiphile value of the solution. There have been a number of studies that have reported an "apparent q " ((Bocharov, Pustovalova et al. 2007; Bocharov, Mineev et al. 2008; Morrison and Henzler-Wildman 2012) where the free detergent population below the DHPC CMC value has not been accounted for. When preparing bicelle samples for the reconstitution of BPV E5_{V2} peptides, a number of sample variables were trialled in order to identify the ideal conditions to obtain high resolution solution NMR spectra, whilst also maintaining the correct secondary structure and oligomeric state. For DMPC/DHPC bicelles, we tested the effects of the q ratio, lipid to protein ratio (LPR), total percentage amphiphile in solution and temperature on the quality of the solution NMR data.

An initial solution NMR sample for BPV E5_{V2} peptide reconstituted into DMPC/DHPC bicelles was prepared with a q value of 0.5, a lipid to protein ratio of 80:1 w/w (220:1 molar) and a percentage total amphiphile value of 10% in 20 mM TRIS buffer (as described in more detail in **Chapter 2.10.1**). To assess the quality of the sample, a 2D ¹⁵N – ¹H HSQC spectrum was recorded (**Figure 6.5**). Whereas previously reported spectra obtained for the BPV E5_{V2} peptide in SDS micelles showed a presence of two oligomeric species (monomer/dimer) dependent upon SDS micelle concentration (King, Oates et al. 2011), in the 2D ¹⁵N – ¹H HSQC spectrum recorded in $q=0.5$ bicelles only six resonances were observed for the six ¹⁵N labelled amino acids within the E5_{V2} peptide. The cross peak tentatively assigned as arising from glycine 11 was of a much lower signal intensity compared to the other five. 2D ¹⁵N – ¹H HSQC spectra were recorded at a range of temperatures, from 25 °C (the T_m temperature for DMPC) to 50 °C. By extracting a 1D plane from each 2D experiment signal to noise was compared (**Figure 6.6**).

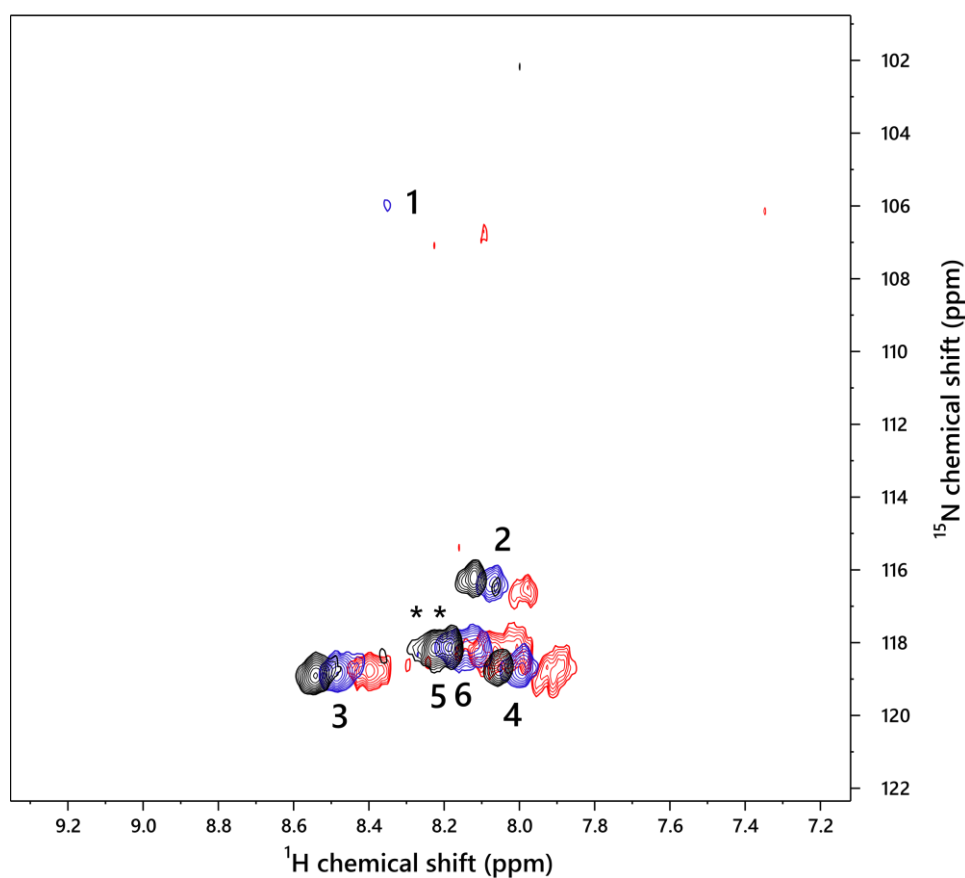


Figure 6.5 2D ^{15}N – ^1H HSQC spectra of BPV E5_{v2} in DMPC/DHPC bicelles at increasing temperature

2D ^{15}N – ^1H HSQC spectra of BPV E5_{v2} peptide reconstituted into $q = 0.5$ DMPC/DHPC bicelles. Spectra recorded on a 700 MHz spectrometer (Bruker, Avance II) at 25 (red), 40 (blue) and 50 °C (black) (298, 313 and 323 K) with a 7.84 μsec ^1H 90° pulse length with 4K planes in F_2 and 64 planes in F_1 for 16 co-added transients.

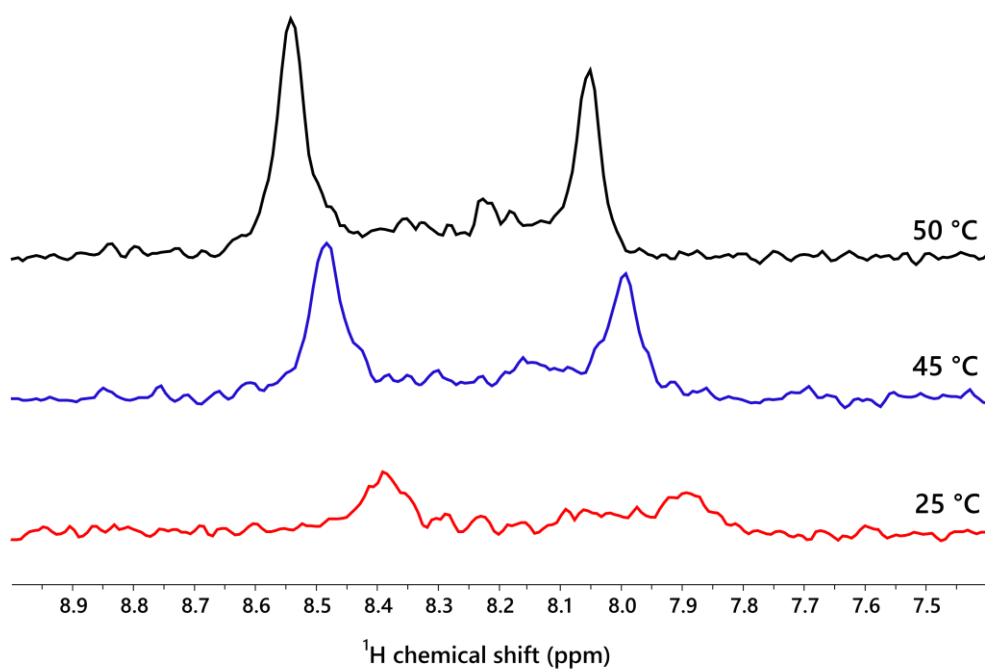


Figure 6.6 Extracted planes from 2D ^{15}N - ^1H HSQC spectra at increasing temperatures

1D row extracted from 2D HSQC experiments at each of the temperatures recorded, at 50 °C, signal and line width were much improved over that of those resonances recorded at 25 °C.

A temperature of 50 °C resulted in 2D ^{15}N – ^1H HSQC spectra with the best line shape and highest signal to noise ratio of all 2D experiments recorded, this was even more evident in the 1D trace extracted, where the signal to noise ratio appeared to double with each increase in temperature. Signal to noise ratios of 32.3 for 50 °C, 15.9 for 40 °C and 8.5 for 25 °C were measured. Similarly, the peak widths at half height for cross peaks observed were narrower at 50 °C (0.040 ppm/28.20 Hz) than at lower temperature (40 °C: 0.052 ppm/37.95 Hz; 25 °C: 0.054/38.46 Hz). Although running 2D ^{15}N – ^1H HSQC at 50 °C gave the best signal to noise ratio and sharper resonances, this temperature could not be sustained on our current spectrometers over a number of days, and a compromise temperature of 40 °C was used to acquire all further bicelle experiments. Solution NMR data for $q=0.5$ bicelles resulted in the observation of the correct number of resonances for each of the labelled amino acids within BPV E5_{v2}, but with lower signal to noise for the cross peak arising from glycine 11. Therefore a second bicelle sample was prepared in an attempt to improve the signal to noise ratio of ^{15}N labelled amide groups and the dispersion of resonances in the spectra, as there

appeared to be some spectral overlap for two of the unassigned resonances within the sample (in contrast to samples prepared in TFE where all signals were well dispersed).

6.3.1 Solution NMR of BPV E5_{V2} in $q=0.33$ bicelles.

The second bicelle sample was prepared with an increased DHPC concentration in order to lower the q ratio to 0.33, as a lower q ratio should result in bicelles with a smaller diameter that tumble faster in solution, leading to an improvement in the spectra obtained. In addition to reducing the q ratio, the peptide concentration was increased and the lipid concentration decreased from 10% total amphiphile to 2% in order to decrease the lipid to protein ratio considerably (from 80:1 to 30:1 w/w). This reduced the viscosity of the sample solution in a further attempt to improve bicelle tumbling rates. A third bicelle sample was also prepared with a $q=0.33$ ratio, but using an alternative method for reconstitution of peptide into bicelles. For $q=0.5$ bicelles, lipids were weighed out individually and dissolved directly into TRIS buffer prior to being mixed together at the correct quantities to form bicelles of the required q ratio, which was then added to a dried film of BPV E5_{V2} peptide. For the third bicelle sample, both DMPC and DHPC were dissolved in TFE and the correct volumes, mixed together and dried down to a film before being resuspended in buffer to form a stock 2% total amphiphile solution of $q = 0.33$ bicelles. This stock bicelle solution was then added to a dried BPV E5_{V2} peptide film in order to form bicelles with a 30:1 w/w (40:1 molar ratio) lipid to protein ratio. All of the bicelle samples were then analysed in order to compare whether any of the changes made in the sample preparation yielded an improvement in signal intensity or line width in the spectra recorded.

An overlay of the 2D HSQC spectra for all three bicelle samples is shown in **Figure 6.7**. When comparing the two $q=0.33$ bicelle samples, the sample prepared by co-dissolving lipids in TFE to prepare a bicelle stock solution in HEPES buffer appeared to have slightly higher signal intensity and sharper resonances than that of the sample prepared using the original method and in TRIS buffer. Therefore this sample was the one used for further analysis of the BPV E5_{V2} peptide. Comparing the three spectra, it was seen that by decreasing the q value from 0.5 to 0.33 and by reducing the lipid to protein ratio to 30:1 w/w, whilst also decreasing the percentage total amphiphile in

solution from to 2%, had a positive effect on the spectrum obtained. The resonances observed from ^{15}N labelled amide groups had greatly improved signal intensity, in particular the resonance tentatively assigned as arising from glycine 11 in the BPV E5_{V2} peptide was now clearly visible above the noise, as was the case in spectra recorded in TFE. Again six resonances were observed for the six isotopic ^{15}N labels incorporated into the peptide.

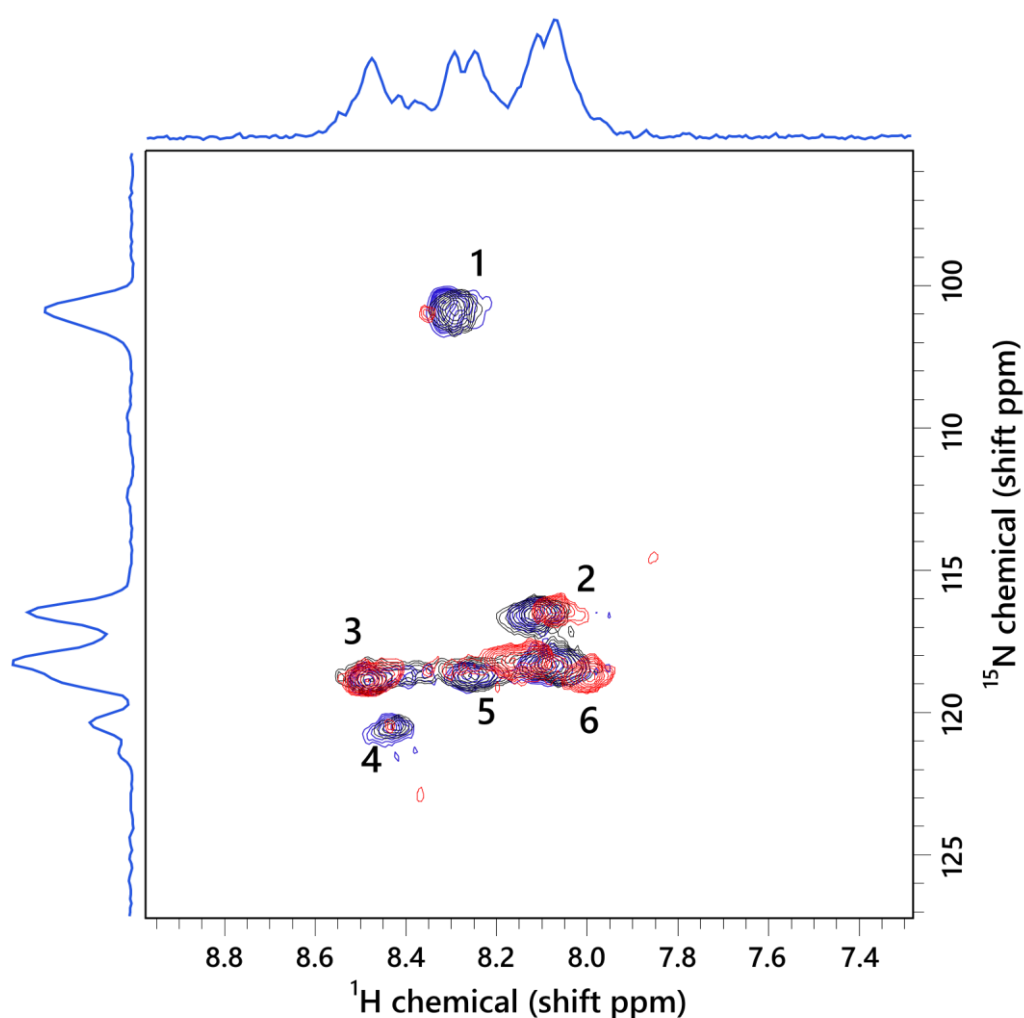


Figure 6.7 Overlay of 2D ^{15}N - ^1H HSQC of BPV E5_{V2} reconstituted into q=0.33 and q=0.5 bicelles

2D ^{15}N - ^1H HSQC spectra of BPV E5_{V2} peptide reconstituted into DMPC/DHPC bicelles with a q of 0.5 (red), 0.33 in TRIS (blue) and 0.33 in HEPES (black). Spectra recorded on a 700 MHz spectrometer (Bruker, Avance II) at 40 °C (313 K). Signal intensity of the q=0.33 bicelle sample was much higher, with the resonance arising from glycine 11 much more apparent than in the q=0.5 bicelle sample.

6.3.2 3D HSQC TOCSY and NOESY analysis of BPV E5_{V2}

To assign the resonances observed in the 2D $^{15}\text{N} - ^1\text{H}$ HSQC spectra recorded for BPV E5_{V2} peptides reconstituted into bicelles, ^{15}N edited $^1\text{H} - ^1\text{H}$ HSQC NOESY and HSQC TOtal Correlation SpectroscopY (TOCSY) spectra were acquired. By running experiments using ^{15}N as the third dimension, signals from DMPC and DHPC lipids in addition to any signal from HEPES buffer were filtered out, enabling only those resonances which were isotopically ^{15}N labelled to be observed. 3D experiments were recorded as “pseudo 3D” experiments, as only a single plane was recorded in the ^{15}N dimension. Since only six amino acids were isotopically labelled within BPV E5_{V2} peptide, the resolution obtained by running “full” 3D experiments was not required and as such experimental time was reduced greatly allowing for a larger number of co-added transients to be recorded.

The $^1\text{H} - ^1\text{H}$ TOCSY sequence is a homonuclear experiment, from which *through bond* information for protons that are coupled together within the same spin system (i.e. all protons within a single amino acid) can be obtained from scalar couplings, the magnitude of which depend upon the number of intervening bonds. Protons from different amino acids give rise to different spin systems as there is no scalar coupling across the amide bond. Analysis of cross peaks arising from different spin systems can be used for the assignment of each amino acid based upon the chemical shifts within each system and known resonance patterns. ^{15}N edited $^1\text{H} - ^1\text{H}$ TOCSY-HSQC spectra were recorded for BPV E5_{V2} in bicelles using short (30 ms) and long (100 ms) TOCSY mixing times. A ^{15}N edited $^1\text{H} - ^1\text{H}$ TOCSY spectrum with a 60 ms spin-lock time is shown in **Figure 6.10**. From $^1\text{H} - ^1\text{H}$ TOCSY-HSQC spectra recorded, a number of broad TOCSY cross peaks were observed, but the number of resonances observed fell short of the number of resonances expected. Tentative assignments of cross peaks observed were made based upon published chemical shift data at the BMRB in conjunction with NOESY-HSQC spectra recorded, however these assignments need verification in the future. Assignment of resonances observed in TOCSY and NOESY spectra of BPV E5_{V2} in bicelles was difficult due to the increase in signal broadening when compared to those recorded in TFE, as lipid bicelles tumble at a slower rate than that of TFE or detergent micelles. At 40 °C (data not shown) weak TOCSY signals were only observed for coupling of HN - αH nuclei in each of the ^{15}N labelled amino acids within the BPV E5_{V2} peptide. Increasing the temperature to 50 °C resulted in a slightly improved spectrum, in which in

addition TOCSY cross peak from a β H proton was observed in the tentatively assigned valine 13/30 spin system. At longer TOCSY mixing times (100 ms), which should allow for longer-range through bond correlations to be observed, no additional cross peaks were observed.

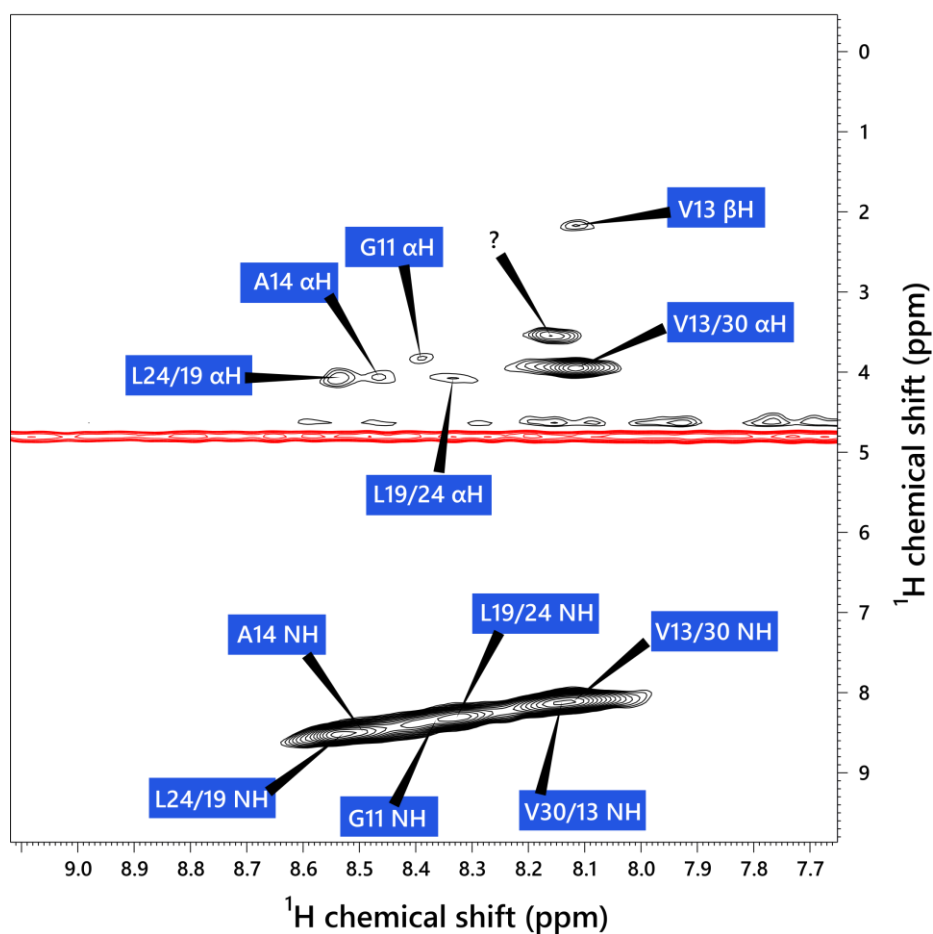


Figure 6.8 ^{15}N edited ^1H - ^1H HSQC TOCSY of BPV E5_{v2} reconstituted into $q=0.33$ bicelles

2D plane of a 3D ^{15}N edited ^1H - ^1H HSQC TOCSY spectrum with a 60 ms TOCSY mixing time of BPV E5_{v2} peptide reconstituted into $q=0.33$ bicelles. Spectrum recorded on a 700 MHz spectrometer (Bruker, Avance II) at 50 °C (323 K). Tentative assignments were made based upon published chemical shift and resonance pattern data.

The ^1H - ^1H NOESY sequence is used to gain information for protons that interact *through space* that are coupled together by dipolar couplings. The intensity of NOE signals observed depends upon the distance between coupled nuclei (where distance is proportional to r^{-6}) with closer nuclei giving higher intensity NOE cross peaks. As with TOCSY experiments, analysis of the cross peaks arising from different spin systems for which NOEs are observed can be used for the assignment of each amino acid based upon the chemical shifts of the resonances observed. However, at longer NOESY mixing times, ambiguity can arise when NOESY spectra are used for assignment purposes.

^{15}N edited ^1H – ^1H NOESY-HSQC spectra were recorded for BPV E5_{v2} over a range of NOESY mixing times from 20 ms to 200 ms. **Figure 6.11** shows a ^{15}N edited ^1H – ^1H NOESY spectrum recorded with a 50 ms mixing time. A number of resonances were observed arising from intra-residue NOEs between protons within each amino acid spin system. Cross peaks observed were broad and overlapping, particularly at longer mixing times. Tentative assignments of cross peaks were made based upon published chemical shift data at the BMRB. Assignment of glycine 11 and alanine 14 spin systems was particularly straightforward based upon the chemical shift of resonances observed. No alanine αH cross peak was observed, but a cross peak for alanine βH was assigned.

Although cross peaks were broad and with some overlap of resonances due to similar chemical shift values, in particular of those thought to arise from labelled leucine 19/24 and valine 13/30, assignment was still attempted. Based upon the predicted molecular model of the BPV E5 homodimer prepared in CHI, the leucine 19 spin system was tentatively assigned as being distinguishable from leucine 24, since leucine 24 residues are predicted to lie at the homodimer interface whereas leucine 19 residues face away from the interface. This could result in a larger signal intensity for interacting leucines with a similar chemical shift and as such tentatively assigned leucine spin system with the higher signal intensity was assigned to leucine 24. The presence of a potential cross peak arising from an NOE between valine 13 and alanine 14, was used to distinguish between valine spin systems. A potential cross peak arising from an NOE between glycine 11 and alanine 14 was also observed.

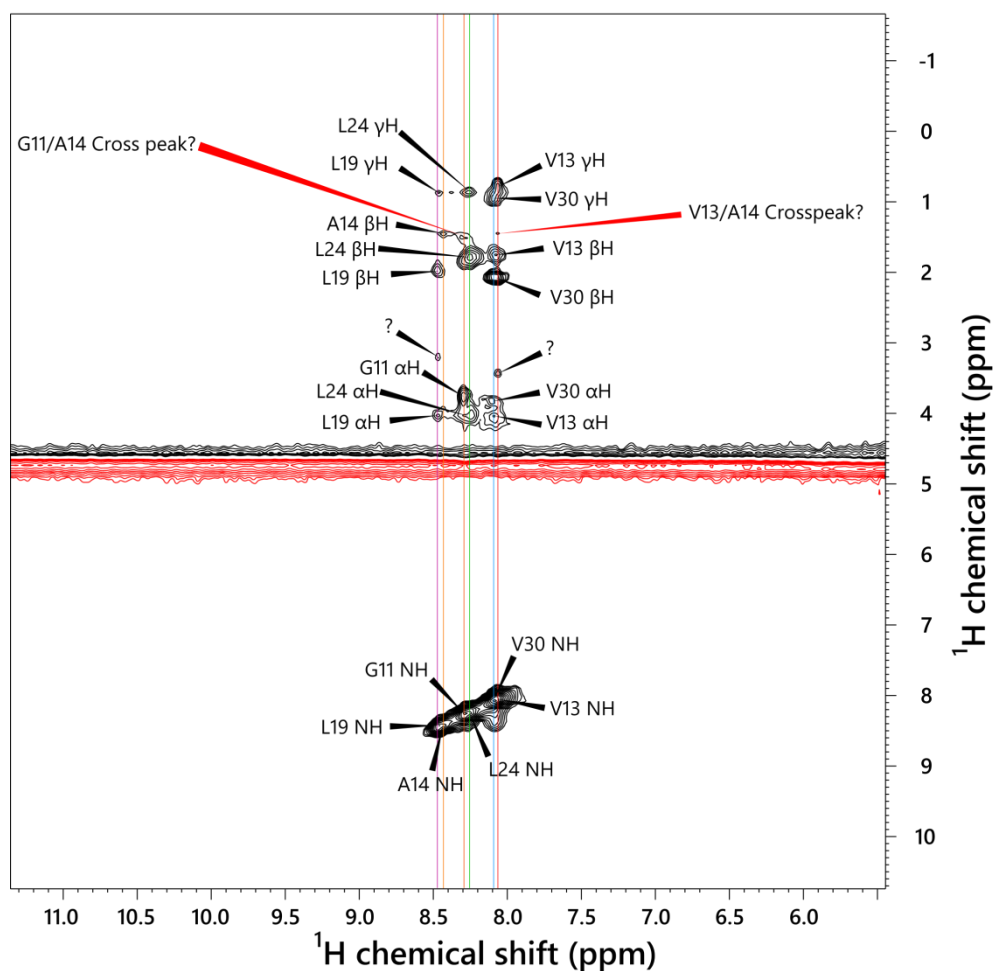


Figure 6.9 ^{15}N edited ^1H - ^1H NOESY-HSQC spectrum of BPV E5_{V2} reconstituted into $q=0.33$ bicelles

2D plane of a 3D ^{15}N edited ^1H - ^1H NOESY-HSQC spectrum of BPV E5_{V2} peptide reconstituted into $q=0.33$ bicelles recorded with a 50 ms NOESY mixing time. Spectrum recorded on a 700 MHz spectrometer (Bruker, Avance II) at 50 °C (323 K). Tentative assignments were made based upon published chemical shift and resonance pattern data. Coloured lines are used to differentiate between observed spin systems.

A combination of chemical shift data from both ^{15}N edited ^1H - ^1H TOCSY and NOESY-HSQC spectra were used to tentatively assign the 2D ^{15}N - ^1H HSQC spectrum of BPV E5_{v2} in $q=0.33$ bicelles (**Figure 6.12**). The majority of resonances were assigned with confidence, in particular those arising from glycine 11 and alanine 14, with the resonances arising from valine 13/30 and leucine 19/24 assigned with reasonable confidence. A chemical shift table based upon experimental data is given in **Table 6.1**. In order to be able to assign all resonances unambiguously data from TOCSY experiments would be required.

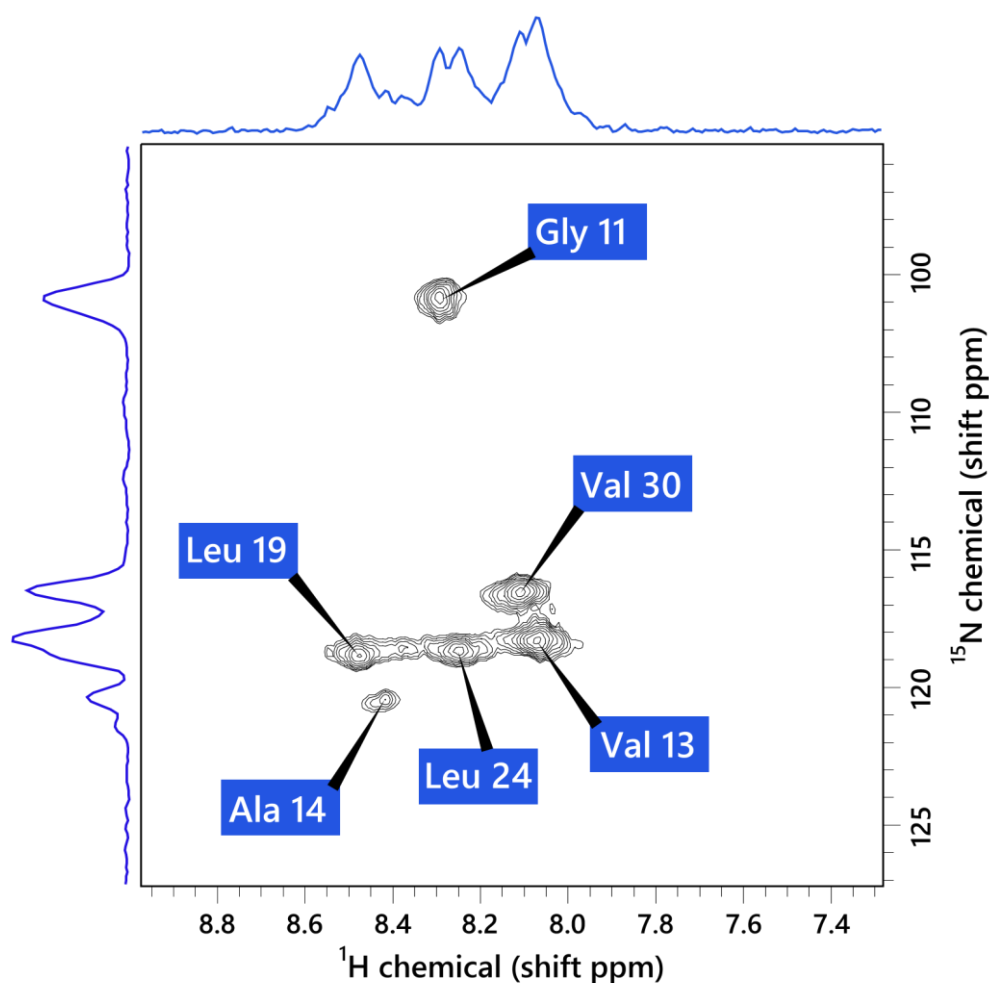


Figure 6.10 2D ^{15}N - ^1H HSQC of BPV E5V2 reconstituted into $q=0.33$ bicelles with tentative amino acid assignments

2D ^{15}N - ^1H HSQC spectrum of BPV E5_{v2} peptide reconstituted into $q=0.33$ bicelles. Spectrum recorded on a 700 MHz spectrometer (Bruker, Avance II) at 40 °C (313 K) Tentative assignments shown on the spectrum were made based upon published chemical shift and TOCSY/NOESY data.

Table 6.1 Chemical shift assignments for ¹⁵N labelled amino acids in BPV E5_{v2} reconstituted in DMPC/DHPC bicelles

Experimental chemical shifts (in ppm) for tentatively assigned amino acids derived from 2D ¹⁵N – ¹H HSQC and ¹⁵N edited 3D ¹H – ¹H HSQC TOCSY and NOESY spectra of BPV E5V2 reconstituted into DMPC/DHPC bicelles.

	N	NH	αH	βH	γH
Gly11	105.88	8.33	3.75	--	--
Val13	118.39	8.06	4.03	1.74	0.80
Ala14	120.57	8.43	3.97	1.44	--
Leu19	118.85	8.47	4.03	1.98	0.86
Leu24	118.70	8.24	4.03	1.80	0.86
Val30	116.51	8.10	3.82	2.09	0.92

6.4 Solution NMR analysis of BPV E5_{LF} in bicelles

An advantage of solution NMR over that of ssNMR is the much narrower line width of resonances obtained in solution NMR spectra in comparison to the broad line widths observed in solid samples. To exploit this property, solution NMR was used to analyse the E5_{LF} peptides originally designed for ssNMR studies. The preliminary results obtained from the singly labelled ¹³C/¹⁵N BPV E5 peptides are presented in this section. In order to probe for any potential through space couplings between labelled leucine and phenylalanine at the BPV E5 homodimer interface using hetero nuclear solution NMR experiments making use of ¹³C isotopically labelled carbon atoms, a bicelle sample was prepared using the singly labelled BPV E5 peptides originally designed for ssNMR analysis. Peptides labelled at leucine 24 (BPV E5_L) and phenylalanine 28 (BPV E5_F) were mixed (BPV E5_{LF}) and reconstituted into q= 0.25 deuterated DMPC/DHPC bicelles at a 30:1 (w/w) lipid to protein ratio for analysis by solution NMR. For the BPV E5_{LF} bicelle sample, deuterated *d*₅₄-DMPC and *d*₂₂-DHPC were used, so as to explore the use of ¹H-¹H TOCSY and NOESY experiments without interference of proton signals from lipid and detergent acyl chains. An initial ¹⁵N – ¹H HSQC spectrum of BPV E5_{LF} reconstituted into q = 0.25 bicelles was recorded (**Figure 6.13**). From the 2D HSQC spectrum obtained, two resonances were observed for the two ¹⁵N labelled amide groups within the sample. Using published chemical shift data an initial assignment of resonances was made (**Table 6.2**). Although the chemical shift values of resonances observed were similar to one another in the ¹H dimension, resonances were resolvable in the ¹⁵N dimension.

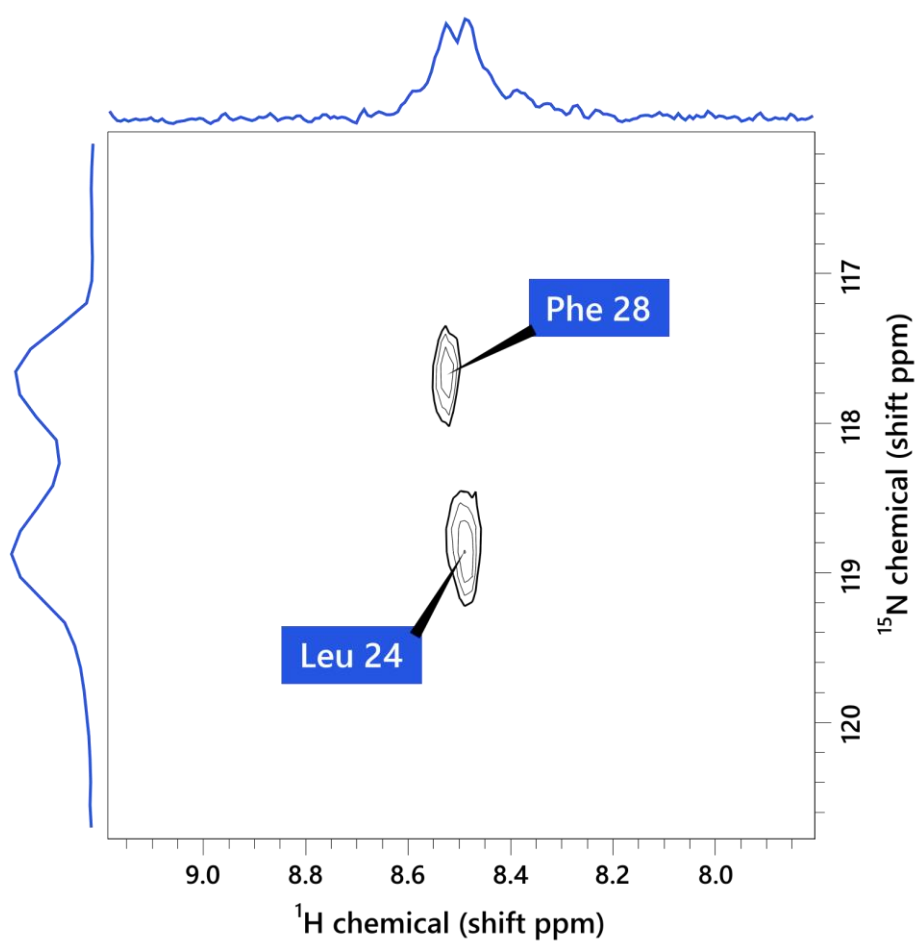


Figure 6.11 2D ^{15}N - ^1H HSQC of BPV E5_{LF} reconstituted into q=0.25 bicelles

2D ^{15}N - ^1H HSQC spectra of $^{13}\text{C}/^{15}\text{N}$ labelled BPV E5_{LF} peptide reconstituted in DMPC/DHPC bicelles with a q of 0.25 at a lipid to protein ratio of 30:1 and 2% total amphiphile. Spectrum recorded on a 600 MHz spectrometer (Bruker, Avance III) at 50 °C (323 K). Two resonances were observed for isotopically labelled leucine 24 and phenylalanine 28 amide groups.

Table 6.2 Tentative chemical shift assignments for labelled BPV E5 Leu 24 and Phe 28 in bicelles

Leu 24 ¹⁵ N chemical shifts (ppm)	
N	118.85
NH	8.48

Phe 28 ¹⁵ N chemical shifts (ppm)	
N	117.76
NH	8.51

Following initial 2D ¹⁵N – ¹H HSQC experiments, homonuclear ¹H – ¹H -TOCSY experiments were acquired for assignment of resonances and ¹H – ¹H -NOESY experiments to investigate potential interactions between labelled leucine and phenylalanine through dipolar couplings (data not shown). These data, recorded at lower field (600 MHz), showed a considerable lack of cross peaks off the diagonal, and no improvement upon increase of temperature, as such the physical properties of isotropic bicelles were investigated.

6.5 Characterisation of DMPC/DHPC bicelles

In an attempt to understand why bicelle samples gave rise to cross peaks in ¹⁵N – ¹H HSQC and ¹⁵N edited ¹H – ¹H HSQC-NOESY spectra, but very few cross peaks in the ¹⁵N edited ¹H – ¹H TOCSY-HSQC and homonuclear correlation spectra, bicelles were analysed by ³¹P NMR and Dynamic light scattering (DLS). Using ³¹P solution NMR, the formation of isotropic bicelles at $q=0.33$ was assessed. DLS was used to measure the approximate size of the bicelle-peptide aggregates formed at this q ratio.

6.5.1 Analysis of bicelles by ^{31}P solution NMR

To confirm the presence of isotropic bicelles in samples prepared using a q of 0.33, ^{31}P solution NMR spectra were recorded on empty bicelle samples prepared with a 2% total amphiphile solution. These empty bicelles were analysed by ^{31}P NMR over the range of temperatures previously used when collecting ^{15}N - ^1H HSQC spectra. In the ^{31}P solution NMR spectra obtained (**Figure 6.12**), two distinct resonances were observed. At all temperatures, the downfield peak intensity was found to be three times that of the upfield peak, similarly the integrated area beneath the higher intensity peak was three times that of the lower intensity upfield peak and integrated area under the downfield peak. Therefore the higher intensity, downfield peak was assigned to the phosphate head group of DHPC and the upfield peak was assigned to DMPC. The intensity and integral volumes correlate well with the experimental q ratio of 0.33 used to prepare bicelle samples. As has been previously demonstrated, DHPC and DMPC ^{31}P spectra recorded in isolation result in resonances of identical chemical shift values, therefore the presence of two distinguishable resonances suggests that the phosphate head groups of DHPC and DMPC are in two distinctly different chemical/magnetic environments (Glover, Whiles et al. 2001).

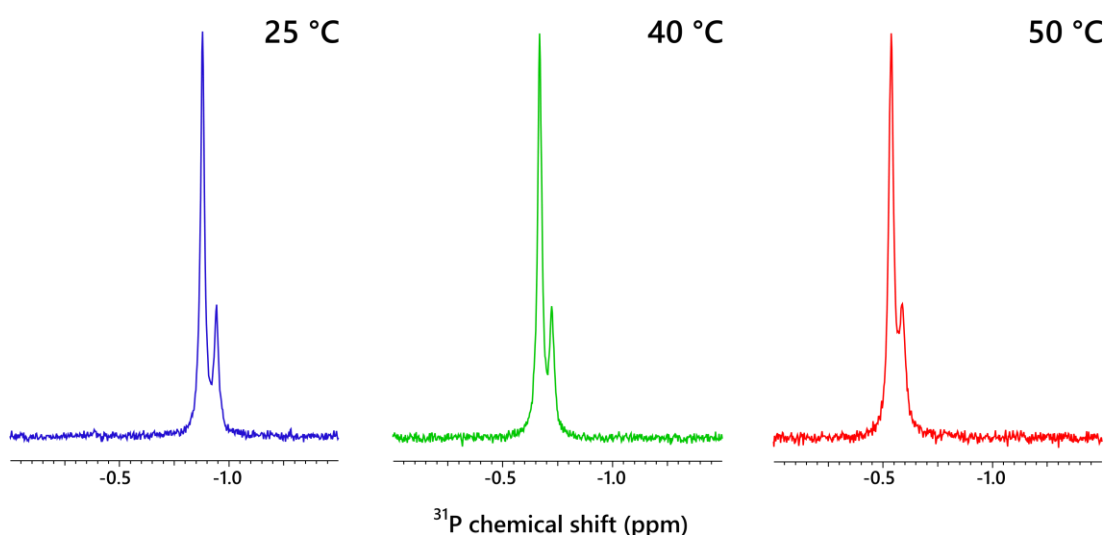


Figure 6.12 ^{31}P solution NMR spectra of $q=0.33$ bicelles at increasing temperature

^{31}P solution NMR spectra of $q = 0.33$ bicelles in 2% total amphiphile solution in 20 mM TRIS buffer, pH 7.4. Spectra were recorded at 600 MHz over a range of temperatures with a spectral window of 100 ppm and referenced by placing the carrier frequency at the centre of the spectrum.

The presence of a downfield resonance is characteristic of the isotropic behaviour of non-aligned bicelles (Marcotte and Auger 2005). The presence of two resonances (from DMPC and DHPC) at all temperatures suggests that when using a q ratio of 0.33, isotropic bicelles were formed.

6.5.2 Analysis of bicelles by DLS

To assess the size of DMPC/DHPC isotropic bicelle samples prepared, and to observe the effect of insertion of BPV E5_{v2} peptide on the size of bicelles following reconstitution, empty and peptide-loaded bicelle samples were analysed using DLS. DLS is a biophysical technique that is commonly used to measure the size of particles in solution by measuring the Brownian motion of particles and correlating that to particle size. The relationship between particle size and its diffusion coefficient due to Brownian motion is defined by the Stokes-Einstein equation (**Equation 10**) and is used to report the hydrodynamic radius of particles. In Equation 10, D_t is the translational diffusion coefficient (m^2/s), k_B is the Boltzmann constant, T the absolute temperature (K), η the viscosity of solution and R_h the hydrodynamic radius.

$$D_t = \frac{k_B T}{6\pi\eta R_h} \quad (10)$$

Equation 10 can be rearranged to give the hydrodynamic radius of a sphere (**Equation 11**)

$$R_h = \frac{k_B T}{6\pi\eta D_t} \quad (11)$$

The hydrodynamic radii of empty $q=0.33$ bicelles in 2% total amphiphile solution, prepared in both TRIS and HEPES buffer, as well as bicelles containing BPV E5_{v2} peptide are given in **Table 6.3**.

Table 6.3 Hydrodynamic radii of empty and bicelles with inserted BPV E5_{v2} peptides

Hydrodynamic radii (R_h) of $q=0.33$ bicelles in 2% total amphiphile solution measured at 40 °C. Values were averaged over three measurements and the standard deviation given. The polydispersity index (PI) as a measure of sample homogeneity is also given.

Sample	R_h (nm)	SD	PI
E5 _{v2} (HEPES)	7.97	0.67	0.85
Empty (HEPES)	5.64	0.24	0.54
E5 _{v2} (TRIS)	13.77	0.79	0.80
Empty (TRIS)	4.39	0.80	0.30

Empty bicelle samples prepared in TRIS buffer resulted in bicelles of the smallest diameter (4.39 nm) and a low PI value of 0.30, indicating that the bicelles in solution were uniform in size. In contrast, bicelles prepared in TRIS buffer and containing BPV E5_{v2} peptides were much larger (13.77 nm) and had a high PI value indicating a lack of homogeneity in the sample, as evidenced by the multiple populations of differing particle sizes observed. For samples prepared in HEPES buffer (as used in solution NMR experiments), the hydrodynamic radii of both empty and peptide-loaded bicelles were similar with empty bicelles at 5.64 nm and with peptide-loaded (7.97). As aggregate size increased, so did the PI, making interpretation of DLS measurements difficult due to heterogeneity of the sample. Ideal PI values in the range of 0.3-0.6 indicate sample homogeneity, with values closer to 1 indicative of highly polydisperse samples, generally not suited to characterisation by DLS. It should be noted that samples were unfiltered so as to not affect bicelle concentration through any potential loss of bicelles via interactions with membranes. As such, non-filtered dispersions may be affected by the

presence of interfering dust particles that may also result in a non-representative average hydrodynamic radius of bicelles size measured.

The hydrodynamic radii of empty bicelle samples, as detected using DLS, appear to be in the range of those in the literature for isotropic bicelles of similar q ratios and total % amphiphile values. Bicelle diameters of 3-6 nm have been reported for empty, $q=0.5$ isotropic bicelles in 10 % total amphiphile solution and 4-10 nm when % amphiphile is reduced to 2.5% (Glover, Whiles et al. 2001). For isotropic bicelles of $q = 0.3$ ratios, diameters of 12 -17 nm have been reported (Lu, Van Horn et al. 2012). The much larger diameters observed in bicelle samples prepared before and after reconstitution of BPV E5_{v2} peptide may be a result of aggregated bicelles/peptide which causes the fusion of bicelles. The larger diameter of the resulting bicelle/protein complexes would tumble at a much slower rate in comparison to smaller isotropic bicelles, leading to reduced signal intensity and broadening of resonances in the spectra recorded. Although larger diameters have been reported, those of peptide-loaded bicelles used to obtain high resolution NMR spectra have been reported to be much smaller, at around 2.7-3 nm (Bocharov, Mineev et al. 2008), suggesting that the bicelles used in this study are just on the edge of this size useable, hence the lack of TOCSY data.

6.6 Summary

Although a powerful technique, ssNMR is a less well established method for protein structure determination in comparison with solution NMR, and is inherently less sensitive resulting in lower resolution spectra. In particular, the use of 3D heteronuclear experiments is commonplace for solution NMR studies, therefore the use of solution NMR for investigating the structure of the BPV E5 membrane protein as a complimentary technique to ssNMR was investigated.

In this chapter the preliminary results for BPV E5 peptide reconstituted into DMPC/DHPC bicelles were presented. Previously to date, solution NMR data obtained for BPV E5 has only been reported in the organic solvent TFE and in SDS detergent micelles (Windisch, Hoffmann et al. 2010; King, Oates et al. 2011). Using fast tumbling

isotropic bicelles as a membrane mimetic system for the reconstitution of BPV E5 peptides for solution NMR studies, we were able to record NMR spectra in a more “native-like” lipid environment, resulting in more biologically relevant data. HSQC spectra of BPV E5_{V2} in isotropic bicelles showed well dispersed resonances for ¹⁵N labelled amide groups indicating that reconstituted BPV E5_{V2} peptide was adopting a folded secondary structure in isotropic DMPC/DHPC bicelles. When reconstituted into bicelles only a single species for BPV E5 peptide was observed, and this is contrast to SDS micelles, where dependant on the detergent concentration the presence of a secondary set of resonances was reported (King, Oates et al. 2011) indicating the presence of protein in a different chemical environment. These secondary resonances were suggested to arise from BPV E5_{V2} monomers upon increasing detergent concentration. For BPV E5_{V2} bicelle samples only a single set of resonances were observed at all *q* ratios and temperatures tested, suggesting that in isotropic bicelles, stable BPV E5_{V2} dimers were being formed.

Resonances observed in 2D ¹⁵N-¹H HSQC spectra of BPV E5_{V2} were tentatively assigned using a combination of 3D ¹⁵N edited ¹H-¹H TOCSY-HSQC and NOESY-HSQC experiments. Typically the data from COSY and TOCSY type experiments are required in order to assign spectra with confidence. Unfortunately, our TOCSY-HSQC spectra contained significantly fewer cross peaks than expected. From TOCSY-HSQC experiments recorded at 40 °C, only NH coupling to αH protons was detected, whereas recording experiments at 50 °C some NH-βH coupling was also observable. Due to spectrometer limitations, ¹⁵N edited ¹H-¹H HSQC-TOCSY and NOESY spectra were only able to be acquired at 40 °C. The slower tumbling of bicelles at lower temperature may have been a reason for the lack of TOCSY signals, as increasing the temperature to 50 °C saw a moderate improvement in signal intensity as well as more observable cross peaks (e.g. βH protons). Therefore running all experiments at 50 °C would have been beneficial when attempting to improve signal to noise for 3D experiments, as running at higher temperatures is a method commonly used for improving the resolution of slower tumbling molecules in solution. Elevating the temperature leads to an overall increase in molecular tumbling rate and reduces the transverse (*T*₂) relaxation rate of slow tumbling molecules, leading to improvement in the line-width of spectra recorded.

Due to unforeseen equipment reliability issues, experiments that were originally recorded at 700 MHz on a spectrometer equipped with a cryoprobe, which cools the probe coil and preamplifier with a stream of helium gas, reducing the thermal noise generated by the electronics within the probe, thereby increasing sensitivity up to tenfold, then had to be recorded at lower field (600 MHz) due to equipment failure. Although it was possible to record experiments at higher temperatures (above 40 °C), the lower field and the lack of a cryoprobe ultimately resulted in increased experimental times and a reduction in signal to noise. This was evident in $^{13}\text{C}/^{15}\text{N}$ experiments recorded on BPV E5_{LF} bicelle samples where a distinct lack of NOE and TOCSY signals were observed.

As slower tumbling bicelles were thought to be the reason for the lack of TOCSY signals observed, the presence and size of isotropic bicelles were analysed by ^{31}P solution NMR and DLS. The results from ^{31}P spectra of bicelle samples, indicated that at all temperatures above the T_m of DMPC, isotropic bicelles were present, although the size of which were not ascertainable without the use of experiments to measure the translational diffusion rate of bicelles (Andersson and Maler 2005; Lu, Van Horn et al. 2012). DLS analyses of bicelles was used to estimate the size of bicelles in solution, both empty and with peptide inserted. In order to analyse bicelle samples with a q of 0.33 and a total % amphiphile value of 2%, samples had to specifically be prepared for DLS analysis as solution NMR samples could not be used directly due to the high concentration of particles in the sample. DLS analysis of empty bicelles would suggest that those prepared in HEPES buffer, as used for solution NMR experiments, had an increased hydrodynamic diameter in comparison to those prepared in TRIS buffer. Upon the addition of peptide to bicelles the diameter of the bicelles recorded almost doubled in size. This may be as a result of peptide aggregation, though not visible from NMR spectra recorded which indicate the presence of a protein in a single chemical environment, therefore the increase in size may be due to the fusion of bicelles in solution.

It has been shown that bicelles in dilute solutions (<15% total amphiphile) grow at a much larger rate than of those in concentrated sample preparations as free DHPC detergent is sequestered from solution into bicelles that have been formed in order to maintain the inaggregate "free" DHPC concentration (Cohen, Thurston et al. 1998). In

this study an initial concentration of 10% total amphiphile was used which was then reduced to 2% for further studies. For future analysis of BPV E5 peptides in bicelles it would be advisable to not only screen for a ratio of q values but also adjust the q value of bicelles in conjunction with the % total amphiphile of solution in an attempt to obtain spectra of higher quality.

6.7 References

- Andersson, A. and L. Maler (2005). "Magnetic resonance investigations of lipid motion in isotropic bicelles." *Langmuir* **21**(17): 7702-7709.
- Bax, A. and N. Tjandra (1997). "High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium." *J Biomol NMR* **10**(3): 289-292.
- Bocharov, E. V., K. S. Mineev, P. E. Volynsky, Y. S. Ermolyuk, E. N. Tkach, A. G. Sobol, V. V. Chupin, M. P. Kirpichnikov, R. G. Efremov and A. S. Arseniev (2008). "Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state." *J Biol Chem* **283**(11): 6950-6956.
- Bocharov, E. V., Y. E. Pustovalova, K. V. Pavlov, P. E. Volynsky, M. V. Goncharuk, Y. S. Ermolyuk, D. V. Karpunin, A. A. Schulga, M. P. Kirpichnikov, R. G. Efremov, I. V. Maslennikov and A. S. Arseniev (2007). "Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger." *J Biol Chem* **282**(22): 16256-16266.
- Cohen, D. E., G. M. Thurston, R. A. Chamberlin, G. B. Benedek and M. C. Carey (1998). "Laser light scattering evidence for a common wormlike growth structure of mixed micelles in bile salt- and straight-chain detergent-phosphatidylcholine aqueous systems: relevance to the micellar structure of bile." *Biochemistry* **37**(42): 14798-14814.
- De Angelis, A. A., S. C. Howell, A. A. Nevzorov and S. J. Opella (2006). "Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy." *J Am Chem Soc* **128**(37): 12256-12267.
- De Angelis, A. A., A. A. Nevzorov, S. H. Park, S. C. Howell, A. A. Mrse and S. J. Opella (2004). "High-resolution NMR spectroscopy of membrane proteins in aligned bicelles." *J Am Chem Soc* **126**(47): 15340-15341.
- Glover, K. J., J. A. Whiles, G. Wu, N. Yu, R. Deems, J. O. Struppe, R. E. Stark, E. A. Komives and R. R. Vold (2001). "Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules." *Biophys J* **81**(4): 2163-2171.
- Hwang, T. L. and A. J. Shaka (1995). "Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients." *Journal of Magnetic Resonance, Series A* **112**(2): 275-279.
- King, G., J. Oates, D. Patel, H. A. van den Berg and A. M. Dixon (2011). "Towards a structural understanding of the smallest known oncoprotein: investigation of the bovine papillomavirus E5 protein using solution-state NMR." *Biochim Biophys Acta* **1808**(6): 1493-1501.
- Lu, Z., W. D. Van Horn, J. Chen, S. Mathew, R. Zent and C. R. Sanders (2012). "Bicelles at low concentrations." *Mol Pharm* **9**(4): 752-761.

- Luchette, P. A., T. N. Vetman, R. S. Prosser, R. E. Hancock, M. P. Nieh, C. J. Glinka, S. Krueger and J. Katsaras (2001). "Morphology of fast-tumbling bicelles: a small angle neutron scattering and NMR study." *Biochim Biophys Acta* **1513**(2): 83-94.
- Luo, P. and R. L. Baldwin (1997). "Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water." *Biochemistry* **36**(27): 8413-8421.
- Marcotte, I. and M. Auger (2005). "Bicelles as model membranes for solid- and solution-state NMR studies of membrane peptides and proteins." *Concepts in Magnetic Resonance Part A* **24A**(1): 17-37.
- Mattoon, D., K. Gupta, J. Doyon, P. Loll and D. DiMaio (2001). "Identification of the transmembrane dimer interface of the bovine papillomavirus E5 protein." *Oncogene* **20**(29): 3824-3834.
- Morrison, E. A. and K. A. Henzler-Wildman (2012). "Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile." *Biochim Biophys Acta* **1818**(3): 814-820.
- Park, S. H., S. Prytulla, A. A. De Angelis, J. M. Brown, H. Kiefer and S. J. Opella (2006). "High-resolution NMR spectroscopy of a GPCR in aligned bicelles." *J Am Chem Soc* **128**(23): 7402-7403.
- Sanders, C. R., 2nd and G. C. Landis (1995). "Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies." *Biochemistry* **34**(12): 4030-4040.
- Sanders, C. R., 2nd and J. H. Prestegard (1990). "Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO." *Biophys J* **58**(2): 447-460.
- Sanders, C. R., 2nd and J. P. Schwonek (1992). "Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR." *Biochemistry* **31**(37): 8898-8905.
- Schleucher, J., M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. Sorensen and C. Griesinger (1994). "A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients." *J Biomol NMR* **4**(2): 301-306.
- Struppe, J., E. A. Komives, S. S. Taylor and R. R. Vold (1998). "²H NMR studies of a myristoylated peptide in neutral and acidic phospholipid bicelles." *Biochemistry* **37**(44): 15523-15527.
- Struppe, J., J. A. Whiles and R. R. Vold (2000). "Acidic phospholipid bicelles: a versatile model membrane system." *Biophys J* **78**(1): 281-289.
- Triba, M. N., M. Zoonens, J. L. Popot, P. F. Devaux and D. E. Warschawski (2006). "Reconstitution and alignment by a magnetic field of a beta-barrel membrane protein in bicelles." *Eur Biophys J* **35**(3): 268-275.
- Vold, R. R., R. S. Prosser and A. J. Deese (1997). "Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides." *J Biomol NMR* **9**(3): 329-335.
- Watts, A. and P. J. Spooner (1991). "Phospholipid phase transitions as revealed by NMR." *Chem Phys Lipids* **57**(2-3): 195-211.
- Windisch, D., S. Hoffmann, S. Afonin, S. Vollmer, S. Benamira, B. Langer, J. Burck, C. Muhle-Goll and A. S. Ulrich (2010). "Structural role of the conserved cysteines in the dimerization of the viral transmembrane oncoprotein E5." *Biophys J* **99**(6): 1764-1772.

7

DISCUSSION AND FUTURE WORK

7.1 Discussion

The determination of membrane protein structure is a challenging process, as is evident by the extremely low number of structures solved to date, in comparison to water soluble proteins. Given their importance, membrane proteins are particularly underrepresented with the total number of unique structures deposited at the Protein data bank (PDB), accounting for less than 1% of all unique structures deposited to date. Solid state NMR (ssNMR) is well placed to provide the means to determine the structure of membrane proteins in "native-like" lipid bilayers, allowing for the study of membrane proteins that would either be too large to study by solution NMR methods, due to the slow tumbling of proteins in a large detergent/protein complex.

The challenge presented when using ssNMR is the preparation of homogenous samples that give rise to well resolved spectra, particularly when the membrane protein structure being studied is embedded in fully hydrated lipid bilayers, which tend to result in spectra with increased resonance line widths over micro/nano crystalline samples, therefore optimal sample preparation methods were sought.

The aim of this study was to evaluate and develop reliable and reproducible sample preparation methods for the reconstitution of small TM proteins into liposomes, in order to study the structure of membrane proteins by ssNMR. Using selectively uniformly $^{13}\text{C}/^{15}\text{N}$ labelled amino acids incorporated at key positions at the homo dimer interface of synthetically prepared peptides, the problem of spectral over-crowding was sought to be overcome. In this study, sample preparation methods for the reconstitution of small α -helical TM proteins into hydrated DMPC/cholesterol lipid bilayers by ssNMR were investigated, using the well characterised membrane protein Glycophorin A (GpA), as a model small α -helical TM protein that has previously been characterised by both

solution and solid state NMR methods, making it an ideal choice as a membrane protein to work with when developing standard protocols. Using GpA, samples were prepared for ssNMR using two separate methods; a “detergent removal” method, which was based upon a previously documented sample preparation method using Bio-beads to reconstitute membrane proteins by removal of detergent from solution. The second method was based upon co-solubilisation of protein and lipids in organic solvent.

Using a range of biophysical techniques, samples prepared using the detergent-removal method were compared alongside those prepared by co-solubilisation. From the biophysical characterisation and solid state NMR spectra obtained using both sample preparation methods, samples prepared by co-solubilisation of peptide with lipid/cholesterol in TFE were of visibly better quality. Samples prepared using the detergent removal method, had resonances arising from a second chemical environment, which using secondary chemical shift analysis was predicted to be due to incorrectly folded or aggregated protein in β -sheet form. Samples prepared using the co-solubilisation method showed little to no sign protein in a secondary environment and was concluded to be the better method for sample preparation.

Using uniformly labelled valine 80 and glycine 83 within GpA peptides, we were able to observe dipolar couplings between the two amino acids through the presence of cross peaks at long mixing times in 2D ^{13}C - ^{13}C DARR correlation spectra. These correlations were attributed to inter-helical interaction as they were only present at long mixing times, but due to the effects of spin diffusion within the peptide chain the possibility of measuring an intra-helical coupling could not be ruled out. Therefore a new labelling strategy was employed using “singly” labelled peptides, with only one labelled amino acid per chain. Using singly labelled GpA peptides, samples prepared by co-solubilisation showed the presence of the same inter-helical couplings, as seen in doubly labelled GpA samples and therefore it was concluded that without doubt that inter-helical couplings were being observed. Therefore we have demonstrated for that GpA in DMPC liposomes with the addition of cholesterol to the lipid bilayers still form dimers similar to those observed in pure DMPC (Smith, Jonas et al. 1994; Smith, Song et al. 2001) and DPC detergent micelles (MacKenzie, Prestegard et al. 1997), this makes the lipid environment in which the GpA protein is embedded more biologically relevant. A

positive step towards making the membrane in which GpA protein is studied more biologically relevant.

Having shown that inter-helical couplings can be observed between singly and doubly labelled peptides, the next step would be to see if we could increase the number of labels incorporated further, thereby reducing the number of samples that would have to be prepared. An advantage of this labelling scheme over using peptides designed with sole spin pairs, is that the number of samples that need to be tested when attempting to study a protein with no existing structure beneficially for when the structure of the protein is unknown as is the case with BPV E5 where no structural data exists only a model prepared from computational analysis using data from mutagenesis studies. Although using singly labelled peptides prepared for the study of the BPV E5 homodimer returned no conclusive results, this was a result of a poor choice of amino acids to label within the peptide. Although limited by choice, the decision to label phenylalanine and tyrosine both of which gave rise to broad unresolvable resonances was in hindsight the wrong decision. In future the incorporation of ^{19}F phenylalanine in the aromatic ring could be used in order to obtain longer range dipolar couplings (Hong, Zhang et al. 2012).

In an attempt to improve the spectral resolution of BPV E5 samples, the effects of increasing the cholesterol on DMPC lipid membranes was investigated as it had previously been reported (Cady, Mishanina et al. 2009) that using higher cholesterol content in conjunction with additional phospholipids such as sphingomyelin (SM) in membranes reduced the fluidity of the membrane at physiological temperatures, allowing for experiments to be conducted at higher temperature. For our samples we chose not to change too many sample parameters from those observed to work for GpA and therefore only the cholesterol content of samples was increased. With no protein signals observable at high temperature, experimental data was recorded at low temperature ($<0\text{ }^{\circ}\text{C}$) as with our previous samples. At low temperature the addition of increased cholesterol lead to the opposite effect of that desired, with an increase in membrane fluidity rather than a decrease. Therefore for future experiments, an investigation on the effect of the addition of increased cholesterol alongside the addition of SM and other phospholipids would be of interest.

For future investigations, the effect of reconstituting GpA into bilayers other than that of DMPC would be of interest. Different phospholipids have different acyl chain lengths and therefore the thickness of the bilayers that they form are also affected (Gallova, Uhrikova et al. 2004).

In order to complement our solid state NMR work, we investigated the use of bicelles for studying membrane protein structure. This was a preliminary study to investigate the feasibility of using solution NMR to complement data obtained by ssNMR. Previous work in our group had concentrated on the use of SDS detergent to solubilise BPV E5 peptides (King, Oates et al. 2011). For this study, bicelles were used as they represent a more "native-like" membrane mimetic environment than that of detergent micelles that can induce curvature stress and affect structure, function or oligomerisation. With SDS detergent micelles our group had reported observing two separate chemical environments upon increasing the SDS concentration. This second population was believed to be due to BPV E5 monomers at high detergent concentrations. In contrast when using bicelles, only one single population was observed at all q ratios. Whilst we were able to obtain heteronuclear HSQC spectra that indicated the presence of a single well folded species whereas in detergent the appearance of a second chemical environment was reported, we were unable to obtain TOCSY data for full assignment of resonances. The cause of the lack of TOCSY signals was believed to be due to bicelles being larger than the fast tumbling isotropic bicelles required for high resolution NMR. DLS data suggests that the bicelle samples prepared were just bigger than those reported in the literature, therefore for future studies the effect of total amphiphile in addition to that of q ratio would be beneficial, as it has been reported that using a lower % total amphiphile in solution can lead to bicelles growing at a faster rate than that at higher concentrations (Lu, Van Horn et al. 2012). The use of larger alignable bicelles for ssNMR could also be a venue for investigation complementary structural data by both solution and solid state NMR.

7.2 References

- Cady, S. D., T. V. Mishanina and M. Hong (2009). "Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding." *J Mol Biol* **385**(4): 1127-1141.
- Gallova, J., D. Uhríkova, M. Hanulova, J. Teixeira and P. Balgavy (2004). "Bilayer thickness in unilamellar extruded 1,2-dimyristoleoyl and 1,2-dierucoyl phosphatidylcholine vesicles: SANS contrast variation study of cholesterol effect." *Colloids Surf B Biointerfaces* **38**(1-2): 11-14.
- Hong, M., Y. Zhang and F. Hu (2012). "Membrane protein structure and dynamics from NMR spectroscopy." *Annu Rev Phys Chem* **63**: 1-24.
- King, G., J. Oates, D. Patel, H. A. van den Berg and A. M. Dixon (2011). "Towards a structural understanding of the smallest known oncoprotein: investigation of the bovine papillomavirus E5 protein using solution-state NMR." *Biochim Biophys Acta* **1808**(6): 1493-1501.
- Lu, Z., W. D. Van Horn, J. Chen, S. Mathew, R. Zent and C. R. Sanders (2012). "Bicelles at low concentrations." *Mol Pharm* **9**(4): 752-761.
- MacKenzie, K., J. Prestegard and D. Engelman (1997). "A transmembrane helix dimer: structure and implications." *Science* **276**(5309): 131-133.
- Smith, S., R. Jonas, M. Braiman and B. Bormann (1994). "Structure and orientation of the transmembrane domain of glycoporphin A in lipid bilayers." *Biochemistry* **33**(20): 6334-6341.
- Smith, S., D. Song, S. Shekar, M. Groesbeek, M. Ziliox and S. Aimoto (2001). "Structure of the transmembrane dimer interface of glycoporphin A in membrane bilayers." *Biochemistry* **40**(22): 6553-6558.

