

CHARACTERIZATION OF THE MOBILE LOOPS IN GM2 ACTIVATOR PROTEIN AND
THE FLAP REGION OF HIV-1 PROTEASE VIA ELECTRON PARAMAGNETIC
RESONANCE TECHNIQUES

By

JEFFREY DAVID CARTER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

© 2012 Jeffrey David Carter

To my grandmother, Nancy, whose love and support I will always cherish; to my Mom and Dad for their encouragement throughout my years of studies.

ACKNOWLEDGMENTS

I would like to especially thank my grandmother Nancy Como for her unconditional love and support, and my parents Diane and David Carter for being wonderful role models. Next, I would like to express thanks to my advisor, Dr. Gail Fanucci for her diligence in training me to become a scientist. I would also like to thank all of members of my doctoral committee, Doctors Nicole Horenstein, Mavis McKenna, Nick Polfer, and David Powell. I would like to thank Alex Angerhofer for time on the E-580 instrument.

I would like to thank all the past and present members of the Fanucci group for their help, friendship, and most importantly keeping me sane through the graduate school experience. In particular, I like to express sincere thanks to Dr. Jamie Kear, Dr. Natasha Pirman, Dr. Jared Lynch, Adam Smith, Dr. Austin Turner, Ian deVera, and Dr. Mandy Blackburn. Each you made graduate school much more enjoyable, and I will always cherish your friendships. I would also like to thank my undergraduate mentee Peter D'Amore for sticking with me through our oftentimes-baffling experiments.

I wish to express my gratitude to Drs. Alan G. Marshall and Jeremiah Tipton of the National High Magnetic Field Laboratory in Tallahassee Florida and Christine Schubert-Wright of the Department of Pharmacology, X-ray Laboratory and University of Virginia Health System for early help with the GM2AP project. Christine and her post-doc, Fang Xu, generated six of the single CYS mutants along with the two double mutants. Finally, I would like to thank NIH AIDS Reagents Program for free access to HIV-1 protease inhibitors, and the National High Magnetic Field Lab IHRP, NIH, NSF, and UF Startup for funding. The work described herein was supported by NSF MBC-0746533 and ARI DMR-9601864, NIH R37 AI28571, the UF Center for AIDS Research, NHMFL-IHRP, NIH R01 GM077232 awarded to Gail E. Fanucci.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	8
LIST OF FIGURES.....	9
LIST OF ABBREVIATIONS.....	15
ABSTRACT	23
 CHAPTER	
1 INTRODUCTION TO THE BIOLOGICAL SYSTEMS: GM2 ACTIVATOR PROTEIN AND HIV-1 PROTEASE.....	25
Scope of the Dissertation.....	25
Introduction to GM2 Activator Protein	30
Monosaccharide Naming and Abbreviation.....	30
Glycolipids, Glycosphingolipids, and Gangliosides.....	30
Glycosphingolipid Degradation.....	32
Sphingolipid Activator Proteins.....	32
Saposin A.....	33
Saposin B.....	34
SAP C	35
SAP D.....	35
GM2 activator protein.....	36
β -Hexosaminidase A	37
Lipid Storage Diseases and Diseases Associated with Abnormal Sphingolipid Metabolism	37
Tay-Sachs disease	37
Sandhoff disease	38
Introduction to HIV-1 Protease.....	38
Introduction to Human Immunodeficiency Virus	38
Genetic Variability	39
HIV-1 Protease.....	40
Conformational Sampling and the Conformational Ensemble of the HIV-1 Protease Flaps	41
Inhibitors of HIV-1 Protease	41
2 INTRODUCTION TO BIOPHYSICAL METHODOLOGIES	57
Introduction	57
Mass Spectrometry.....	57
Fluorescence Spectroscopy.....	59
Circular Dichroism Spectroscopy.....	60

Electron Paramagnetic Resonance Spectroscopy	61
Site-Directed Spin-Labeling	61
(1-Oxyl-2,2,5,5-Tetramethyl- Δ 3-Pyrroline-3-Methyl) Methane- thiosulfonate (MTSL)	62
4-Maleimido-TEMPO (MSL)	63
Non-native amino acid p-acetyl-L-phenylalanine spin label	63
Continuous-Wave Electron Paramagnetic Resonance Spectroscopy (CW EPR)	64
Introduction and history	64
CW EPR sample requirements	65
CW EPR theory	65
Nitroxide spectral line shapes	66
Pulsed Electron Paramagnetic Resonance Spectroscopy	68
Introduction and history	68
Sample requirements	69
DEER theory	69
E580 spectrometer DEER setup	75
3 SURFACE LOOP CONFORMATIONAL FLEXIBILITY OF THE GM2 ACTIVATOR PROTEIN IS NOT ALTERED BY LIGAND BINDING BUT IS NECESSARY FOR LIPID LIGAND EXTRACTION FROM BILAYER SURFACES	100
Introduction	100
Materials and Methods	103
Materials	103
Methods	103
Expression of GM2AP constructs	104
GM2AP purification protocol	104
Spin-labeling	105
Crosslinking	106
Results and Discussion	107
4 MASS SPECTROMETRIC ANALYSIS AND CONFIRMATION OF DISULFIDE BONDING AND REPORTER SITE LABELING EFFICIENCY IN GM2AP	134
Introduction	134
Methods	135
Direct Infusion and Reversed-Phase Capillary Liquid Chromatography	135
Mass Spectrometry	135
Sequential Digestion with Trypsin, GluC, and AspN	138
Data Analysis and Informatics	138
Results and Discussion	139
5 DOUBLE ELECTRON-ELECTRON RESONANCE STUDIES OF HIV-1 PROTEASE	154

Introduction and Previous Work.....	154
Materials and Methods.....	160
Materials.....	160
Methods.....	161
Expression of HIV-1 protease	161
Purification of HIV-1 protease	161
Incorporation of the unnatural amino acid	164
DEER experiments	165
DEER data analysis	165
Population validation	165
Results and Discussion.....	166
Effect of Sample Freezing Conditions on the Conformational Ensemble of HIV-1PR.....	166
Single Point Drug-Induced Mutations Confer Alterations in the Conformational Ensemble of HIV-1 Protease.....	168
Electron Paramagnetic Resonance Studies of HIV-1 Protease Via Incorporation of the Unnatural Amino Acid Acetyl Phenylalanine.....	172
6 FUTURE DIRECTIONS	209
APPENDIX: CW AND PULSED EPR PARAMETERS	212
LIST OF REFERENCES	213
BIOGRAPHICAL SKETCH.....	227

LIST OF TABLES

<u>Table</u>		<u>page</u>
1-1	Common monosaccharide abbreviations.	43
1-2	Representative diseases associated with abnormal sphingolipid metabolism....	43
1-3	A comparison of HIV-1 and HIV-2.	43
1-4	FDA-approved drugs for the treatment of HIV/AIDS.....	43
3-1	Summary of Double CYS R1 GM2AP Construct Distance Experiments	119
3-2	Differential Scanning Calorimetry Values For GM2AP Constructs	119
3-3	GM2AP Cross-linking Ligand Extraction Data	119
5-1	Distance parameters for HIV-1PR C (apo) and various freezing conditions.	174
5-2	Distance parameters for HIV-1PR C (IDV) and various freezing conditions.	174
5-3	Distance parameters for HIV-1PR (TPV) and various freezing conditions.....	175
5-4	Distance parameters for HIV-1PR (CaP2) and various freezing conditions.....	175
5-5	Summary of distance parameters obtained from DEER EPR of HIV-1PR.....	176
A-1	Typical CW EPR parameters used for work reported within.....	212
A-2	Typical pulsed EPR parameters used for DEER reported within.	212
A-3	4-pulse table used for DEER experiments performed in deuterated matrices. .	212

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Glycolipid structure, where Y = lipid	44
1-2 General structure of a glyceroglycolipid.....	44
1-3 General structure of a glycosylphosphatidylinositol	44
1-4 Structure of sphingosine	45
1-5 Glycosphingolipid structure.....	45
1-6 General structure of a ganglioside.....	45
1-7 Ganglioside GM2 structure.....	46
1-8 Glycosphingolipid degradation pathway	47
1-9 Crystal structure of Saposin A (PDB ID 2DOB).	48
1-10 Unit cell crystal structure of Saposin B (PDB ID 1N69)	48
1-11 Unit cell crystal structure of Saposin C (PDB ID 2QYP)	48
1-12 Unit cell crystal structure of Saposin D (PDB ID 2RB3).....	49
1-13 Representation of ganglioside hydrolysis	49
1-14 Cartoon structure of GM2AP showing the 4 disulfide bonds in red (PDB 1G13).	50
1-15 Cartoon representation of GM2AP binding lipid membranes and extracting its specific substrate GM2 (PDB 1G13).....	50
1-16 Structure of β -hexosaminidase A (PDB ID 2GJX).	51
1-17 HIV-1 life cycle.....	52
1-18 Cartoon representation of the gag and gag-pol polyproteins cleaved by HIV-1 protease during viral maturation.	53
1-19 Structure of Subtype B HIV-1 Protease (PDB ID 2BPX).....	53
1-20 Flap conformations of HIV-1 protease captured by molecular dynamics simulations.....	54
1-21 Points of inhibition within the HIV-1 viral life cycle.....	55

1-22	Structures of the FDA-approved protease inhibitors.....	56
2-1	Schematic diagram of a simple mass spectrometer with a sector-type mass analyzer.....	91
2-2	Schematic diagram of a standard fluorescence spectrometer.....	91
2-3	Schematic diagram of a standard circular dichroism instrumental.....	92
2-4	Effect of choice of spin label on the derivative EPR line shape of HIV-1PR Subtype B _{si}	92
2-5	Site-directed spin-labeling scheme showing the reaction of MTSL with a free thiol group of an engineered CYS residue.....	93
2-6	Structure of MSL spin label.	93
2-7	Spin-labeling scheme for the incorporation of the K1 spin label moiety.....	94
2-8	Scheme for electron transition in a magnetic field	94
2-9	Cartoon representation of EPR.	95
2-10	Dependence of EPR spectral line shape on motion, correlation time, and degree of anisotropy.....	96
2-11	The 2 + 1 pulse DEER sequence.	96
2-12	The 3-pulse DEER sequence	96
2-13	The 4-pulse DEER sequence	97
2-14	Absorption spectra for a nitroxide spin label.....	97
2-15	Sample dipolar evolution curve showing locations of raw dipolar modulation	98
2-16	Selection of regularization parameter in TKR DEER analysis method	98
2-17	Examples of processed DEER data.	99
3-1	Structures of GM2AP and MTSL.	120
3-2	The three unique monomers (A, B, C) of apo GM2AP (PDB ID 1G13).	120
3-3	The red spheres in monomer A of GM2AP (PDB 1G13) represent the C α positions of the reporter sites chosen for SDSL EPR experiments.....	121
3-4	Stack plot of EPR line shapes from six sites in the flexible loops of GM2AP....	121

3-5	DEER data from GM2AP double mutants.	122
3-6	GM2AP construct I72C/S130C long pass filtered and background subtracted dipolar echo curve	123
3-7	GM2AP construct I72C/S130C bound to GM2 long pass filtered and background subtracted dipolar echo curve	124
3-8	GM2AP construct I66C/S126C long pass filtered and background subtracted dipolar echo curve	125
3-9	GM2AP construct I66C/L126C bound to GM2 long pass filtered and background subtracted dipolar echo curve	126
3-10	Differential scanning calorimetry for GM2AP constructs A60R1 and L126R1..	127
3-11	Overlay of area normalized EPR spectra and MOMD simulated spectra of the ten spin labeled GM2AP constructs.....	128
3-12	GM2AP chromatograms from protein purification.	129
3-13	Representative EPR spectra of GM2AP L126R1 and I66R1 as a function of temperature	130
3-14	Linear regression of the ratio of the fractions of mobile and immobile L126R1 populations plotted as a function of inverse temperature	131
3-15	Representative diagram of the dansyl-DHPE fluorescent lipid extraction assay	132
3-16	Tethering results from double mutant GM2AP constructs	133
3-17	Representative diagram of the gel filtration GM2 lipid extraction assay. Gel filtration assay for GM2 extraction	133
4-1	16.5% Tris-HCL Biorad pre-cast gel of GM2AP constructs purified for MS analysis	146
4-2	Broadband 14.5 T electrospray positive-ion of intact wild-type, L126C-MSL, I66C-MSL, N136C-MSL, I66C/L126C-MSL	147
4-3	Parent ion FTICR data for GM2AP-MSL	148
4-4	Protease digestion scheme for GM2AP L126-MSL.	149
4-5	FTICR results for GM2AP L126-MSL peptide.....	150
4-6	Scheme presenting the theoretical disulfide bound peptides after sequential proteolytic digestion.....	151

4-7	CID MS/MS disulfide identification.....	152
4-8	ECD MS/MS disulfide identification.	153
5-1	HIV crystal structure	176
5-2	DEER results of subtype B HIV-1Pr with and without the FDA-approved inhibitor.....	177
5-3	DEER experiments are typically run at cryogenic temperatures.....	177
5-4	Locations of polymorphisms in PRE and POST, with respect to Subtype B (LAI).....	178
5-5	Unnatural amino acid p-acetylphenylalanine and structurally similar amino acid tyrosine.	178
5-6	Chemical reaction for Incorporation of unnatural amino acid to form spin-labeled side-chain K1.	178
5-7	pSUPAR3 plasmid map.....	179
5-8	Sequence alignment for HIV-1 protease LAI with constructs B _{si} , C _{si} , PRE _i , I63P _{si} , A71V _{si} , and POST _i	179
5-9	HIV-1PR Subtype C apo frozen at -20.....	180
5-10	HIV-1PR Subtype C apo frozen in liquid nitrogen.....	181
5-11	HIV-1PR Subtype C apo frozen in isopentane	182
5-12	HIV-1PR Subtype C apo frozen in isopentane from 37C.....	183
5-13	Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, apo after various freezing conditions	184
5-14	HIV-1PR Subtype C TPV-bound frozen at -20 °C.....	185
5-15	HIV-1PR Subtype C TPV-bound frozen in liquid nitrogen.....	186
5-16	HIV-1PR Subtype C TPV-bound frozen in isopentane	187
5-17	HIV-1PR Subtype C TPV-bound apo heated to 37 °C then frozen in isopentane.....	188
5-18	Stacked distance distribution profiles of HIV-1PR Subtype C K55R1 bound to TPV	189

5-19	HIV-1PR Subtype C IDV-bound frozen at -20. Background subtracted dipolar echo curve.....	190
5-20	HIV-1PR Subtype C IDV-bound frozen in liquid nitrogen.....	191
5-21	HIV-1PR Subtype C IDV-bound frozen in isopentane	192
5-22	HIV-1PR Subtype C IDV-bound apo heated to 37 °C then frozen in isopentane.....	193
5-23	Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, bound with IDV	194
5-24	HIV-1PR Subtype C CA-p2-bound frozen at -20°C.....	195
5-25	HIV-1PR Subtype C CA-p2-bound frozen in liquid nitrogen	196
5-26	HIV-1PR Subtype C CA-p2-bound frozen in isopentane	197
5-27	HIV-1PR Subtype C CA-p2-bound apo heated to 37 °C then frozen in isopentane.....	198
5-28	Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, bound with CA-p2.....	199
5-29	DEER results for drug pressure mutations	199
5-30	Population analysis and molecular dynamics simulations for HIV-1 PR.....	200
5-31	HIV-1PR PRE apo.....	201
5-32	HIV-1PR POST apo.....	202
5-33	HIV-1PR Subtype B L63P apo.....	203
5-34	HIV-1PR Subtype B A71V Apo.....	204
5-35	The relative populations (%) for closed, semi-open, tucked/curled, and wide-open conformations of PRE apo and POST apo.	205
5-36	Mass spectra of spin-labeled PRE.....	205
5-37	Mass spectra of spin-labeled POST	206
5-38	Mass spectra of spin-labeled L63P.....	206
5-39	Mass spectra of spin-labeled A71V	207
5-40	Analysis of mass spectra for HIV-1 protease constructs.	207

5-41	Bsi-Amber. Long pass filtered and background subtracted dipolar echo curve overlaid with the TKR regenerated echo curve from DeerAnalysis.....	208
6-1	HSQC NMR spectra of uniformly labeled ¹⁵ N GM2AP upon addition of 4x molar excess GM2 micelles.....	210
6-2	Crystallization trials of wild-type GM2AP	211

LIST OF ABBREVIATIONS

(+)mRNA	plus-sense messenger RNA
Å	angstrom
AIDS	Acquired Immunodeficiency Syndrome
AmpR	ampicillin resistance
APV	Amprenavir
ARV	antiretroviral drugs
AT	adenine-thymine
ATV	Atazanavir
BEV	baculovirus
BME	β-Mercaptoethanol, 2-Mercaptoethanol
BM(PEG) ₂	1,8-bis-maleimidotriethyleneglycol
BM(PEG) ₃	1,11-bis-maleimidotriethyleneglycol
bp	base pair
BSA	bovine serum albumin
B _{si}	subtype B stabilized inactivated
°C	Celsius
CD	circular dichroism
CD4	Cluster of Differentiation 4
CI	chemical ionization
CID	collision-induced dissociation
CRF	circulating recombinant form
C _{si}	subtype C stabilized inactivated

CW	continuous wave
CYS	cysteine
Da	Dalton
dB	decibel
DEER	double electron-electron resonance
DNA	deoxyribonucleic acid
DRV	Darunavir
DTME	dithio-bismaleimidoethane
<i>E. coli</i>	<i>Escherichia coli</i>
EI	entry inhibitor
Env	envelope
EPR	electron paramagnetic resonance
ER	endoplasmic reticulum
ESEEM	electron spin echo envelope modulation
ESI	electrospray ionization
ESR	electron spin resonance
FDA	Food and Drug Administration
FI	fusion inhibitors
FIV	feline immunodeficiency virus
FPLC	fast protein liquid chromatography
FPV	Fosamprenavir
FRET	fluorescence resonance energy transfer
FT-ICR	Fourier transform ion cyclotron resonance

FWHM	full-width at half-maximum
G	Gauss
Gal	galactose
GalN	galactosamine
GalNAc	<i>N</i> -acetylgalactosamine
GHz	gigahertz
Glc	glucose
GlcA	glucuronic acid
GlcCer	glucosylceramide
GlcN	glucosamine
GlcNac	<i>N</i> -acetylglucosamine
gp 120	glycoprotein 120
gp 41	glycoprotein 41
HAART	Highly Active Antiretroviral Therapy
HIV	Human Immunodeficiency Virus
HIV-1	Human immunodeficiency virus type 1
HIV-1PR	HIV-1 protease
HIV-2	Human immunodeficiency virus type 2
HPLC	high-pressure liquid chromatography
IAP	3-(2-Iodoacetamido)-PROXYL
IASL	4-(2-Iodoacetamido)-TEMPO
IDA	iminodiacetic acid
IDV	Indinavir

II	integrase inhibitor
IHRP	In-house research program
IMAC	immobilized metal ion affinity chromatography
Int	integrase
IP	isopentane
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
J	Joule
K	Kelvin
kD	kilodalton
LB	Luria-Bertani media
LC	liquid chromatography
LN ₂	liquid nitrogen
LPV	Lopinavir
LTR	long terminal repeat
LTQ	linear quadrupole
M	major group
MALDI	matrix assisted laser desorption ionization
MC	Monte Carlo
MCS	multiple cloning site
MD	molecular dynamics
MDR769	Multi-drug resistant variant 769
min	minute

mL	milliliter
mM	millimolar
mol	mole
Mol. wt.	molecular weight
MS	mass spectrometry
mS	milli-Siemens
ms	milli-second
MSL	4-Maleimido-TEMPO
MTSL	(1-oxyl-2,2,5,5-tetramethyl- Δ 3-pyrroline-3-methyl)-methanethiosulfonate
MHZ	megahertz
NaOAc	sodium acetate
NHLBI	National Heart, Lung, and Blood Institute
NHMFL	National High Magnetic Field Lab
NIH	National Institute of Health
NFV	Nelfinavir
Ng	nanogram
NK	natural killer
nL	nanoliter
Nle	norleucine
nm	nanometer
NMR	nuclear magnetic resonance
NNRTI	non-nucleoside reverse transcriptase inhibitor
NOE	Nuclear Overhauser Effect

NRTI	nucleoside reverse transcriptase inhibitor
NSF	National Science Foundation
NTA	nitrilotriacetic acid
O	outlier group
OD	optical density
OI	Opportunistic infection
ORF	open reading frame
PDB	Protein Data Bank
PEG	polyethylene glycol
PI	protease inhibitors
pmol	picomol
ppm	part per million
PMPR	pentamutated protease
R1	MTSL
RNA	ribonucleic acid
rms	root mean square
rpm	rotations per minute
RT	reverse transcriptase
RTV	Ritonavir
SapA	Saposin A
SapB	Saposin B
SapC	Saposin C
SapD	Saposin D

SDS	sodium dodecyl sulfate
SDSL	site-directed spin-labeling
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIV	Simian Immunodeficiency virus
SIVcpz	Simian Immunodeficiency virus, chimpanzee
SIVsm	Simian Immunodeficiency virus, sooty mangabey
SNR	signal to noise ratio
SPR	surface plasmon resonance
SU	surface
SQV	Saquinavir
T	Tesla
TKR	Tikhonov Regularization
T_m	melting temperature
T_M	phase memory time
TM	transmembrane
TPV	Tipranavir
tRNA	transfer RNA
TSD	Tay-Sachs disease
UF	University of Florida
μL	microliter
UNAIDS	Joint United Nations Programme on HIV/AIDS
UV	ultraviolet
UV-Vis	ultraviolet visible spectroscopy

VMD	visual molecular dynamics
WHO	World Health Organization
WT	wild-type

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CHARACTERIZATION OF THE MOBILE LOOPS IN GM2 ACTIVATOR PROTEIN AND
THE FLAP REGION OF HIV-1 PROTEASE VIA ELECTRON PARAMAGNETIC
RESONANCE TECHNIQUES

By

Jeffrey David Carter

August 2012

Chair: Gail E. Fanucci
Major: Chemistry

The work in this dissertation focuses on two unrelated proteins, namely GM2 activator protein (GM2AP) and HIV-1 protease (HIV-1PR), each playing a role in health and disease. GM2AP malfunction can be involved in the lipid storage disease Tay-Sachs, and HIV-1PR is the enzyme responsible for maturation of the human immunodeficiency virus (HIV). Site-directed spin-labeling (SDSL) paired with electron paramagnetic resonance spectroscopy (EPR) is an excellent technique for examining the conformations and conformational changes of a protein, and was utilized here, among other techniques, to gain information about the proteins' respective mobile region.

GM2AP is thought to sample three distinct conformations, although the functional importance of these conformers has not yet been defined. Various X-ray structures with different lipid substrates reveal multiple conformations of a mobile loop region, and these conformations suggest a functional conformation change that may occur upon lipid ligand or vesicle binding. Here, multiple conformations of the mobile loop *in solution* were detected by EPR, indicating that the conformations identified in various crystal

structures are present in solution and may play a functional role in ligand binding or interactions with lipid vesicles. Distance measurements from EPR, however, did not lead to detection of local conformational changes in the loop regions as a result of ligand binding. Furthermore, functional assays showed that immobilization of the loops inhibited the ability of GM2AP to bind its ligand, indicating that although GM2AP may not undergo a global conformational change as a result of ligand binding, loop plasticity is essential to its role as a general lipid transfer protein.

HIV-1PR functions as a homodimer with two mobile loops regulating access of substrate to the active site cavity. The double electron electron resonance (DEER) results reported here for HIV-1PR show, most notably, that single drug-induced mutations even outside of the active site pocket are sufficient to cause a substantial shift in the flap conformations likely playing an important role in drug-resistance and efficacy. The conformational ensemble for the parental subtype B suggests that the flaps favor the semi-open conformation. The addition of the single mutation A71V, and to a slightly lesser extent L63P, shifts the flaps' conformational ensemble to favor a closed conformation. The results indicate a possible mechanism for combining primary and secondary mutations to effect drug resistance, where conformational ensemble equilibria are altered which may then impact enzyme efficiency and altered inhibitor-binding affinities.

CHAPTER 1 INTRODUCTION TO THE BIOLOGICAL SYSTEMS: GM2 ACTIVATOR PROTEIN AND HIV-1 PROTEASE

Scope of the Dissertation

This dissertation describes experiments carried out with two protein systems, namely GM2 activator protein and HIV-1 protease. Chapter 1 is entitled *Introduction to Biological Systems*, and it serves as an introduction to the biological aspects of the research included within. First, the biological system of the GM2 activator protein is described in detail. This includes first describing and providing structures of the glycolipids, glycosphingolipids, and gangliosides. Next, the sphingolipid degradation pathway is described, including relevant enzymes and accessory proteins, and lipid storage diseases that can occur as a result of a genetic malfunction in any number of proteins. Following this is a section describing the structure and function of the sphingolipid activator proteins, including saposins A-D, and the GM2 activator protein, as well as β -hexosaminidase A. The crystal structures of each of these proteins are shown and examined with respect to space group and unit cell. Tay-Sachs and Sandhoff disease are then discussed as two relevant GM2AP-related lipid storage diseases.

Following the introduction to the GM2 activator protein is an introduction to HIV-1 protease. Starting with a discussion of the worldwide impact, human immunodeficiency virus (HIV) and the autoimmune deficiency syndrome (AIDS) are discussed. A brief introduction to the virus and its life cycle are provided, followed by a discussion regarding points of inhibition within the viral life cycle. A summary of current FDA-approved drugs is given, including structures and details about each of the nine FDA-

approved protease inhibitors, namely Indinavir, Ritonavir, Amprenavir (Fosamprenavir), Atazanavir, Darunavir, Saquinavir, Tipranavir, Nelfinivir, and Lopinavir.

Chapter 2, entitled *Introduction to Biophysical Methodologies*, provides a summary of the biophysical techniques utilized here. Several different methodologies are described, including mass spectrometry, fluorescence spectroscopy, circular dichroism spectroscopy (CD), site-directed spin-labeling, continuous-wave (CW) and pulsed electron paramagnetic resonance spectroscopy (EPR). In addition to the introductions given in Chapter 2, each research chapter will have a detailed materials and methods section for all experiments contained within that respective chapter.

Chapter 3 is entitled *Surface Loop Conformational Flexibility of the GM2 Activator Protein is not Altered by Ligand Binding but is Necessary for Lipid Ligand Extraction from Bilayer Surfaces*. Site-directed spin-labeling was utilized to investigate the flexible loop regions of GM2AP. We were able to engineer a ninth and tenth cysteine in the protein and isolate homogeneous samples of spin labeled GM2AP. Analysis of the EPR spectral line shapes and MOMD simulations for spin labels in the disordered chain segment and both loops are consistent with the various conformations seen in X-ray structures. Multiple conformations of the mobile loop *in solution* were detected by EPR, indicating that the conformations seen in the crystal structure are present in solution and may play a functional role in ligand binding or interactions with lipid vesicles. Although, distance measurements from EPR were not able to detect local conformational changes in the loop regions as a result of ligand binding. The halo state of R1 mutants was confirmed by differential scanning calorimetry, which observed a 0.9 °C increase in the thermotropic unfolding of GM2AP constructs with GM2 bound. Furthermore, functional

assays showed that immobilization of the loops inhibited the ability of GM2AP to bind its ligand, indicating that although GM2AP may not undergo a global conformational change as a result of ligand binding, loop plasticity is essential to its role as a general lipid transfer protein.

Chapter 4 is entitled *Mass Spectrometric Analysis and Confirmation of Disulfide Bonding and Reporter Site Labeling Efficiency in GM2AP*. Mass spectrometry enables direct identification of posttranslational modifications, including disulfide bonds (Tipton, Carte et al 2009; Bean, Carr et al, 1992; Badock, Raida et al, 1998; Gorman, Wallis, 2002; Xu, Zhang et al, 2008; Yen, Joshi et al, 2002; Nair, Nilson et al, 2006; Popolo, Ragni et al, 2008). Because of the analytical challenge associated with verifying multiple disulfide bonds, a multitier assay was developed (Tipton, Carter 2009). Following the identification of the intact protein a systematic approach was used to verify proper folding and labeling of a recombinantly overexpressed and MSL-labeled L126C GM2AP construct. Furthermore, verification of disulfide connectivity was performed from sequential trypsin, Glu-C, and Asp-N enzymatic digestions. The data acquired from all three sequential digests by accurate mass measurement MS and MS/MS sequence tags resulted in quick identification of all four disulfide peptides as well as the sequence location of the MSL-Cys. The result is a high-confidence MS assay that will be easy to apply to other spin labeled proteins containing multiple disulfide linkages.

Chapter 5 is entitled *Double Electron-Electron Resonance Studies of HIV-1 Protease*. SDSL DEER was used to further the examination of HIV-1PR flap conformational ensembles. DEER data are generally collected at cryogenic temperatures due to a phase memory time that is too short for detection of the spin

echo at room temperature. The question is often raised on whether the conformations trapped at 65K accurately represent the thermodynamic conformations sampled. Distances between spin-labeled sites in the flap region of an apo HIV-1PR Subtype C construct, as well as CA-p2, Indinavir (IDV) and Tipranavir (TPV)-bound HIV-1PR Subtype C constructs, were measured following four different freezing conditions: a slow -20 °C freeze, an intermediate liquid nitrogen freezing freeze, and rapid freezes from both 25 °C and 37 °C using isopentane in combination with liquid nitrogen. Results indicated very small differences in the captured conformational ensemble of HIV-1PR with varying freezing methods. Additional studies were done to determine the conformational ensembles for four protease constructs, namely PRE and POST, and L63P and A71V secondary mutants. PRE and POST were derived from a pediatric subject infected via maternal transmission with HIV (Ho, Coman et al., 2008) before and after PI treatment; respectively. The *gag-pol* alleles were isolated from serial blood samples obtained over 7 years, starting before therapy initiation (PRE) and after the development of multiple drug resistance following 77 weeks of PI therapy with Ritonavir (RTV) and an additional 16 weeks with Indinavir (IDV) (POST) (Barrie, Perez et al., 1996). PRE contains polymorphisms caused by the normal process of genetic drift. Results indicate that PRE adopts a predominately closed flap conformation, but that the emergence of drug-pressure selected mutations after (POST) causes the conformational sampling to favor a semi-open population. Also investigated were the effects of two secondary mutations, L63P and A71V on the conformational ensemble. L63P is hypothesized to alter the conformational sampling resulting in a more closed state in the PRE construct. A71V, a common secondary mutation seen in multi-drug

resistant constructs (Clemente, Hemraini et al., 2003), is thought to induce a more closed conformation. A minimal amount of early work was done to incorporate and site-directed spin-label an unnatural amino acid into HIV-1PR at position 55 of the flaps. This method differs from the traditional approach to site-directed spin-labeling, in which a thiol-reactive spin label is attached to a native or non-native CYS residue. Previous work from the lab of Wayne Hubbel reported an orthogonal labeling strategy that does not label any functional group found within the 20 naturally occurring amino acids (Fleissner, Brustad et al, 2009). In this method, Fleissner utilized the genetically encoded unnatural amino acid p-acetyl-L-phenylalanine (p-AcPhe) and reacted it with a hydroxylamine reagent to yield a nitroxide side chain termed K1. They were able to demonstrate success using seven different mutants of T4 lysozyme, each containing a single p-AcPhe at a solvent-exposed helix site. It was shown that the EPR spectra of the K1 mutants had higher nitroxide mobilities than the spectra of mutants containing the R1 side chain. Here we were successful in expressing HIV-1 protease K55p-acetyl-L-phenylalanine in large-scale inclusion bodies, purifying, refolding (with mediocre success) in acidic solution, and spin labeling; finally, a low-quality DEER dataset was obtained.

Chapter 6 is called *Future Directions*, and discusses early protein crystallization trials and NMR work on GM2AP. The results from SDSL EPR, DSC, and fluorescence experiments of GM2AP CYS constructs indicate that the flexible loops do not undergo conformational exchange as a result of ligand binding. But, loop mobility is an important attribute in the ability of GM2AP to extract its multiple ligands. To further investigate how ligand binding alters the average conformation of GM2AP, uniformly ^{15}N labeled

protein was generated for solution nuclear magnetic resonance (NMR) heteronuclear quantum coherence (HSQC) measurements. The binding of GM2 does not induce a shift in most of the resonances, only a few peaks are seen to shift or disappear, indicating that GM2AP does not undergo a global conformational shift as a result of ligand binding. NMR assignments for GM2AP are lacking, and future NMR studies are aimed at mapping out the specific regions in GM2AP that are altered during interactions with ligands. In an attempt to overcome the overlapping NMR resonance peaks from the fast conformational exchange sub μ s timescale, crystallizations trials have been started in attempted to gain amino acid specific assignments from solid state NMR.

Introduction to GM2 Activator Protein

Monosaccharide Naming and Abbreviation

Monosaccharide, in Greek, means “single sugar.” These sugars are the most basic form of carbohydrate found in nature. Monosaccharides are named and often abbreviated, and a list of common monosaccharide abbreviations is shown in Table 1-1.

Glycolipids, Glycosphingolipids, and Gangliosides

Glycolipids are part of a family of compounds called the glycoconjugates. They are glycosyl derivatives of lipids with a carbohydrate group attached, as shown in the basic glycolipid structure in Figure 1-1. For more information on glycolipids, the reader is referred to several excellent reviews (Stults, Sweeley et al. 1989). The term glycolipid refers broadly to a compound that contains at least one monosaccharide residue, often more, bound by a glycosidic linkage to a hydrophobic moiety such as a sphingoid or ceramide, among others. These molecules are often found on the surface of cells to provide surface markers for cellular recognition. There exist several subtypes of glycolipids, including glyceroglycolipids, glycosphingolipids, and

glycosylphosphatidylinositols. Glyceroglycolipids, which include galactolipids and sulfolipids, are glycolipids with a glycerol backbone. A representative structure is shown in Figure 1-2. Glycosylphosphatidyl inositols (GPIs) are glycolipid molecules that contain saccharides glycosidically linked to an inositol moiety of a phosphatidylinositol or glycoproteins (Ferguson and Williams 1988; Low and Saltiel 1988). The general structure of a GPI is shown in Figure 1-3. GPIs are widespread in nature, and often serve as protein anchors on cell surfaces. Generally, GPI anchors are covalently linked to the C-terminal end of a protein or peptide via an amide linkage (Ferguson and Williams 1988). Glycosphingolipids, which include cerebrosides, gangliosides, globosides, sulfatides, and glycosphosphosphingolipids, are a subtype of glycolipids containing at least one monosaccharide and the amino alcohol sphingosine. The structure of sphingosine is shown in Figure 1-4, and the general structure of a glycosphingolipid is shown in Figure 1-5. Gangliosides, also known as sialoglycosphingolipids, are the most complex of the glycolipids. They are composed of glycosphingolipids with one or more sialic acid groups linked to the sugar group. Found predominantly on the surface of nerve cells, many different gangliosides have been identified, whose structural differences tend to involve the structure of the sugar. The oligosaccharide chains of gangliosides actually contain N-acetylneuraminic acid (NeuNAc), an acetylated derivative of sialic acid, which imparts a negative charge. The ceramide is linked via its C-1 residue to a β -glucosyl residue, which is bound to a β -galactosyl residue. The general structure of a ganglioside is shown in Figure 1-6. Examples of gangliosides include GM1 and GM2. The names given for gangliosides are a description of their structure. The letter G refers to ganglioside, and the subscripts

M, D, T and Q indicates the number of sialic acids (mono-, di-, tri and quatra-sialic acid). The numerical subscripts 1, 2 and 3 refer to the carbohydrates attached to the ceramide, where 1 equals GalGalNAcGalGlc-ceramide, 2 equals GalNAcGalGlc-ceramide, and 3 equals GalGlc-ceramide. The structure of ganglioside GM2 is shown in Figure 1-7.

Glycosphingolipid Degradation

Glycosphingolipids are degraded by exoglycosidases and hydrolases. In addition to these enzymes, non-catalytic sphingolipid activator proteins are required to activate and deliver the substrate to the enzyme for cleavage. Some activator proteins function by binding to GSLs and forming aqueous-soluble complexes, which provide aqueous-soluble enzymes access to a lipid that would otherwise be in a hydrophobic environment. Figure 1-8 shows a diagram of part of the glycosphingolipid degradation pathway. For more information, readers are referred to many excellent reviews (Sandhoff and Kolter 1996).

Degradation of the glycosphingolipids takes place in lysosomal compartments within the cell (Sandhoff and Kolter 1995). GSLs are endocytosed, trafficked and sorted through early and late endosomal compartments on the way to the lysosome, where specific enzymes sequentially cleave sugar groups, eventually producing ceramide, as shown in Figure 1-8, which is finally deacylated to sphingosine (Griffiths, Hoflack et al. 1988). More than ten different enzymes are involved in this degradation process.

Sphingolipid Activator Proteins

Saposins, an acronym for sphingolipid activator proteins (SAPs), function to activate glycosphingolipid degradation by interacting with lipid molecules and presenting them for hydrolysis. In other words, SAPs are non-enzymatic accessory proteins

required for sphingolipid hydrolysis by specific hydrolases. There exist five known SAPs, including saposins A-D and GM2 activator protein (Darmoise, Maschmeyer et al. 2010). For an excellent review of the sphingolipid activator proteins, one can refer to a 2010 *Advances in Immunology* chapter called *The Immunological Functions of Saposins*, by Alexandre Darmoise et al. (Darmoise, Maschmeyer et al. 2010).

The *PSAP* gene encodes for a highly conserved 524 amino acid precursor glycoprotein called prosaposin, which is proteolytically processed to liberate four cleavage products called SAPs A-D (Furst, Machleidt et al. 1988; O'Brien, Kretz et al. 1988). These proteins have a high level of homology. Each domain contains approximately 80 amino acid residues, each with two prolines, six cysteine residues, nearly identical disulfide connectivity, and similar sites of glycosylation, with one in saposins B, C, and D, and two in saposin A. Saposins A-D are mostly α -helical in structure, as seen in the ribbon diagrams shown in Figures 1-9 – 1-12. GM2AP is encoded by different gene. With a β -cup topology, GM2AP is not structurally homologous to SAPs A-D (Darmoise, Maschmeyer et al. 2010).

Saposin A

SAP-A activates the hydrolysis of 4-methylumbelliferyl- β -glucoside, glucocerebroside, and galactocerebroside (Morimoto, Martin et al. 1989). SAP-A deficiency leads to the late-onset lipid storage disease called globoid cell leukodystrophy that is very similar to Krabbe disease, a condition caused by a deficiency in galactosylceramide β -galactosidase (Morimoto, Martin et al. 1989; Matsuda, Vanier et al. 2001; Spiegel, Bach et al. 2005; Ahn, Leyko et al. 2006; Locatelli-Hoops, Remmel et al. 2006; Darmoise, Maschmeyer et al. 2010). SapA has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.00

Angstroms and a space group described as $P 2_1 2_1 2$. A ribbon diagram of the crystal structure is shown in Figure 1-9.

Saposin B

Saposin B was the first of the saposin activator proteins to be discovered, and is often considered a non-specific activator protein (Mehl and Jatzkewitz 1964). It is necessary for the hydrolysis of sulfatides by arylsulfatase A, as well as Gb3 and Ga2Cer by α -galactosidase A, and LacCer by galactosylceramide β -galactosidase (Wenger, DeGala et al. 1989; Schlote, Harzer et al. 1991; Ahn, Faull et al. 2003; Darmoise, Maschmeyer et al. 2010). SapB also works with GM2AP for degradation by GM2 by β -hexosaminidase A (Wilkening, Linke et al. 2000). As an added function, SapB also binds phosphatidylinositol and transfers the phospholipid between biological membranes (Ciaffoni, Tatti et al. 2006). Defects in SapB can lead to several lipid storage disorders, such as metachromatic leukosystrophy (MLD) (Li, Kihara et al. 1985; Sun, Witte et al. 2008). SapB has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.20 Angstroms (Ahn, Faull et al. 2003). A ribbon diagram of the crystal structure is shown in Figure 1-10. The protein is a homodimer with the lipid-binding cavity found between the monomers. The space group is described as $P 3_1 2 1$ (Ahn, Faull et al. 2003). The unit cell contains three independent peptide chains called A, B, and C (depicted as white, grey, and black, respectively, in Figure 1-10) that form two distinct homodimers. The first dimer, called dimer AB is formed by chains A and B (white and grey, respectively), and the second dimer, called dimer CC' is formed by two C chains. Similar to other members of the Saposin family, the V-shaped fold of each monomer is stabilized by two disulfide bonds (Ahn, Faull et al. 2003).

SAP C

Saposin C was the second saposin to be discovered and stimulates the hydrolysis of glycosphingolipids by glycosylceramidase and galactosylceramidase (Ho and O'Brien 1971). Defects in SapC can lead to a juvenile form of the lipid storage disease called Gaucher disease, where B-GlcCer is accumulated (Ho and O'Brien 1971). SapC has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.45 Angstroms (Rossmann, Schultz-Heienbrok et al. 2008). A ribbon diagram of the crystal structure is shown in Figure 1-11. The protein is a homodimer with boomerang-shaped monomers and a space group described as $C222_1$ (Rossmann, Schultz-Heienbrok et al. 2008).

SAP D

Saposin D is arguably the least-studied saposin to date but is by far the most abundant saposin in normal tissue (O'Brien and Kishimoto 1991). SapD activates the hydrolysis of ceramide by acid ceramidase. Defects in SapD lead to accumulation of α -hydroxy fatty acid ceramides (Matsuda, Kido et al. 2004). SapD is also a membrane disrupter. By binding membranes which contain anionic lipids such as BMP and PI, SapD facilitates small vesicle formation (Ciaffoni, Salvioli et al. 2001). SapD has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.10 Angstroms (Rossmann, Schultz-Heienbrok et al. 2008). A ribbon diagram of the crystal structure is shown in Figure 1-12. The protein is a homodimer that exists as either a substrate-free closed helix bundle or in an open ligand-bound conformation with the lipid binding cavity found between the monomers (Rossmann, Schultz-Heienbrok et al. 2008).

GM2 activator protein

The GM2 Activator Protein (GM2AP) is an 18 kDa accessory protein that plays a fundamental and essential role in the specific catabolism of the ganglioside GM2 to GM3, as shown in Figure 1-13 (Conzelmann and Sandhoff 1979; Furst and Sandhoff 1992). The high affinity of GM2AP for GM2 is based on recognition of the oligosaccharide moiety and the ceramide tail. GM2AP is thought to bind GM2 in intralysosomal vesicles and present the oligosaccharide head group for hydrolytic cleavage. β -Hexosaminidase A (Hex A) specifically cleaves the GalNAc sugar group from the ceramide, producing GM3. Because GM2AP does not actually catalyze a chemical reaction, it is considered an accessory protein, not an enzyme, in the catabolism of GM2. Mutations in GM2AP lead to an accumulation of GM2 in the lysosomes, causing the lysosomal storage diseases Tay Sachs or the AB variant of Sandhoff's disease (Conzelmann and Sandhoff 1979). GM2AP has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.00 Angstroms (Wright, Li et al. 2000). A ribbon diagram of the crystal structure is shown in Figure 1-14, and the positions of the four disulfide bridges are shown in red. The protein functions as a monomer. The space group is described as $P 2_1 2_1 2_1$ and the unit cell contains four separate monomers (Wright, Li et al. 2000). The structure consists of an eight-stranded antiparallel β -cup with an accessible hydrophobic cavity large enough for binding large lipid acyl chains (Wright, Li et al. 2000).

Two possible mechanisms for GM2AP function have been presented (Sandhoff and Kolter 1996). The first describes GM2AP acting as a "liftase," where GM2 is recognized and lifted from the bilayer thereby presenting it to the water-soluble enzyme for cleavage. However, it is also possible that GM2AP binds the GM2 ligand and forms

a water soluble complex whereby the enzymatic reaction would take place free in solution. A cartoon model of the latter hypothesis is presented in the cartoon representation in Figure 1-15. Note that the red circles denote a conformational change in the mobile loop region from an open conformation to a closed conformation.

β -Hexosaminidase A

β -Hexosaminidase A (HexA) is found in the lysosomes and is enzyme responsible for the degradation of the GM2 ganglioside. HexA has been crystallized and the crystal structure solved by X-Ray diffraction to a resolution of 2.80 Angstroms (Lemieux, Mark et al. 2006). A ribbon diagram of the crystal structure is shown in Figure 1-16. The protein functions as a heterodimer. The space group is described as C2 and the unit cell contains four separate monomers (Lemieux, Mark et al. 2006).

Lipid Storage Diseases and Diseases Associated with Abnormal Sphingolipid Metabolism

Lipid storage diseases, also referred to as gangliosidoses, are brought on by mutations in genes coding for enzymes involved in ganglioside metabolism. The enzyme deficiency causes accumulation of the corresponding sphingolipid in the cell, disrupting cellular function. A list of representative diseases associated with abnormal sphingolipid metabolism is given in Table 1-2.

Tay-Sachs disease

Tay-Sachs disease (TSD), also referred to as the B-variant (or type 1) of GM2 gangliosidosis, or Hexosaminidase A deficiency, is a very rare, inherited, progressive neuronal disorder. The disease can onset in infants, adolescence, or adulthood, but is far more common and severe during infancy. Symptoms vary widely, but include slow development, loss of motor skills, vision and hearing loss, mental disability, and

paralysis. A cherry-red spot in the eye is characteristic of Tay-Sachs disease. Children with Tay-Sachs disease usually live only into early childhood.

Mutations in the HexA gene, which encodes for an enzyme called β -hexosaminidase A, are responsible for the disease. The condition is inherited in an autosomal recessive pattern (both copies of the gene harbor mutations). These genetic mutations are found most frequently in people of Ashkenazi Jewish heritage. β -hexosaminidase A is responsible for breaking down the GM2 ganglioside. If the function of this lysosomal enzyme is hindered, GM2 will accumulate to toxic levels in the neurons of the brain and spinal cord, leading to neuronal destruction.

Sandhoff disease

Sandhoff disease, also called Hexosaminidase A and B deficiency or Jatzkewitz-Pilz syndrome, is a rare genetic lipid storage disease. The autosomal recessive disorder is linked to a buildup of GM2 ganglioside, which is toxic at high levels, which leads to progressive damage to the central nervous system via destruction of nerve cells. Sandhoff disease can onset in infants, juveniles, or adults, but is most common during infancy. Symptoms of Sandhoff disease are nearly identical to those of Tay-Sachs disease, and include mental disability, muscle deterioration, blindness, deafness, seizures, and cherry red spots in the retina. Currently, Sandhoff disease does not have any treatment or cure.

Introduction to HIV-1 Protease

Introduction to Human Immunodeficiency Virus

Human immunodeficiency virus (HIV), a member of the retrovirus family, is the cause of Acquired Immunodeficiency Syndrome (AIDS), a disease characterized by a suppression of the immune system (Weiss 1993). There are two strains of HIV called

HIV-1 and HIV-2, with HIV-1 being far more common (Gilbert 2003). A comparison of these HIV strains is given in Table 1-3 (Reeves and Dorns 2002; Gilbert 2003). Furthermore, HIV-1 is categorized into different groups (or different viral lineages), subtypes (taxonomic groups), and circulating recombinant forms (CRFs) (recombinant viral forms) (Kantor, Shafer et al. 2005). HIV infection is classified into four stages, the fourth and final called AIDS, which is characterized by CD4 cell count, as well as presence of opportunistic infections (OIs), which are typically normal infections that can become life-threatening for the patient with the compromised immune system (Buchbinder, Katz et al. 1994; Lawn 2004). There are many drugs currently used to slow the progression of the disease, but there is currently no cure (Schneider, Gange et al. 2005; Douek, Roederer et al. 2009; Noble 2009).

The life cycle of HIV-1, depicted in Figure 1-17, is dependent upon a human host cell harboring the CD4 cell surface receptor. After a targeted fusion event, the virion are uncoated and the contents unloaded into the host cell. The viral RNA is reverse transcribed by viral enzyme, and the resultant DNA is imported into the nucleus where it is integrated into the host cell chromosome, transcribed, and translated (Smith and Daniel 2006). These proteins then associate on the outer membrane of the host cell, triggering a budding event that releases immature viral particles. Maturation takes place when HIV-1 protease cleaves the *gag* and *gag-pol* polyproteins, depicted in Figure 1-18.

Genetic Variability

HIV-1 has a high level of genetic variability, when generally compared to many other types of viruses. This fact is based upon several governing factors, including its fast replication cycle and its high rate of mutation. It has been determined that 10^9 to

10^{10} virions can be generated daily, with approximately 3×10^{-5} mutations per nucleotide base per cycle of replication (Robertson, Hahn et al. 1995; Rambaut, Posada et al. 2004). Because of the wide genetic variability that occurs as a result, HIV is classified into groups. Within the major strain HIV-1, there are three main groups, namely M, N, and O. These classifications are based upon differences in the *env* region of the genome (Thomson, Perez-Alvarez et al. 2002). Group M is by far the most common, and is generally further divided into subtypes. These classifications are based upon differences throughout the entire genome (Carr, Foley et al. 1998). The most common HIV-1 are subtype C, comprising approximately 47% of infection worldwide, Subtype A/CRF02_A/G, comprising approximately 26.7%, Subtype B, comprising approximately 12%, Subtype D, comprising approximately 5%, CRF01_A/E, comprising approximately 3%, and 5% comprising all other subtypes and CRFs (Osmanov, Pattou et al. 2002). The majority of HIV research is performed using Subtype B, which is most commonly found in the United States and Europe (Perrin, Kaiser et al. 2003).

HIV-1 Protease

Human immunodeficiency virus type 1 (HIV-1) protease (HIV-1PR) is a homodimer of two 99 amino acid monomers. The enzyme, an aspartic protease, is required for viral maturation of HIV-1, thus its inhibition is a major target for AIDS antiviral therapy (Ashorn, McQuade et al. 1990). Access of the *gag* and *gag-pol* polyprotein substrates to the active site cavity is mediated by two β -hairpins commonly called the flaps. The protease flaps have been shown to undergo a large conformational change during substrate binding and catalysis (Todd, Semo et al. 1998; Ishima, Freedberg et al. 1999; Todd and Freire 1999; Todd, Luque et al. 2000; Velazquez-Campoy, Todd et al. 2000; Ishima, Ghirlando et al. 2001; Freedberg, Ishima et al. 2002; Ishima, Torchia et al.

2003; Louis, Ishima et al. 2003). Over 300 structures of HIV-1PR have been reported in the Protein Data Bank (PDB), the first of which was reported in 1989 (Miller, Schneider et al. 1989). The structure of HIV-1 (top and side views), in both ribbon and space-filling form, is shown in Figure 1-19. This model also clearly demonstrates that the active site is highly inaccessible to large peptide substrates when the flaps are in the closed conformation. Molecular dynamics simulations have provided some of the most direct evidence of the necessary conformational changes within the flap region to facilitate ligand binding and catalysis. Additionally, molecular dynamics (MD) simulations performed by Simmerling et al. have pointed to three distinct flap conformations in HIV-1 protease, as illustrated in Figure 1-20 (Hornak, Okur et al. 2006; Hornak, Okur et al. 2006).

Conformational Sampling and the Conformational Ensemble of the HIV-1 Protease Flaps

The flaps of HIV-1 protease have been shown to undergo a large conformational change upon binding and catalysis, and the sum of the conformations sampled is described as the conformational ensemble. The preferred conformations of the protease flaps change upon addition of inhibitor or substrate (Galiano, Blackburn et al. 2009), incorporation of naturally occurring or drug-pressure selected polymorphisms into the protease (Kear, Blackburn et al. 2009), or the conditions of the protein solution are changed [Blackburn, unpublished].

Inhibitors of HIV-1 Protease

Several points throughout the HIV-1 life cycle are exploited for inhibition, including viral fusion, reverse transcription, integration, and maturation steps, as shown in Figure 1-21. There are currently 31 antiretroviral drugs (ARVs) approved by the U.S. Food and

Drug Administration (FDA) for Highly Active Antiretroviral Therapy (HAART) treatment of HIV infection, and these are shown in Table 1-4. The ARVs are classified into fusion inhibitors (FIs), nucleoside and non-nucleoside reverse transcriptase inhibitors (NRTIs and NNRTIs), integrase inhibitors, and protease inhibitors. HAART treatment usually consists of combined treatment with several types of inhibitors.

The research discussed within this dissertation deals specifically with the protease inhibitors. There are nine FDA-approved HIV-1 protease inhibitors currently being used in the treatment of HIV-1 infection, namely Ritonavir, Indinavir, Lopinavir, Darunavir, Saquinavir, Tipranavir, Atazanavir, Fosamprenavir (the pro-drug of Amprenavir) and Nelfinavir. Structure of these protease inhibitors are given in Figure 1-22. Protease inhibitors, first developed in 1989, generally function to target the active site of the protease by mimicking the tetrahedral intermediate of the substrate.

Table 1-1. Common monosaccharide abbreviations

Name	Symbol
<i>N</i> -acetylgalactosamine	GalNAc
<i>N</i> -acetylglucosamine	GlcNac
<i>N</i> -acetylneuraminic acid	NeuAc or Neu5Ac
Galactosamine	GalN
Galactose	Gal
Glucosamine	GlcN
Glucose	Glc
Glucuronic acid	GlcA

Table 1-2. Representative diseases associated with abnormal sphingolipid metabolism

Disorder	Enzyme Deficiency	Accumulating Molecule
Tay-Sachs disease	HexA	GM2
Tay-Sachs AB variant	GM2AP	GM2
Sandhoff disease	HexA and HexB	GM2, globoside
GM1 Gangliosidosis	β -galactosidase-1	GM1
Gaucher disease	glucocerebrosidase	glucocerebroside
Krabbe disease	galactocerebrosidase	galactocerebroside
Fucosidosis	α -fucosidase	pentahexosylfucoglycolipid

Table 1-3. A comparison of HIV-1 and HIV-2

Strain	Virulence	Infectivity	Prevalence	Inferred Origin
HIV-1	High	High	Global	Chimpanzee
HIV-2	Lower	Low	Western Africa	Sooty Mangabey

Table 1-4. FDA-approved drugs for the treatment of HIV/AIDS

PIs	NRTIs	NNRTIs	EIs	IIs
Saquinavir	Delavirdine		Zidovudine	Enfuvirtide
Raltegravir				
Ritonavir	Efavirenz		Emtricitabine	Celsentri
Indinavir	Nevirapine		Lamivudine	
Nelfinavir		Stavudine		
Amprenavir		Abacavir		
Lopinavir		Didanosine		
Atazanavir		Tenofovir		
Tipranavir				
Fosamprenavir				
Darunavir				

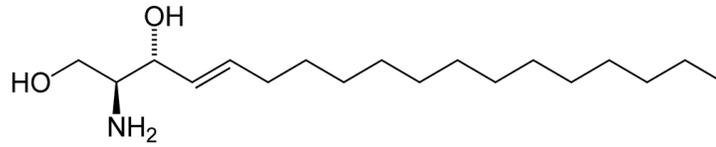


Figure 1-4. Structure of sphingosine

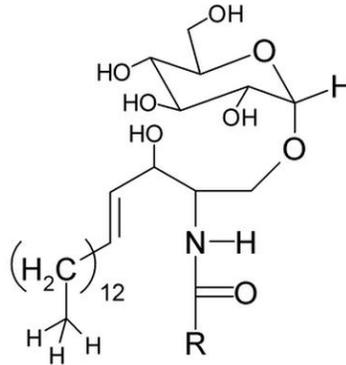


Figure 1-5. Glycosphingolipid structure

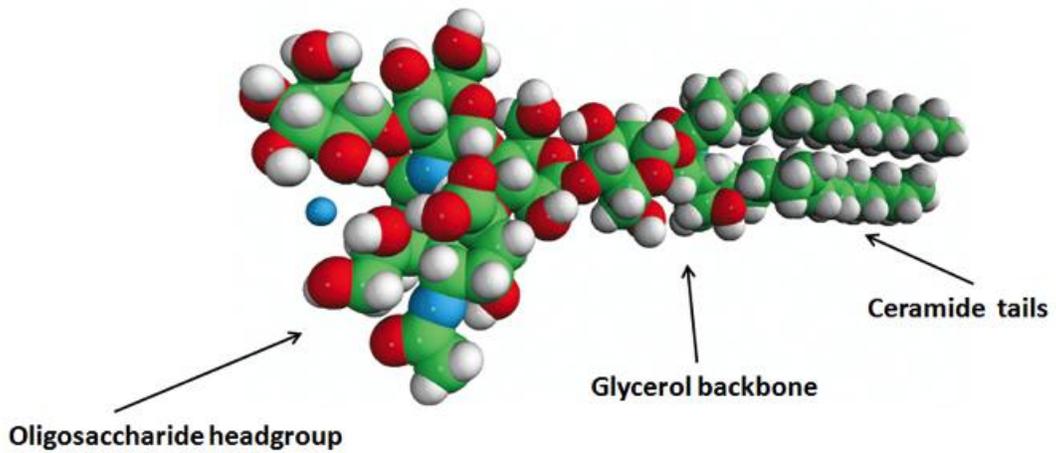


Figure 1-6. General structure of a ganglioside

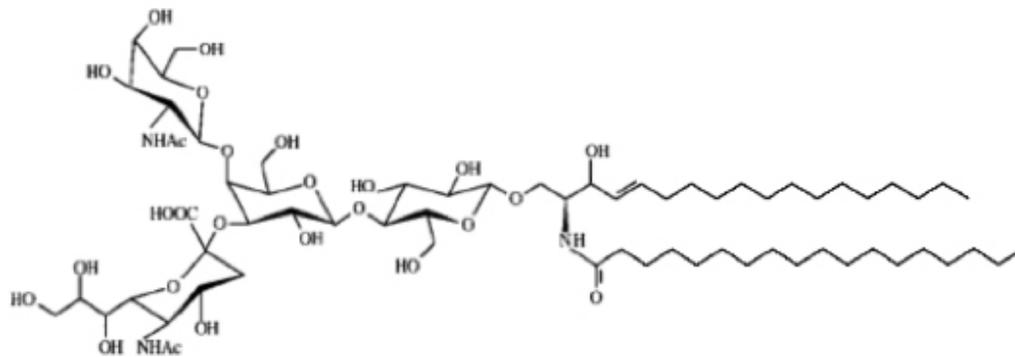


Figure 1-7. Ganglioside GM2 structure

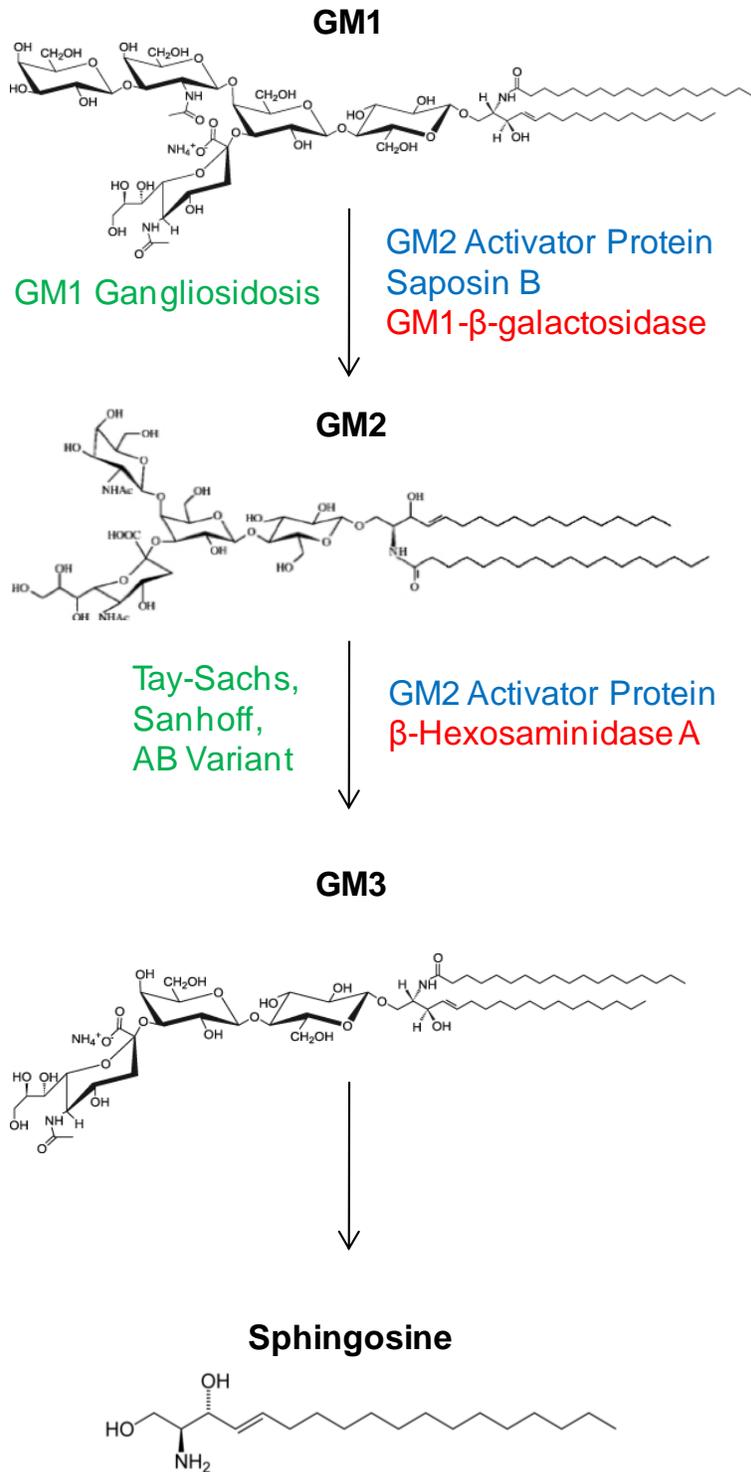


Figure 1-8. Glycosphingolipid degradation pathway. In green are the names of corresponding lipid storage diseases corresponding with each step in the degradation pathway. Likewise, in blue are the names of necessary accessory proteins and in red are the names of enzymes involved in the catabolic reactions.



Figure 1-9. Crystal structure of Saposin A (PDB ID 2DOB). Image rendered in VMD.

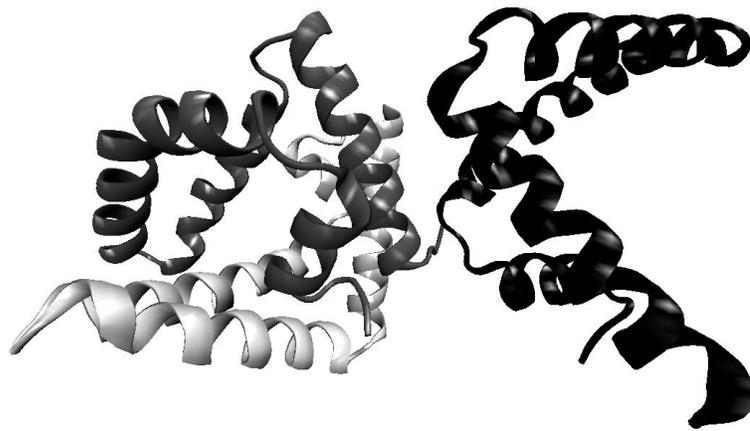


Figure 1-10. Unit cell crystal structure of Saposin B (PDB ID 1N69)

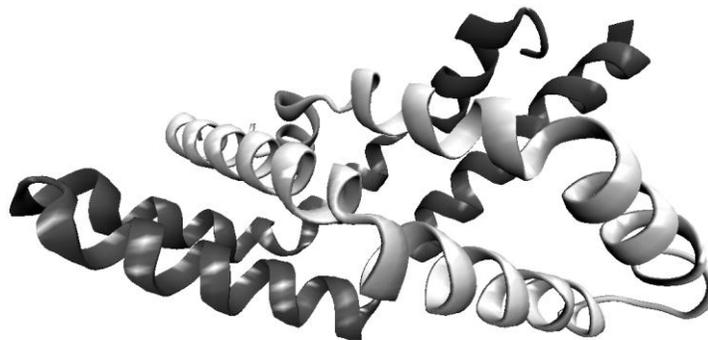


Figure 1-11. Unit cell crystal structure of Saposin C (PDB ID 2QYP)

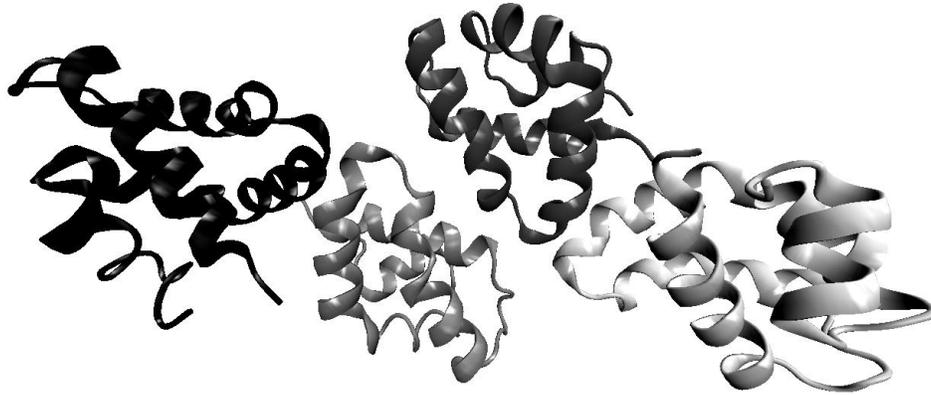


Figure 1-12. Unit cell crystal structure of Saposin D (PDB ID 2RB3)

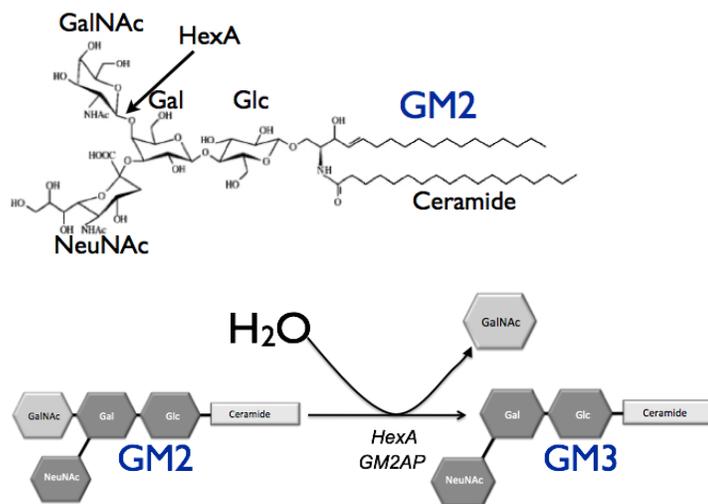


Figure 1-13. Representation of ganglioside hydrolysis. A) GM2 ganglioside lipid structure, B) cartoon scheme for hydrolytic cleavage of GM2 to GM3.

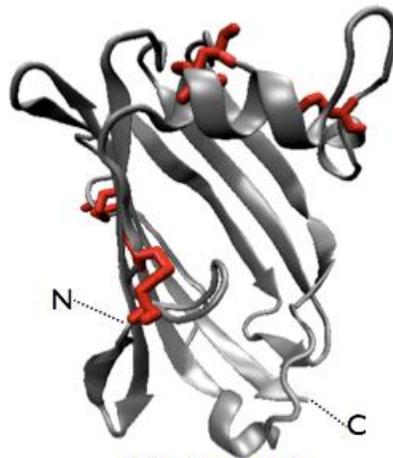


Figure 1-14. Cartoon structure of GM2AP showing the 4 disulfide bonds in red (PDB 1G13).

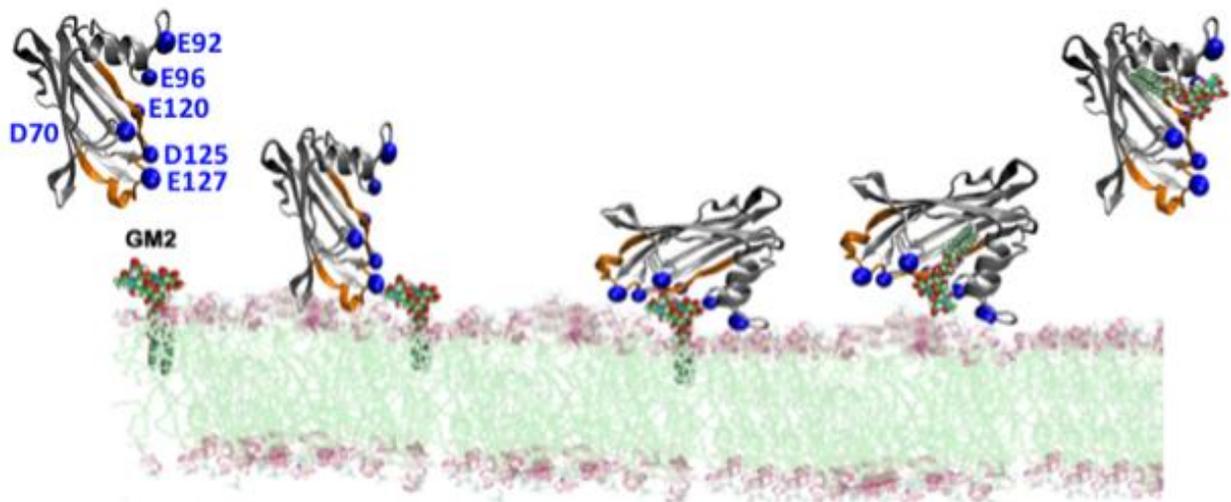


Figure 1-15. Cartoon representation of GM2AP binding lipid membranes and extracting its specific substrate GM2 (PDB 1G13).



Figure 1-16. Structure of β -hexosaminidase A (PDB ID 2GJX).

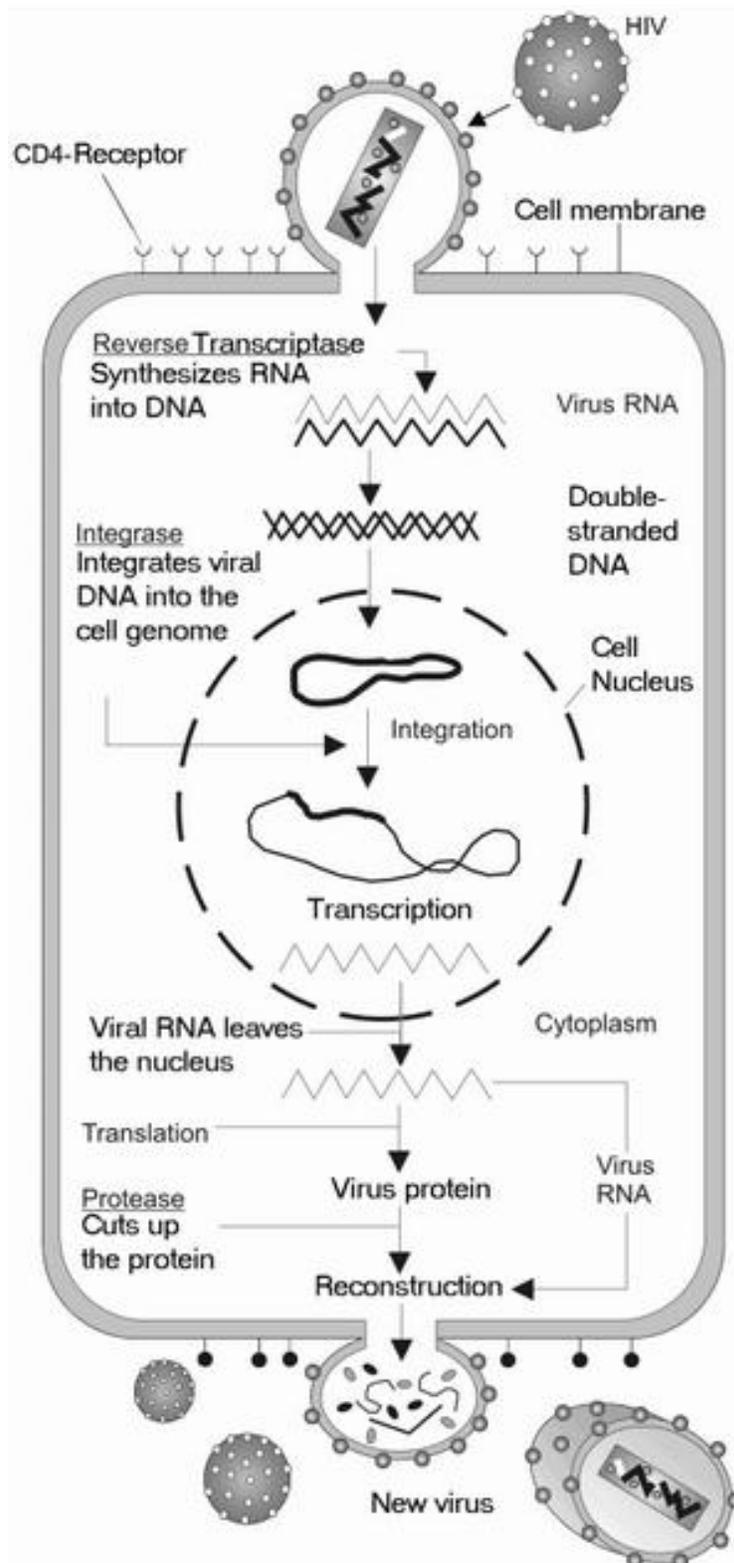


Figure 1-17. HIV-1 life cycle.

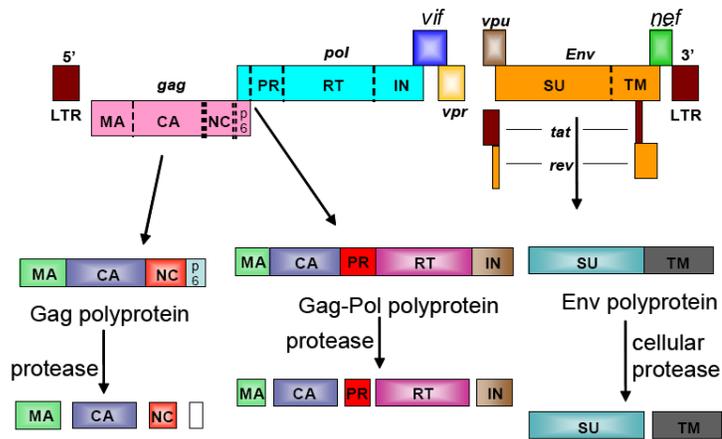


Figure 1-18. Cartoon representation of the gag and gag-pol polyproteins cleaved by HIV-1 protease during viral maturation.

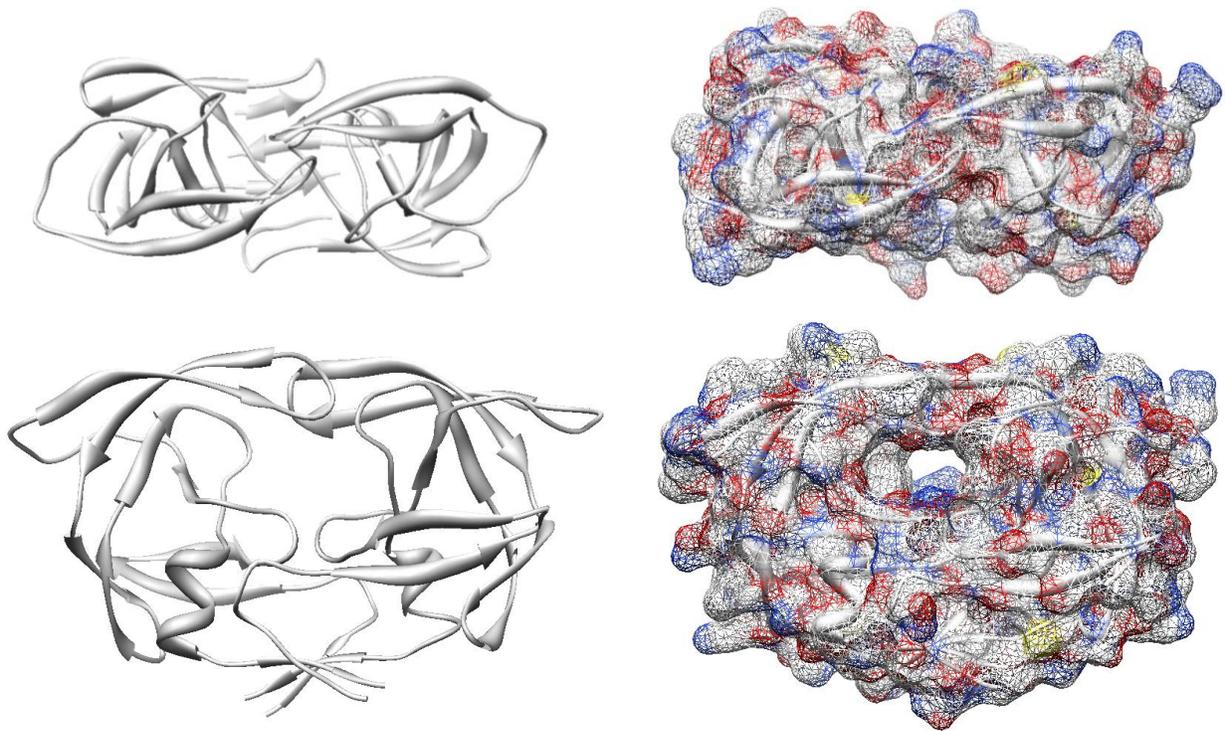


Figure 1-19. Structure of Subtype B HIV-1 Protease (PDB ID 2BPX), A) Ribbon diagram and B) space filling model of top view of Subtype B HIV-1PR. C) Ribbon diagram and D) space filling model of side view of Subtype B HIV-1PR.

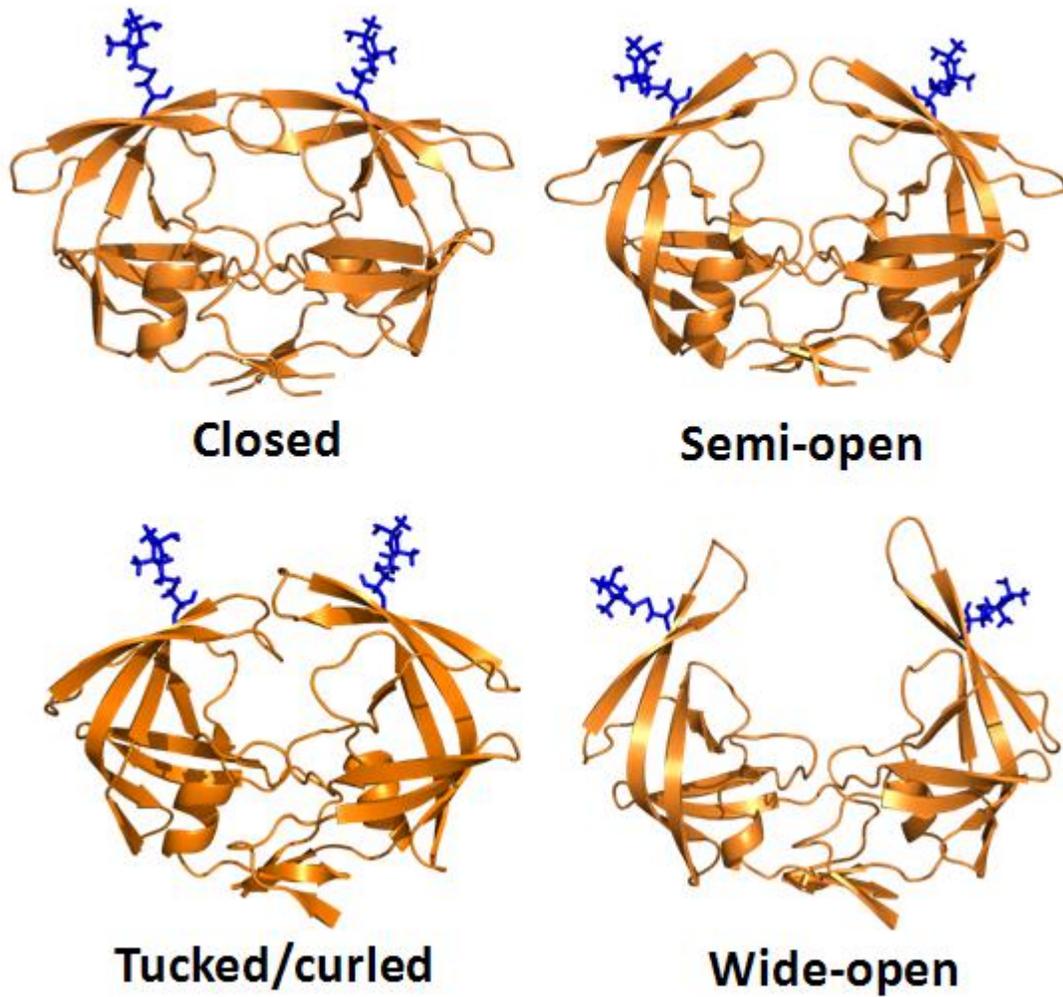


Figure 1-20. Flap conformations of HIV-1 protease captured by molecular dynamics simulations; A) closed, in presence of inhibitor, B) semi-open, and C) wide-open.

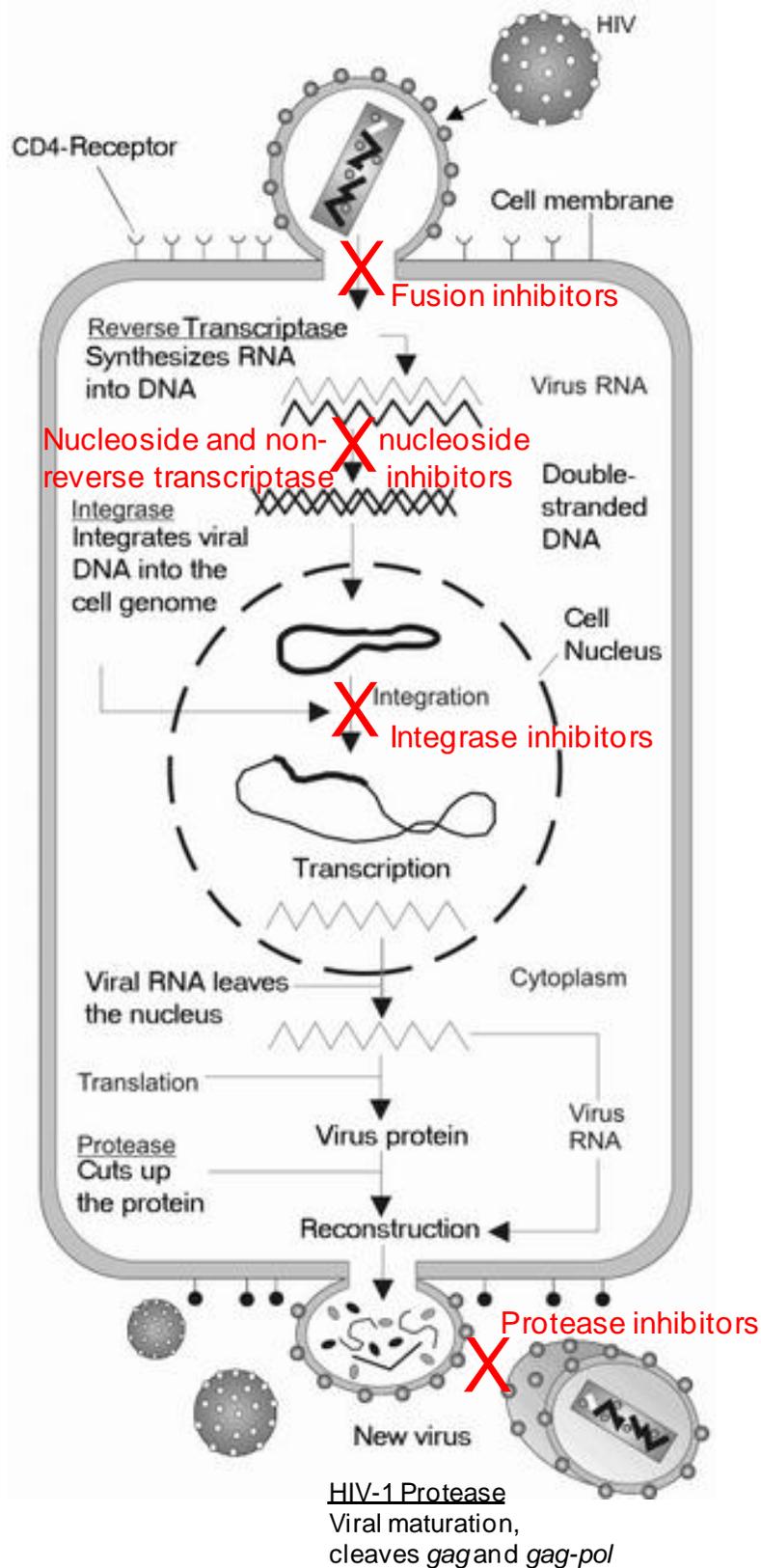


Figure 1-21. Points of inhibition within the HIV-1 viral life cycle.

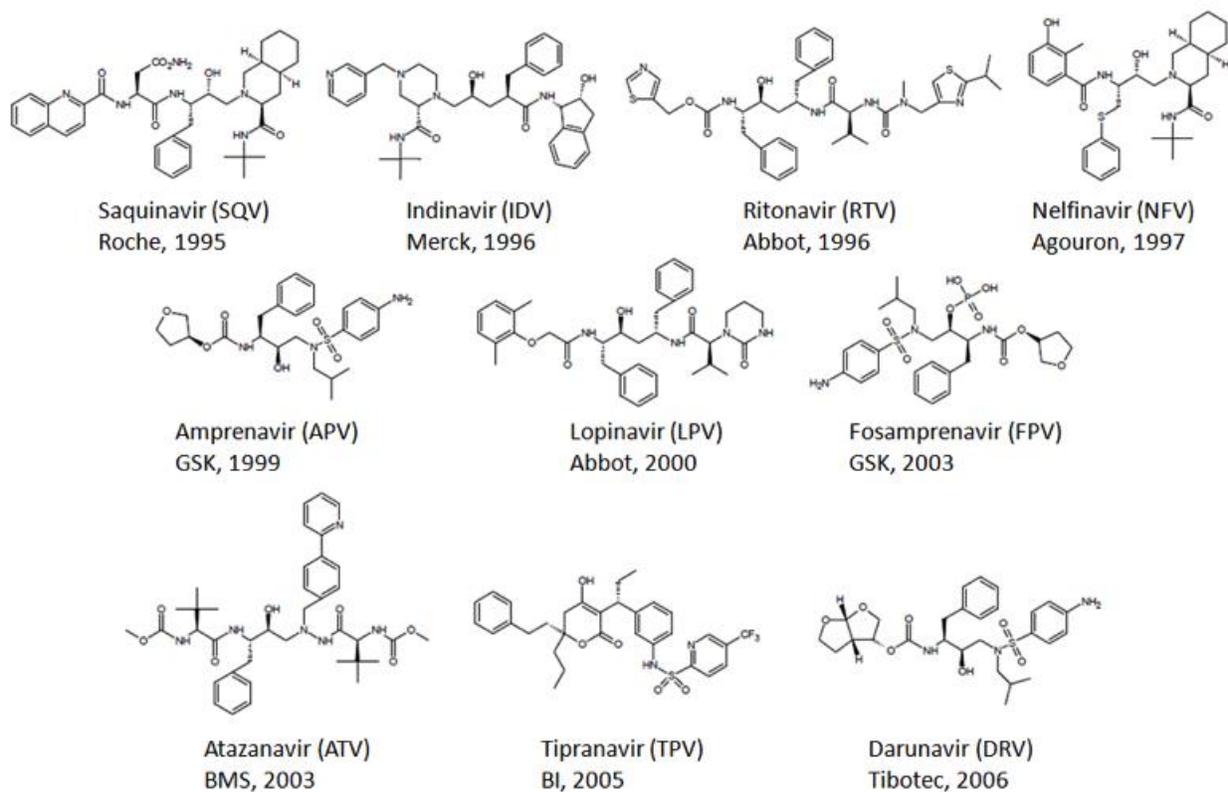


Figure 1-22. Structures of the FDA-approved protease inhibitors.

CHAPTER 2 INTRODUCTION TO BIOPHYSICAL METHODOLOGIES

Introduction

Several different methodologies were utilized in this study, including mass spectrometry, fluorescence spectroscopy, circular dichroism spectroscopy (CD), site-directed spin-labeling, continuous-wave (CW) and pulsed electron paramagnetic resonance spectroscopy (EPR). General overviews to these methods are given in the following sections. Additionally, each research chapter will have a detailed materials and methods section for all experiments contained within.

Mass Spectrometry

Mass spectrometry (MS) is a rich, analytical technique that can be used to determine the exact elemental make-up or chemical structure of a molecule or protein (Sparkman 2000; Hoffman and Stroobant 2007). The fundamental principle behind MS is ionization of the molecule followed by identification of the mass-to-charge ratios. The first application of MS to the study of peptides or proteins came in 1958 (Beckey 1969), but arguable the most significant MS achievement relative to the study of proteins came with the development of matrix-assisted laser desorption/ionization (MALDI) by Hillenkamp and Karas in 1985 and the work from several others that immediately followed (Dougherty 1981; Karas, Bachmann et al. 1985; Karas, Bachman et al. 1987; Tanaka, Waki et al. 1988). Two types of approaches are generally used when using MS to examine a peptide or protein. These include the “top-down” approach in which intact proteins are ionized and examined, and the “bottom-up” approach in which proteins are enzymatically digested into peptide fragments then analyzed (Hoffman and Stroobant 2007).

A schematic diagram of a simple mass spectrometer is shown in Figure 2-1. All MS instruments are composed of three basic parts, the first of which is the ion source (Hoffman and Stroobant 2007). The ion source functions to convert the analyte molecules into ions. Techniques for ionization are variable, and include electron (EI) and chemical ionization (CI) for gases and vapors, electrospray (ESI) and MALDI for liquid and solid samples, among many others. In CI, the sample is ionized via chemical reactions that occur during collisions between ions and molecules in the ion source. In EI, electrons interact with molecules to produce ions (Hoffman and Stroobant 2007).

After ionization, the ions are transported via electric or magnetic field to the analyzer, which is the second integral part of a typical MS instrument. The mass analyzer sorts ions by mass-to-charge ratio. Newton's second law of motion (Eq. 2-1) and Lorentz' force law (Eq. 2-2) describe the dynamics of charged particles in applied fields under vacuum. The Lorentz force is described as the force applied on a charge due to an applied field, where F is the force in newtons, E is the electric field in volts per meter, B is the magnetic field in tesla, q is the electric charge of the particle in coulombs, v is the instantaneous velocity in meters per second, and \times is the vector cross product. Newton's second law of motion describes the relationship between force (F), mass (m), and acceleration (a), and is only valid at ion velocity much lower than the speed of light. Using the expressions given in Equations 2-1 and 2-2, the differential equation shown in Equation 2-3 can be used to describe the motion for charged particles. If the particle's initial conditions are known, this classic equation completely describes the motion of the particle in space and time with respect to mass-to-charge (m/Q). Most commonly, this ratio is presented in terms of the dimensionless m/z ,

where z is the number of elementary charges on the ion (e), as described by Equation 2-4. There exist many common types of mass analyzers, including the sector field mass analyzer, the time-of-flight (TOF), the quadrupole, and the linear quadrupole ion trap. Each type of mass analyzer has unique pros and cons, and the choice of analyzer is highly sample dependent (Hoffman and Stroobant 2007).

The third and final necessary component of all MS instruments is the detector. The detector provides information regarding the mass-to-charge ratios and the quantity of each specific ion present (Sparkman 2000). Standard detectors function in one of two ways; either it records the charge induced or the current produced when the ion passes by or hits a surface (Hoffman and Stroobant 2007).

$$F = ma \quad (2-1)$$

$$F = Q(E + v \times B) \quad (2-2)$$

$$(m/Q)a = E + v \times B \quad (2-3)$$

$$z = Q/e \quad (2-4)$$

Fluorescence Spectroscopy

Fluorescence spectroscopy, also called fluorometry or spectrofluorometry is a form of spectroscopy that analyzes fluorescence from a sample. For a thorough review on fluorescence spectroscopy, the reader is referred to the Principles of Fluorescence Spectroscopy textbook, written by J. R. Lakowicz (Lakowicz 2006). This technique utilizes a beam of ultraviolet light to excite electrons via photon absorption within a given system, causing an emission of light. After a species is excited from its ground state to a higher energy state, the species then emits a photon as it returns to the ground electronic state (Lakowicz 2006). Analysis of the frequency and intensity of emission leads to information regarding the structure of the different vibrational levels. A typical

experiment results in an emission spectrum, where the various wavelengths of fluorescent light emitted by a sample are monitored while the excitation wavelength is kept constant (Lakowicz 2006).

A typical fluorescence instrument functions by passing an excitation light source through a monochromator and into the sample cell, thereby exciting molecules within the sample. Typical light sources include lamps, lasers, and photodiodes. The excited molecules fluoresce and that light passes through a second monochromator to the detector. The function of the monochromator is to transmit light at a specific wavelength and intensity. The detector is generally positioned 90° from the incident light to reduce effects from scattering and improve signal to noise (Lakowicz 2006).

Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy is an optical biophysical technique used to measure differences in absorption of left and right-handed circular polarized light by optically active materials (Atkins and Paula 2005). This technique allows for the determination of the average global secondary structure of a peptide or protein. When being used for this application, CD is performed in the far UV regions of 180 - 260 nm. A CD signal arises when a peptide bond is in a structured environment, and left and right-handed circularly polarized light is absorbed to different extents by chiral amino acids (Atkins and Paula 2005). The raw data is usually reported in ellipticity with units of millidegrees (θ), which are most often converted to mean residue ellipticity $[\theta]$ ($\text{deg cm}^{-2} \text{dmol}^{-1} \text{residue}^{-1}$), as given in Equation 2-5, where M_r is the protein molecular weight, c is the protein concentration (in mg/mL), l is the cuvette path length and NA is the number of amino acids in the protein (Atkins and Paula 2005).

Advantages of CD spectroscopy include that the technique requires very small amounts of protein, is non-destructive, is highly sensitive toward small changes in secondary structure, and require very little data processing. In addition, even small changes in overall secondary structure can be monitored very accurately. A glaring disadvantage, however, is that the technique is not site-specific. In other words, only the overall structure can be determined with no indication of what regions contain what elements of secondary structure. In biophysical studies of proteins, CD is often used as positive evidence of proper protein folding, but is not a definitive method. A schematic diagram of a typical circular dichroism instrument is shown in Figure 2-3 (Atkins and Paula 2005).

$$[\theta] = \frac{\theta \times 100 \times Mr}{C \times l \times NA} \quad (2-5)$$

Electron Paramagnetic Resonance Spectroscopy

Site-Directed Spin-Labeling

Site-directed spin-labeling (SDSL) is a technique utilized to incorporate a paramagnetic tag, or label, into a protein or other biological molecule at a specific site in order to facilitate electron paramagnetic resonance (EPR) studies (Griffith and McConnell 1965; Stone, Buckman et al. 1965). SDSL, first introduced in the mid-1960s by Griffith and McConnell, et al., is an excellent technique for studying the conformations and conformational changes of proteins. Most often, SDSL is carried out by engineering a cysteine residue into the protein of interest via manipulation of the DNA through a technique called site-directed mutagenesis. The engineered cysteine residue can then be site-specifically modified with a spin label.

There are many spin labels commonly used in EPR studies today, some of which include (1-Oxyl-2,2,5,5-Tetramethyl- Δ 3-Pyrroline-3-Methyl) Methane-thiosulfonate (MTSL), 4-Maleimido-TEMPO (MSL), 3-(2-Iodoacetamido)-PROXYL (IAP), and 4-(2-Iodoacetamido)-TEMPO (IASL). Each of these spin labels are nitroxide-based. Each one contains a free electron associated with either a five or six-membered ring with four methyl groups to help defer collision-induced reactions with the radical. Used in this study were three spin labels, MTSL, MSL, and the non-native amino acid p-acetyl-L-phenylalanine. The following sections provide details on each of those labels. Each of the spin labels contain a flexible linker of between four and six bonds that connect the nitroxide to the CYS residue, and a head group of varying bulkiness. Dr. Luis Galiano performed an experiment to examine the differences in the CW EPR line shapes for HIV-1PR labeled at the K55C position with each of four different spin labels. All spectra were collected at the same temperature with the same concentration of protein; thus the differences in the line shapes are due to differences in inherent spin label mobility. That data was reported in his 2008 dissertation and is shown in Figure 2-4 (Galiano 2008).

(1-Oxyl-2,2,5,5-Tetramethyl- Δ 3-Pyrroline-3-Methyl) Methane-thiosulfonate (MTSL)

MTSL is the most commonly used spin label (Hubbell, McHaourab et al. 1996; McHaourab 1996; Hubbell, Gross et al. 1998; McHaourab, Kalai et al. 1999; Hubbell, Cafiso et al. 2000; Hubbell 2000). MTSL is incorporated into a protein via thiol-based chemistry; a general reaction scheme is shown in Figure 2-5, where the modified cysteine residue is referred to as R1 (Berliner, Grunwald et al. 1982). The aqueous exposed spin-labeled side chain can be easily accommodated by most sites within proteins and has been shown to be no more perturbing than most other single amino acid substitutions (Hubbell, McHaourab et al. 1996).

Figure 2-5 also serves as a graphical representation of the χ^4/χ^5 model for the intrinsic mobility of the MTSL label, which was developed in the laboratory of Wayne Hubbell at UCLA and is based upon the spin label positioned in a solvent exposed site on an α -helix in T4 Lysozyme (Langen, Oh et al. 2000). MTSL has a 5-bond flexible linker, but the χ^4/χ^5 model predicts that the hydrogen atom on the α -carbon interacts with the δ -sulfur, thus restricting the motion of the first three bonds in the linker, limiting the rotatable degrees of freedom to the χ^4 and χ^5 bonds.

4-Maleimido-TEMPO (MSL)

In comparison to MTSL, MSL, whose structure is shown in Figure 2-6, has a much bulkier head group. This slows the overall intrinsic mobility of the label, and lends to the more restricted anisotropic EPR line shape, as shown in Figure 2-4. Another key difference between MTSL and MSL is the mechanism of attachment. Again, MTSL has a thiol reactive moiety that leads to an S-S bond between the nitroxide and the CYS residue. MSL, on the other hand, bonds to the CYS residue via a C-S attachment. This fact makes the MSL label more stable, with respect to permanence of the attachment. Under certain experimental conditions, this realization becomes very important.

Non-native amino acid p-acetyl-L-phenylalanine spin label

A non-native amino acid p-acetyl-L-phenylalanine and spin label, were generously received from Mark Fleissner and Wayne Hubbell at UCLA, who demonstrated its utility using T4 lysozyme (Fleissner, Brustad et al. 2009). This label differs from traditional nitroxide spin labels based upon its unique method of incorporation that does not rely upon the addition of a non-native cysteine residue. This is particular useful for proteins that have native cysteines that are either functionally active or are involved in structurally important disulfide bonds. Here, the genetically encoded unnatural amino

acid p-acetyl-L-phenylalanine (p-AcPhe) is reacted with a hydroxylamine-based reagent, resulting in a nitroxide side chain called K1.

The structure of K1 and a basic scheme of spin label attachment are shown in Figure 2-7. The K1 side chain has a higher degree of freedom than does the R1 side chain, and this fact is indicated by an EPR line shape indicative of a higher degree of mobility and a faster correlation time (Fleissner, Brustad et al. 2009). The unnatural amino acid, which can be site-specifically incorporated in using the nonsense amber codon into *Escherichia coli* (Wang, Zhang et al. 2003), *Saccharomyces cerevisiae* (Chin, Cropp et al. 2003), and mammalian cells (Liu, Brock et al. 2007). p-AcPhe contains a keto functional group, a group that is not present in any of the common amino acids that is reactive in aqueous solution with hydroxylamines (Brustad, Lemke et al. 2008).

Continuous-Wave Electron Paramagnetic Resonance Spectroscopy (CW EPR)

Introduction and history

Electron paramagnetic resonance (EPR) spectroscopy is the study of absorption of microwave radiation by a paramagnetic species in the presence of an external magnetic field. Today, EPR spectroscopy is used in a wide variety of applications for the detection of free radicals and paramagnetic species (Eaton, Eaton et al. 1998). In the study of protein systems, EPR is often used in conjunction with site-directed spin-labeling, which will be discussed in details in the upcoming sections.

EPR spectroscopy was discovered and first described in 1944 by Soviet physicist Yevgeny Konstantinovich Zavoisky of Kazan State University in Kazan, Russia, when EPR signals were detected in several salts, including hydrous copper chloride, copper sulfate, and manganese sulfate. He was also said to have observed nuclear magnetic

resonance (NMR) in 1941 but dismissed early results on the basis of lack of reproducibility (Eaton, Eaton et al. 1998). Bloch and Purcell first published on NMR in 1946 and subsequently shared the Nobel Prize in physics in 1952.

CW EPR sample requirements

For all work reported within, the CW EPR methodology was performed with loop-gap resonator at X-band frequencies with a center field of approximately 3480 Gauss. These experiments required nanomole quantities of spin-labeled protein. Typical sample sizes were 10 microliters of 50-250 μM protein.

CW EPR theory

The unpaired electron, found most often in a free radical or paramagnetic species, experiences a net dipole. This dipole is due to a magnetic moment that arises from spin angular momentum of quantum number $m = \pm \frac{1}{2}$. In the simplest case of the free electron in solution, the electron's magnetic moment is degenerate, meaning that the energy of the two spin states, $m_s = \pm \frac{1}{2}$, is equal. When a magnetic field is applied, the electron can align itself either parallel or antiparallel to the field. This phenomenon is referred to as the Zeeman Effect. The Zeeman Equation (Eq. 2-6) describes the difference between the energy levels E_α and E_β (ΔE), where g is the spectroscopic g -factor, β_e is the Bohr magneton, and B is the strength of the applied magnetic field (Weil, Bolton et al. 1972; Poole 1983).

The Bohr magneton ($9.274 \times 10^{-24} \text{ JT}^{-1}$), is a proportionality constant defined in Equation 2-7, where e is the electric charge, \hbar is the Planck's constant divided by 2π ($1.054 \times 10^{-34} \text{ J}\cdot\text{s}$), m_e is the mass of the electron ($9.109 \times 10^{-31} \text{ kg}$). Energy is absorbed, i.e., resonance occurs, when the applied energy is equal to the difference in

energy levels E_α and E_β ; this is achieved by maintaining a constant frequency and sweeping the magnetic field. The energy diagram describing this graphically is given as Figure 2-8, while Figure 2-13B shows a typical derivative absorption spectrum for a free electron in solution (Weil, Bolton et al. 1972; Poole 1983). When EPR is performed in conjunction with SDSL, the free electron ($m_s = \pm 1/2$) associated with the nitroxide interacts with the nuclear spin from nitrogen ($m_I = 1$) via hyperfine coupling. As such, both levels E_α and E_β split into three Hyperfine energy levels, affording the system three allowed energy transitions. The energy diagram for this type of system, along with the corresponding derivative absorption spectrum is given in Figure 2-9 (Weil, Bolton et al. 1972; Poole 1983).

$$\Delta E = h\nu = g\beta_e B \quad (2-6)$$

$$\beta_e = e\hbar / 2m_e \quad (2-7)$$

Nitroxide spectral line shapes

The EPR spectral line shape is sensitive to local secondary structural elements, local dynamics, and conformational changes, all of which have an impact on correlation time (Figure 2-10), and because of this, conformational changes can be easily monitored with SDSL EPR (Altenbach, Flitsch et al. 1989; Farahbakhsh, Altenbach et al. 1992; Altenbach, Yang et al. 1996; McHaourab, Lietzow et al. 1996; Hubbell, Gross et al. 1998; Altenbach, Cai et al. 1999; Altenbach, Klein-Seetharaman et al. 1999; Barnes, Liang et al. 1999; McHaourab, Kalai et al. 1999; Hubbell, Cafiso et al. 2000; Mollaaghababa, Steinhoff et al. 2000; Columbus 2001; Columbus, Kalai et al. 2001; Columbus, Kalai et al. 2001; Barnakov, Altenbach et al. 2002; Columbus and Hubbell 2002; Columbus and Hubbell 2002; Columbus and Hubbell 2004; Qin, Feigon et al. 2005; Fanucci and Cafiso 2006; Fanucci and Cafiso 2006; Guo, Cascio et al. 2008).

The correlation time of the paramagnet is described by three modes of motion, namely rotational correlation time (τ_R), internal correlation time (τ_i), and local dynamics and backbone fluctuations (τ_B). Rotational correlation time is the rate at which the protein is tumbling in solution; that rate is clearly affected by the size and tumbling volume of the paramagnetic macromolecule, as well as the viscosity, temperature and presence of solutes in solution. The internal correlation time (τ_i) describes the mobility of the spin label, and is affected by the size and structure of the nitroxide spin label, intrinsic spin label mobility and the torsional oscillations within the spin label. The local dynamics and backbone fluctuations (τ_B) are characterized by structure (secondary and tertiary), in addition to conformational changes involving the paramagnetic center.

The CW EPR spectral line shape is highly sensitive to changes in correlation time and anisotropy. For sites that undergo fast isotropic motion and a low rotational correlation time, the resultant spectra have sharp, narrow peaks (top). As motion becomes more restricted and rotational correlation time increases, resonance peaks decrease in sharpness and line broadening occurs. This phenomenon is shown in Figure 2-10, in the intermediate, slow (middle), and rigid spectra (bottom).

CW EPR line shapes are analyzed in order to extract parameters that describe the motion of the spin label, as described in the previous sections. Perhaps the simplest and most frequently used parameter is the ΔH_{pp} , or the peak-to-peak width of the central resonance line, most commonly reported in units of Gauss. As shown in Figure 2-10, the line shape resulting from fast, isotropic motion shows sharp, narrow resonance lines that broaden as motion becomes more restricted and anisotropic. The value of ΔH_{pp} , therefore, increases with anisotropy and motion restriction. Scaled

mobility is essentially a normalized ΔH_{pp} and is reported in order to compare ΔH_{pp} values from data collected on different instruments. Scaled mobility, Equation 2-4, is calculated by normalizing experimental values of ΔH_{pp} (δ_{exp}) using spectral line widths of the most mobile (δ_m) and immobile (δ_i) spectrum (Hubbell, Cafiso et al. 2000).

$$M_s = \frac{\delta_{exp}^{-1} - \delta_i^{-1}}{\delta_m^{-1} - \delta_i^{-1}} \quad (2-8)$$

Pulsed Electron Paramagnetic Resonance Spectroscopy

Introduction and history

The basic concepts of pulsed EPR are very similar to those of pulsed nuclear magnetic resonance (NMR), with only minor perturbations. Pulsed EPR was first described by W. B. Mims of Bell Laboratories in the 1960's, when a modulation of the echo amplitude in a Ce³⁺-doped CaWO₄ crystal was observed for the first time. In the same year, the Lockheed group in California observed echo modulations on a single crystal of Ce³⁺-doped lanthanum magnesium nitrate (Schweiger and Jeschke 2001). It was not until the 1980's that pulsed EPR started to undergo rapid development, and in 1982 that X-band pulsed EPR spectrometers became commercially available. Pulsed EPR has since gained popularity as a valuable tool in a wide variety of applications in the fields of chemistry, physics, and biology.

The specific pulsed technique utilized in this research is called pulsed electron double resonance (PELDOR), or double electron-electron resonance (DEER). DEER experiments use two different microwave frequencies to measure the strength of the coupling between two electron spins; in our case, to measure the distance between two nitroxide spin labels. The following sections will provide details regarding sample

requirements, theory of DEER, pulse sequences, data analysis and interpretation of results, and error analysis.

Sample requirements

As with CW EPR, the pulsed EPR methodology utilized here on the ELEXSYS E580 with MD4 or MD5 dielectric resonator requires nanomole quantities of spin-labeled protein. Typical sample sizes are a total of 100 μL , including 70 μL of 150 - 300 μM protein and 30 μL of d-8 glycerol. Certain samples also contain 3 μL protease inhibitor.

DEER theory

DEER experiments generally utilize two different microwave frequencies in order to extract information regarding the distance between two paramagnets by measuring the strength of the dipolar coupling. In all cases described in this work, that distance refers to the spacing between two nitroxide spin labels at specific sites within either HIV-1 protease or GM2 activator protein. In theory, the DEER technique is capable of extracting distances in the range of 1.5 – 8 nm, however it is seldom used to reliably extract distances greater than 5 nm. For the purposes of describing the theory behind these experiments, the spin label at position 1 will be referred to as spin subset A, and the spin label at position 2 will be referred to as spin subset B.

The DEER technique can be described using a simplified Hamiltonian (Equation 2-6). The simplified Hamiltonian is comprised of three sets of terms, including the Zeeman terms for the spin subsets A and B, $\Omega_A S_z^A$ and $\Omega_B S_z^B$, respectively, and the dipolar coupling term $\omega_{ee} S_z^A S_z^B$. The coupling term is proportional to the inverse cube ($1/r^3$) of the distance between the spins, where r is said distance.

DEER experiments can be carried out using several different pulse sequences. The simplest and first to be developed were the 2 + 1 and 3-pulse DEER, shown in Figures 2-11 and 2-12, respectively. The 2 + 1 sequence is applied at a single frequency. The 3-pulse DEER consists of a two-pulse echo-forming sequence ($\pi/2$ - π) separated by a constant delay τ at a frequency ω_1 , in addition to a pump pulse (π) at a frequency ω_2 with a variable time t with respect to the first observer pulse. Both the 2 + 1 and 3-pulse DEER, however, experience a dead time of approximately 64 ns, during which the system cannot record data. This dead time severely limits accurate experimental measurements (Pannier, Veit et al. 2000).

All pulsed EPR work reported in this dissertation made use of the common 4-pulse DEER sequence at X-band frequencies. The standard 4-pulse DEER sequence, shown in Figure 2-13 and developed by Pannier *et al.*, negates the dead time problem by introducing a refocusing 180° π pulse in the frequency $\omega_{(1)}$ (Pannier, Veit et al. 2000). The Hahn Echo sequence is used to produce an echo for the spins in resonance with the microwave frequency $\omega_{(1)}$, and consists of a $\pi/2$ pulse that flips the magnetization in the XY plane. After a time τ_1 , a π pulse is applied to invert the magnetization, and after a second time τ_1 an echo is produced along the Y axis. After a time τ_2 , the π pulse refocuses the A subset of spins. An additional π pulse at the microwave frequency ω_2 flips the spin subset B, which in turn affects the spin subset A. Because the subsets A and B are dipolar coupled, the resultant spin echo oscillates with the dipolar coupling frequency ω_{ee} (Pannier, Veit et al. 2000).

For all work described here, the distance of interest lies between two identical spin labels on a single protein macromolecule. Figure 2-14 shows the positioning of both the

pump pulse and observer sequence in the field-swept echo experiment. Because the paramagnets are identical spin labels, the g values are very similar and the spectral features of the labels overlap; because of this, selective excitation of a single label is impossible. Due to the larger number of excited spins at the central transition, the pump pulse ω_2 is positioned on the central resonance line, providing an optimal S/N ratio. The observer pulse is applied to the frequency corresponding to the low field resonance, which, for nitroxide spin labels, is approximately 26G (72 MHz) below the central resonance at X-band.

Given the concentration of protein used for DEER experiments (150 - 300 μ M), it is certain that spins from the subset "A" will come in inter-molecular contact with spins from the subset "B" from separate proteins, resulting in a random distribution of spin-spin interactions. The sum of these inter-molecular interactions gives rise to a background signal in the form of an exponential decay. Equation 2-10 describes the DEER signal, where $V(t)$ represents the raw experimental dipolar modulation, $B(t)$ represents the background contribution from inter-molecular spin-spin interactions, and $F(t)$ is a function describing the background subtracted dipolar modulated echo data. After background subtraction, the function $V(t)$ can be written as Equation 2-11, where $\kappa(t)$ is the kernel function given in Equation 2-12.

In order to ensure that the background subtraction is done to the most accurate extent, a form of self-consistent analysis was performed on all data presented in this dissertation. A program called DeerSim is used to generate a theoretical dipolar evolution that is completely free of any background contributions. This theoretical echo curve is overlain with the background subtracted echo curve. If the two curves overlay

exactly, then the background subtraction is accurate and the data is free of contributions from inter-protein spin-spin distances. Figure 2-15 shows typical dipolar modulated echo data from a system containing two separate nitroxide spin labels, including the raw oscillation, the linear function $B(t)$ (in red) which describes the contribution from background, and the background subtracted echo data with fit (grey).

All DEER data presented in this dissertation was analyzed using the DeerAnalysis software. The DeerAnalysis software uses both shell factorization and Tikhonov regularization in order to extract distance information from the dipolar modulated echo curves. For a comprehensive discussion on DEER data collection and analysis, the reader is referred to a number of sources (Jeschke 2002; Jeschke 2002; Jeschke 2004; Chiang 2005; Jeschke, Chechik et al. 2006; Jeschke and Polyhach 2007), including the DeerAnalysis2008 user's manual, which is available online at www.epr.ethz.ch.

Though several methods exist by which to analyze DEER data, including the direct Fourier transform and Monte Carlo analyses, this chapter will focus specifically on Tikhonov regularization (TKR) (Tikhonov 1943; Hansen 1998; Chiang, Borbat et al. 2005). TKR analysis is based upon function in Equation 2-13, which is used in order to find the best answer to an ill-posed problem by balancing the quality of the fit with the smoothness of the solution by varying the regularization parameter λ , where P is the probability of the spin-spin distance, K and L are operators, and S is the experimental data vector. Each individual λ value is given by Equation 2-14. $\eta(\lambda)$ and $\rho(\lambda)$ are given in Equations 2-15 and 2-16. Following TKR analysis at many values of λ , $\log \eta(\lambda)$ is plotted against $\log \rho(\lambda)$. The aforementioned plot is referred to as the L-curve. Good quality DEER data will result in an L-curve with a sharp point, often referred to an

“elbow” (like that of an “L”). The L-curve is subsequently used to determine the optimal regularization parameter by selecting the value of λ found at the elbow point within the “L.” Figure 2-16 shows a standard L-curve, with regions of the L-curve designated as under-smoothed, optimal, and over-smoothed. Again, selection of the optimal regularization parameter allows for extraction of the most accurate distance profile.

Once the correct regularization parameter has been chosen, the optimal distance profile can be selected. The distance profile can be examined for most probable distance, average distance, and the full-width and half-maximum intensity (FWHM). The most probable distance is the distance at which the intensity is the greatest point in the distance profile. The FWHM is the breadth, in Angstroms, at half of the maximum intensity and is indicative of flexibility within the system. Figure 2-17 shows two examples of dipolar modulated echo curves with the corresponding distance profiles. The echo curve in black translates to a distance profile with a greater FWHM, most probable distance, and average distance than does the echo curve in grey. This would tend to indicate that the distance between spin labels in the “grey” sample are, on average, closer together and held in a more rigid position than the spin labels in the “black” sample.

A more quantitatively descriptive method of analysis can be performed, and this process will be termed the Gaussian reconstruction process. In the Gaussian reconstruction process, individual sub-populations are identified and used to regenerate the distance distribution profiles in an effort to define the conformational ensemble. A conformational ensemble is the total of all the individual and distinct structural populations. To define a conformational ensemble, one must report the most probable

distance, the FWHM, and the relative percent of the total population for each individual Gaussian-shaped population. Once the distance profile has been adequately regenerated, a theoretical dipolar modulation is generated and overlain with the background-subtracted experimental dipolar modulation. If the two echo curves overlay exactly, the populations used to regenerate the distance profile are used to define the conformational ensemble.

One final and additional step is performed as a method of error analysis. Minor populations are suppressed one at a time, and in linear combinations. These suppressed ensembles are then used to generate a new theoretical echo to overlay with the background subtracted experimental echo. If the echo curves still overlay perfectly, the population is said to have been successfully suppressed, and that population is discounted from a contributing member of the conformational ensemble. If, on the other hand, the suppression of one or more populations changes the echo such that it no longer overlays with the background subtracted dipolar modulation, that population is said to contribute to the total conformational ensemble.

$$\hat{H} = \Omega_A \hat{S}_Z^A + \Omega_B \hat{S}_z^B + \omega_{AB} \hat{S}_z^A \hat{S}_z^B \quad (2-9)$$

$$V(t) = F(t) B(t) \quad (2-10)$$

$$V(t) = \{1 - [1 - \Delta D(t)]\} B(t) \quad (2-11)$$

$$K(t, r) = \int_0^1 \cos [(3x^2 - 1) \omega_{dd} t] dx \quad (2-12)$$

$$\Phi[P] = \|KP - S\|^2 + \lambda^2 \|LP\|^2 \quad (2-13)$$

$$G_{\lambda}(P) = \|S(t) - D(t)\|^2 + \lambda^2 \left\| \frac{\partial}{\partial r^2} P(r) \right\|^2 \quad (2-14)$$

$$\rho(\lambda) = \|S(t) - D(t)\|^2 \quad (2-15)$$

$$\eta(\lambda) = \left\| \frac{\partial}{\partial r^2} P(r) \right\|^2 \quad (2-16)$$

E580 spectrometer DEER setup

This section serves as a detailed operating manual and experimental setup protocol for DEER experiments run on the Bruker EleXsys E580 Spectrometer in the laboratory of Alex Angerhofer. For a more comprehensive manual, readers are referred to the ELDOR manual. General concepts discussed here include sample preparation, equipment and cryogenics setup, cavity tuning, preliminary experiments and DEER experimental setup, closing out an experiment, and shutting down the spectrometer and pumps.

Two days prior to start of DEER experiments, request the liquid Helium dewar from the department of physics cryogenic website (<http://www.phys.ufl.edu/~cryogenics/>). Follow the 'Order Helium' link to reach the online liquid helium request form, shown below in Figure 2-x. Fill in each box and press 'Send in Request.' The type of need is 'Cooldown.' Under special needs, select 'Non-magnetic.' In the comments section, you should mention how many days you will need to keep the dewar, and also clarify that we are room 303F. Check the level of nitrogen in the dewar in CLB 416. If necessary, order more nitrogen by calling (352) 372-8417 and requesting a low-pressure nitrogen dewar, using account number A2504. Be sure to specify if any empty dewars will be ready for pick-up.

On the day of DEER experiments, prepare purified protein sample for DEER run by buffer exchanging the sample with deuterated buffer using a 5 mL Desalting column. Deuterium (^2H) extends time of spin dephasing in DEER experiment. Wash the desalting column with nanopure water (nH_2O), 1M NaCl, nH_2O , 0.5 M NaOH, nH_2O , and NaOAc (pH 5) using 3 to 4 column volumes of each (15 to 20 mL). Equilibrate column with 10 mL deuterated NaOAc (pH 5). Inject 1 mL of sample (ideally OD280 = 2.5 to 2.7). Collect 1 mL of flow-through in a 1.5 mL tube. Inject 0.5 mL of deuterated NaOAc, collecting the last 0.5 mL of flow-through. Continue injecting 1.5 mL of deuterated NaOAc while collecting the buffer-exchanged sample in another tube. Inject 0.5 mL of nH_2O while collecting the last 0.5 mL of buffer-exchanged sample in a third tube. Concentrate the 2 mL sample to approximate OD reading of 3.0. Thoroughly wash and vacuum-dry the EPR tube to avoid cross-contamination of the sample. Mix 70 μL of sample with 30 μL of deuterated glycerol (glassing agent), and any inhibitors or substrate necessary for the experiment. Insert the mixed sample into a clean EPR tube using a skinny piece of tubing connected to a small syringe.

In order to set up the equipment on the day of the DEER experiment, check to be sure correct that the resonator and cryostat are correctly installed. See Alex Angerhofer and refer to ENDOR manual for assistance. Set up temperature reference by filling the reference dewar with liquid nitrogen and placing it in front of the cavity. Position the blue reference line into the dewar through the cork stopper. Turn on the water supply. Do so by opening four valves, two for water supply and two for water return. The valves are located overhead on opposite sides of the room, and are currently connected to green water hoses. Pump down the cavity. Close off the diffusion pump by closing both the

screw-valve and the lever-valve. Open the diffusion pump by-pass. Close off the cavity by closing the black valve on the front of the cryostat. Be sure the vacuum line is connected to the cavity. Turn on the rough pump (switch embedded in the black cord) and turn on the vacuum gauge (set to sensor 1). When the vacuum gauge reads 10^{-3} Torr, open the black valve in front of the cavity. Let the vacuum return to 10^{-3} Torr and let it sit for ~ 10 mins. Next open the valves to the diffusion pump, and let the vacuum return to 10^{-3} Torr and let it sit for ~ 10 mins. If diffusion pump is cool, turn it on by plugging it in and flipping the switch on the front of the pump. If diffusion pump is hot/warm, leave the pump off and make sure water supply is all the way open. NEVER TURN ON THE DIFFUSION PUMP IF IT IS HOT TO AVOID OVERHEATING. Switch vacuum gauge to sensor 2. Let vacuum sit at $10^{-6} - 10^{-9}$ Torr for ~20 mins. Set up cryostat cool down. Retrieve the assigned He dewar from the helium storage room on the third floor of Leigh Hall by disconnecting the computer and the recovery line before moving the dewar. Be sure to close the recovery valve while dewar is disconnected from the wall. Once the dewar is in CLB 303, connect its computer power cord and connect its He recovery line to the central He recovery (transparent He recovery line from cryostat is always connected to central He recovery). Make sure the recovery valve is open (must remain open throughout the entire DEER experiment). Set up the dewar computer for transfer. Press the far right button on the computer. Select transfer → user=303F → begin transfer. Insert transfer line into He dewar. Be sure the valve on the transfer line is open. Set the transfer line into position at the top of the dewar. Open both the screw-valve and the lever-valve and slowly insert the transfer line. It is important that this step is done slowly as the warm transfer line heats the helium as they

come into contact and the helium boils off. The faster the transfer line is inserted, the more helium you lose and the more pressure builds. Try to keep the pressure below 50 in of water. If pressure builds too fast, you can pull the transfer line back up a little and resume lowering the transfer line once the pressure has dropped. Keep lowering the transfer line until it is all the way down. Close the screw-valve tight around the transfer line. Connect the yellow hose to the top of transfer line (it should be already connected to the bottom valve of He flow network). Insert transfer line into the cryostat. Remove the plug from the cryostat, hold the transfer line so that it is straight and aligned with the probe, and slowly insert it into the probe. Once the transfer line is all the way in, tighten the nut halfway to leave the nut loose in order to purge the air out of the probe with He gas. Start cool-down. Make sure all the connections in the He system have been made. Open the valve on the board that connects the transfer line to the system (bottom valve). Turn on the He pump (small blue pump behind the magnet). Open the valve on the board that connects the pump to the board (third up from bottom). Open the black-handled valve on the wall that leads to the He recovery system. Let system pump for ~10 mins.

Tighten the nut to begin liquid He transfer (slightly wiggle the transfer line as you gently tighten the nut, making sure that it is all the way tightened). Turn on the temperature control unit. Set the parameters of the temperature control device to the following settings: prop = 21 (default is 20). Leave int and der at default values. Set the temperature to 65K. Press the auto button in the block for the heater. Let system equilibrate to reach desired temperature. This can take awhile, if you walk away, be sure to come back in time to slow the gas flow as you reach the set temperature.

Remember, reference thermocouple of the temperature control unit (blue cord) has to be in the dewar with liquid nitrogen to give you correct temperature reading. While cooling, open the needle valve all the way, in order to get a He flow of approximately 40 psi. Once cooled, start closing the needle valve on the transfer line or on the 1st bronze round handle in increasingly small increments to find the optimal (and minimal) He flow (usually ~6-10 psi) necessary to maintain 65K. Watch heater bars on the thermo regulator display (3-4 bars is ok, 5-6 bars means that He flow is too fast). If the gas flow is too high and the heater has to work too hard, it can and will burn out.

Turn on the magnet and console. Turn on the water cooler that chills the magnet. Turn on the magnet by hitting a green button, then hit the white reset button. Turn on the console by hitting the green button. Wait for the console to load (about a minute). This is signaled by a green READY button. Set up the computer. Log-on to the computer (see Angerhofer to have an account created for you, have a secure password ready). Open Konsole., and type 'Xepr' and hit enter. Go to the Acquisition menu and select 'connect to spectrometer'. Enter 'e580' and click OK. Insert sample into probe. Prepare sample tube and adapter. Take a small piece of teflon tape and curl it up into a very small sphere. Drop this into the sample tube. Gently and carefully flick the tube to get the Teflon to the bottom of the tube. The Teflon ball absorbs the pressure of solution expansion, preventing tube breakage. Find an adapter that fits the tube well. The tube should not fall out once you took it out of the magnet (must fit tightly). Insert the sample tube and adapter into the sample wand by screwing it. The bottom of the sample tube should be inserted into the threaded end of the adapter. Doing otherwise will damage the adapter. Adjust the sample tube so that the active part (middle) of the

sample is 39 mm below the bottom of the adapter if using the MD5 resonator, or 55 mm below the bottom of the adapter if using the MD4 resonator. For any other resonator, be sure to check the manual. Make sure the position is above the teflon ball. Tighten the threads to secure the tube.

Freeze sample, first wiping off the sides of the tube and wand with a Kimwipe to remove fingerprints and oil. Fill an open-top dewar with liquid N₂ from CLB 416 nitrogen tank. Dip the tube into the liquid nitrogen at the most acute angle possible. This helps prevent the sample tube from breaking as a result of water expansion due to freezing. Leave the tube in liquid N₂ until just before you are ready to insert it into the cavity. Insert sample into cavity. Turn off the valve that leads to the He pump (third up from the bottom). It is critical that the vacuum is not on when the cavity is open as you will suck moist air into a cold cavity where it will freeze and form ice crystals. It is also critical that you swap out the wands quickly to minimize the amount of air that the cavity is exposed to and care must obviously be taken to ensure that you do not break your sample tube by hitting on the edge of the nut as you are putting your sample into the cavity. Ice on sample tube will not affect the signal, but sample will be stuck inside. When the vacuum gauges on the board read zero, loosen the nut around the sample wand. Carefully remove the sample wand (be sure to pull the wand straight up as pulling at an angle will break the sample tube). Quickly but very carefully, insert the other sample wand with your sample tube in it. If you are not placing sample you can just put a wand inside to seal the cold cavity. Lower it gently to the very bottom and re-tighten the nut. Make sure this nut is all the way closed, a leak here will introduce moisture into the very cold cavity where it will freeze your sample into place (the only solution to this is to

completely warm up the cavity and disassemble it, dry it, and put it back together.

Lower the 'Q lever'.

Tune cavity, by opening the FT Bridge window by clicking on the button that looks like a switch. Under the Bridge Configure tab, select CW mode and stab on. Open the Tuning window and the FT Bridge. In the Tuning window, turn on the dual trace and the reference arm. Set to Tune, then set the Attenuation to 15dB. If the dip is not seen on the screen, adjust the Frequency slider bar until it is found. Slide the Bias slider bar all the way to the right. Adjust Signal Phase so that the dip has a flat baseline and is on a plateau. Slide the Bias slider bar all the way to the left. Turn the monitor around so that you can stand in front of the spectrometer and see the monitor at the same time. Slowly lift the Q lever to its highest position (use small increments) or until the dip is fully broadened. Adjust the Frequency (use the small black arrows) to center the dip (now broadened). Center the stabilizing frequency (External Stabilizer slider bar) by using the arrows in the Receiver Unit tab of the FT Bridge. Click Operate and close the bridge tuning windows. In the FT Bridge panel Bridge Configuration tab, switch to Pulse mode. Close window. In the lower right hand corner of the screen, "Operation" and "Leveled and Calibrated" should both be green.

Create a folder for a new experiment. My computer → My Home → xepFiles → Data → create folder with descriptive name. Check defense pulse. Create an experiment by opening Experiment window (icon that has EXP written in a comic strip balloon) → Pulse tab → select Advanced → Enter experiment name → hit Create. Make sure that the amplifier is off (near zero values on display mean that it is on standby) and that the max attenuation power is set at 60dB (in the FT Bridge). Click the

Clutch button (icon that resembles a pair of cogwheels), which will send parameters to the hardware. Set up pulse tables. Open the Parameter panel (button 23), then drag the window to the upper left corner of the screen. Open SpecJet window (will appear on the upper left corner), and the FT Bridge panel. In the FT Bridge window (Receiver Unit tab), set Video Gain = 66 dB, and Video Band = 25 dB. In SpecJet, set Averages to 256. In the Parameters window (Patterns tab) select '+X' channel, set the pulse 1 Position to 200ns and the Length to 32ns. Select Acquisition trigger, then from the Channel selection menu, set the Length to 2ns in pulse position 1. Click Start in the Parameter window, then click Run in the SpecJet window. If the Start button is "grayed out," it means you forgot to press the clutch button. Increase Video Gain if pulses are not visible. Make sure that the 2ns pulse is in the middle of defense pulse!!! Adjust the Phase in the FT bridge if defense pulse is not clearly showing in one of the lines. Click Stop in both the SpecJet and Parameters windows. If the amplifier is already ON or in Standby, switch to Operate and skip to "Check for cavity ringdown". If amplifier is off, make sure the Max Attenuation Power is set to 60 dB and turn on the TWT amplifier (white toggle). Let the amplifier sit for 2 mins to warm-up. Test defense pulse once again before switching to Operate (Click Start in Parameters window, and Run in SpecJet window). Click Stop in both the SpecJet and Parameters windows. Push the white 'Operate' button on the amplifier. It takes 30-40 sec for the amplifier to reach -11.8. Check for cavity ringdown. Click Start in the Parameter window and Run in SpecJet. Turn down the High Power Attenuation slowly (this increases power) to check for cavity ringdown (usually to about 6dB). If there is significant ringdown, stop immediately and return the attenuation to 60 dB. Check the Q lever position. If problem

persists, seek help from Dr. Angerhofer. Use Video Gain if needed to adjust the size of the signal in the window and adjust signal phase to see peaks better. Ideally, the intensity should only change within the bounds of the defense pulse. Outside of the defense should just be noise. If there is significant intensity outside the defense pulse, (1) you can damage the detector and (2) it will interfere with your signal. Set Attenuation back to its maximum value (60dB). (Shift+click =10, Ctrl+click =1) Click Stop in both the SpecJet and Parameters windows.

The next objectives are to determine the center field of the spectrum, to know the appropriate attenuation, to determine the values of T2, and roughly, T1. First, we collect a field swept echo spectrum (FS) to determine the pump and observe frequencies. Then, we measure T2 to determine the correct value for the D2 parameter. Collect a field-swept spectrum. Set Up Pulse Table. Open Parameters. Select "+X" channel. Set the Position in (1) to 200ns and (2) to 600ns (400 ns for protonated matrices) and the respective Lengths to 16 and 32 ns. Pulse 1 is a 90° pulse while pulse 2 is 180°. We keep the lengths and positions of both pulses constant because we want to maintain a particular bandwidth of excitation to excite 2 particular populations we need. Thus, to calibrate pulses, we need to adjust power instead of the length later. Select Acquisition trigger and set Length to 2ns. Click Start in the Parameters window and Run in SpecJet window. Optimize and phase echo. Lower Attenuation to 6 dB to see the echo. Use Video Gain to adjust the signal intensity (maximum gain ~66 gives the best S/N ratio but a lower value should be used if needed to prevent clipping of the echo). Adjust the Channel offsets (click on Settings in the SpecJet window) so that both the real and imaginary components are near the bottom

of the screen in the SpecJet window. To optimize echo intensity, change the Center Field to maximize the echo (Parameters window → Field tab → Center Field). Adjust the Phase to make the imaginary channel (bottom) dispersive and symmetric. Set the Center Field position to be the center field transition (usually ~3460 G). Set Sweep Width to 200 G. Decrease high power attenuation (Ctrl+click) until the real echo has the maximum intensity (~0.1 - 6 dB) and phase again. This is simultaneously increasing power of both pulses (1) and (2). The attenuation that generates the largest echo is the correct power to make the first pulse a 90-degree pulse, and the second, 180°. Adjust video gain if needed. Re-phase if necessary. Determine echo location and width. To acquire only the echo you will have to determine the start and end positions of the echo using Acquisition trigger. Adjust the Acquisition trigger Position (will be ~1470 ns) so that the echo starts at 0 ns in SpecJet window. Write this down as “Acquisition Trigger position”. Then adjust Acquisition trigger Position so that the other end of echo is at 0ns (try not to clip-off the echo tail). As much as possible, include more echo rather than less). Calculate the difference (usually 200-300 ns) and write this down as Pulse Length under Acquisition trigger. Return the trigger position to the beginning of echo and length of the pulse width as the length. Click stop in both windows. Set up the remaining parameters. Set the integrator to 4.0 ns. Set shots per point to 50. In the Acquisition tab, set the x-axis to magnetic field (x-axis size is 1024 points by default). Close all windows. Press play to collect the spectrum. Click on FS (if needed) to show full screen. Write down Center Field (G)/Frequency (GHz) during run. Save the file into folder you created before for this experiment (experiment initials are FS). Export as a .DAT file (File → Export → Change .DTA file extension to .DAT) . Here, we collect the

echo-detected field-swept absorbance spectrum of the nitroxide radical. From this graph of echo intensity as a function of field, we have to determine and write down the new center field where the echo intensity is maximal.

Measure the T_m of the sample, by opening the Parameters window. In the Acquisition tab, select Time as the x-axis. Under Center Field enter the new center field you determined from the previous experiment. Select Acquisition trigger in the drop down menu and set the Position displacement (pos disp) to 16 ns. In the Patterns tab, select "+x" channel and set the Position displacement (pos disp) in the second column to 8 ns. Close the window and press Play to collect the spectrum. You will get exponential decay of your echo intensity as function of time between two pulses. The waves in the beginning are ESEEM modulations of ^2H in solvent. Write down actual Frequency (GHz). Save file (experiment initials are T_m) and export. Fit the line to single exponential decay: Fitting → Exponential → Exponential decay → hit the Fit button several times until Tau stops changing. $\text{Tau} = T_2$; expected to be within the 2000-16000 ns range. It is hard to get a right fit so you can determine T_2 by eye - it will be a point on curve a little before the flattening of the curve. T_2 must to be greater than or equal to d_2 , which is always 3000 ns. Record the Tau value.

Determine the pump and observe frequencies. DEER Set Up: Create Experiment. Click the Experiment button. Select Pulse tab → DEER → Enter a name of experiment (with a D ending) → Click Create. Type experiment name. In FT Bridge window set the Attenuation to its maximum (60 dB; tantamount to lowest power) → Close. Click Clutch button to send parameters to hardware. Set Up Pulse Table. Open the Parameters window, SpecJet and FT Bridge. In the Set up tab: Set the shot repetition time (SRT,

SRT correlates with T1 and is twice as long. When SRT can be doubled and the echo does not change, the $\sim T1$ because the spins have enough time to relax.) to 4000. Set $\pi/2$ equal to 16ns. Set d0 to 0 ns (will be determined shortly). Set d1 to 400 ns. Set d2 to the determined T2 currently 3000 ns is d2 of choice).¹ This can usually be a little shorter than the actual T2. The trade off is between measuring longer distances (longer d2) and getting better signal to noise (shorter d2). Set π ELDOR to 32 ns. Set d3 to 100 ns. Drop the High Power Attenuation (in the FT Bridge) to the previously determined level for the best 90° pulse. Set the Video Gain to 66 dB. In the Field tab change Center field (CF) and make the Sweep Width equal to 160G. Click on Start and Run. Phase Echo. Adjust the channel offset (click on Settings button in Specjet) and use the *2 to see the imaginary component more clearly. In SpecJet window, phase echoes (via Signal Phase) until the imaginary signal (the bottom one) is dispersive. Click Stop in both SpecJet and Parameters windows. Collect the spectrum. Select 8-step phase cycling to get only the DEER echo. This type of phase cycling effectively eliminates the two stimulated echoes. Close all windows. Press Play. Save as DEER Setup experiment (experiment initials are DSU) and export to ASCII as .DTA file. Determine Echo position and location. Determine the width of the echo by pointing the cursor to the beginning and end of the echo. First, write down d0 and PG parameter. The latter is essentially the width of the echo.

Field-swept echo-detected spectrum. In the Field Sweep tab, set the pulse gate (or gate; PG) to the echo width. Set d0 to the value that you determined for the left edge of the echo. Set shots per point to 100 and Sweep Width to 160G. Select 2-step phase cycling. Leave other parameters at their default values. Close all windows.

Collect Spectrum. Press Play and write down the current Center Field/Frequency.

Save file with an FSD initial.

Next, calculate ν_{ELDOR} . The pump pulse frequency should be set to frequency that corresponds to the maximum of the center field transition. The observe frequency should be set to the maximum of the low field transition. Determine the distance between the two transitions (ΔB) by subtracting the center field from the low field. Convert this to MHz ($\Delta\nu$) multiplying the field difference (in Gauss) by 2.83. The ELDOR frequency (pump) is obtained by subtracting ($\Delta\nu$) from the frequency used to collect the field swept spectrum (because it corresponds to the center field transition).

Sample calculation:

Frequency = 9.72589GHz

Center Field = 3462G

Low field = 3435G

$\Delta B = CF - LF = 3262 \text{ G} - 3435 \text{ G} = 26 \text{ G}$ (this is the lowest possible value for ΔB)

$\Delta\nu = \Delta B * 2.83 \text{ MHz/G} = 26 \text{ G} * 2.83 \text{ MHz/G} = 72 \text{ MHz}$ ($\Delta\nu$ should be $>65 \text{ MHz}$)

$\nu_{\text{ELDOR}} = \nu_{\text{current}} - \Delta\nu = 9.72589 \text{ GHz} - .072 \text{ GHz} = 9.71869 \text{ GHz}$ (ELDOR freq.)

Next, set up DEER experiment. Set up pulse parameters. In the 4-P DEER tab, set field position to the low field value you have found. Set the ELDOR frequency to ν_{ELDOR} you just calculated. Set the Attenuation to 0 dB to get DEER signal. Set shots per point to 100. Set $dx = 12$ (or 16). dx is time increment for pi-pulse in ω_2 frequency channel (big dx means less steps), $dx = \text{Tau} (\sim 3000) / \# \text{ points} (255) = \sim 12$ (usually). Select 2-step phase cycling. Set number of scans to 500000 (this will automatically be changed to the maximum possible number of scans, which is 132000). Close all

windows. Collect the spectrum. Press Play. Write down experiment start time and date. Stay for a couple of scans to check for the DEER signal that looks like oscillations of baseline (baseline should not be flat!) If you see it, pour all the liquid N₂ left into reference thermocouple dewar (enough for 2-3 day experiment). Let your experiment run until you have good signal-to-noise. An average run takes 2-16 hours for a “good” sample and 2 days for a “bad” sample. If you don’t get a decent signal within 24 hours do not expect substantial signal improvement in the next 24 hours. Recall that S/N is proportional to the square root of the number of scans. Fill both the temperature reference dewar and diffusion pump cooling dewar approximately every 12 hours. This is imperative for the maintaining the temperature of the experiment, as well as the safety of the pump.

To close out an experiment, write down start and stop time of DEER experiment, date of experiment, and number of scans. Hit Stop experiment when enough scans have been collected. Wait for the program to end the phase cycling (try NOT to touch Run button because it will start to overwrite your experiment). File → Save (give a descriptive name like “FS” for field swept or “DEER” for DEER experiment) → Export to ASCII .DAT file. Click on Clutch to stop sending parameters to the spectrometer. Open Parameters window. In the 4-Pulse DEER tab set Attenuation to 30dB, then Close. FT Bridge → Receiver Unit → Set High Power Attenuation to 60dB → Close. Bridge Configuration → CW mode Open Tuning window → Set Attenuation to 60 dB → Switch to Tune then Stdby → Close all windows. Turn off TWT amplifier by pushing Standby (green button) (If you are done with all your DEER experiments, you can also turn off the amplifier with the white toggle switch (put in Standby before turning off)).

Close third valve up from the bottom in order to break connection with the Helium pump. Loosen nut and take out the sample carefully to avoid breaking the tube. Replace with a sample wand. Tighten nut. Open valve to He pump.

Warm up the system. If all DEER experiments are complete, it is time to warm up the system. There are two ways in which to do this, one that takes approximately 1.5 hours and another that takes approximately 12 – 15. If there is no time rush, the slow method is simpler. For slow method: Turn off blue helium pump and close needle valve on the top of the Helium dewar. Reset the temperature control box by turning it off and turning it back on. You have now shut off the helium and set up the system to warm up naturally. After approximately 12 hours, when the system has reached room temperature, it is safe to remove the transfer line from the cavity. If a quick take-down is required, a faster method is available which utilizes the system heater. **CRUCIAL STEP:** If warming up using heater, it is necessary to leave the Helium flowing at a very low rate. **DO NOT WARM UP WITH HEATER WITH NO HELIUM FLOW**, but rather just decrease flow rate. Now begin to raise the target temperature on the temperature control box. It is important that this be done in small steps (approximately 5 – 10K at a time), as not to blow the heating element or otherwise cause damage. Either way, once the system has reached room temperature, it is safe to proceed. The needle valve atop the helium dewar should be in the closed position and the helium pump should be turned off. Disconnect the cryogenics. The next step is to remove the transfer line from the room-temperature cavity. Close the black valve under where the transfer line meets the cavity, closing off the cavity from the vacuum pumps. Unscrew the transfer line and slowly and carefully remove by pulling straight out. Take caution that fragile tip of the

transfer line does not get bent when removing. Place the protective cover on the end of the transfer line and screw into place. Replace the black cap onto the end of the cavity. Disconnect the yellow hose from the top of the transfer line. Loosen nut on top of dewar. With gloves on, lift the transfer line out of the dewar, then immediately tighten the nut atop the dewar once the transfer line is clear. Place transfer line in storage place. Close release valve on dewar and wall and disconnect release line. End transfer on the dewar computer, then disconnect computer from dewar. Return dewar to storage room and reconnect to wall. Open release valve on dewar. Reconnect dewar computer to computer in dewar storage room. Shut down the spectrometer and pumps. Go to the Acquisition window. Disconnect from spectrometer. Exit program. Turn off devices in the following sequence: TWT amplifier → spectrometer → water circulator → water supply for water circulator. Turn off diffusion pump. Leave rough pump and water supply/return until diffusion is cool. Turn off rough pump. Turn off water supply/return.

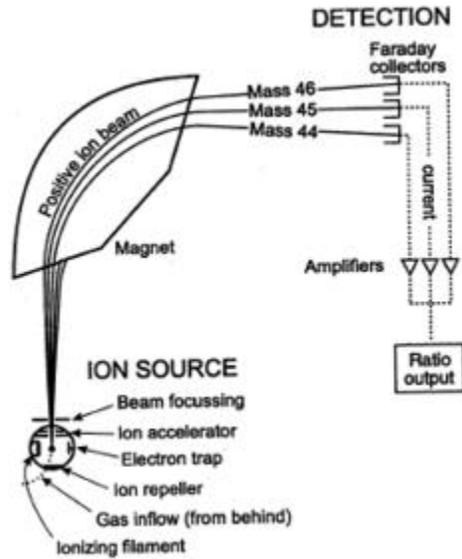


Figure 2-1. Schematic diagram of a simple mass spectrometer with a sector-type mass analyzer. Figure adapted from Wikipedia, and was free for distribution and reprint.

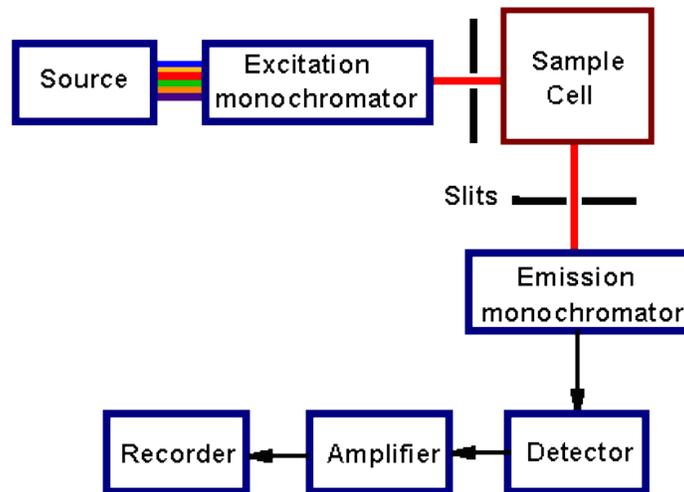


Figure 2-2. Schematic diagram of a standard fluorescence spectrometer.

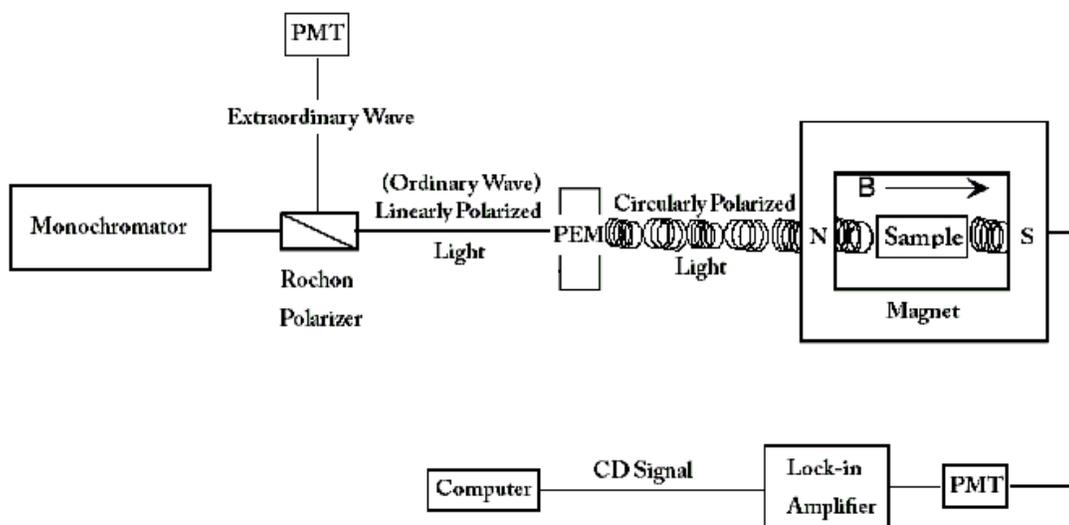


Figure 2-3. Schematic diagram of a standard circular dichroism instrumental set-up. Image adapted from Wikipedia and was free for distribution and reprint.

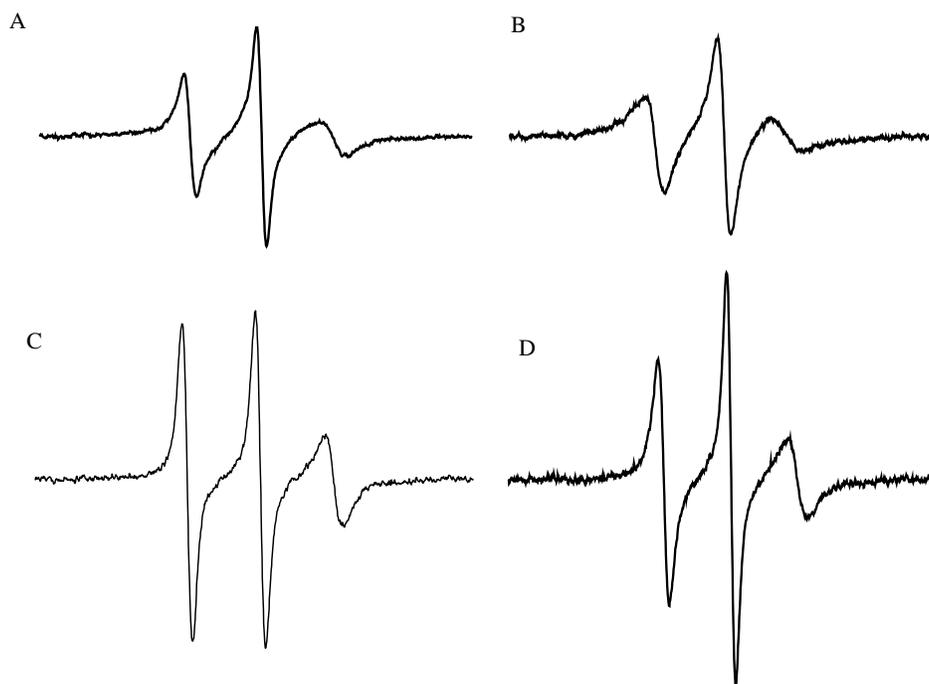


Figure 2-4. Effect of choice of spin label on the derivative EPR line shape of HIV-1PR Subtype B_{si} with A) MTSL, B) MSL, C) IASL, and D) IAP. All spectra were collected at X-band frequency with 100 Gauss sweep width.

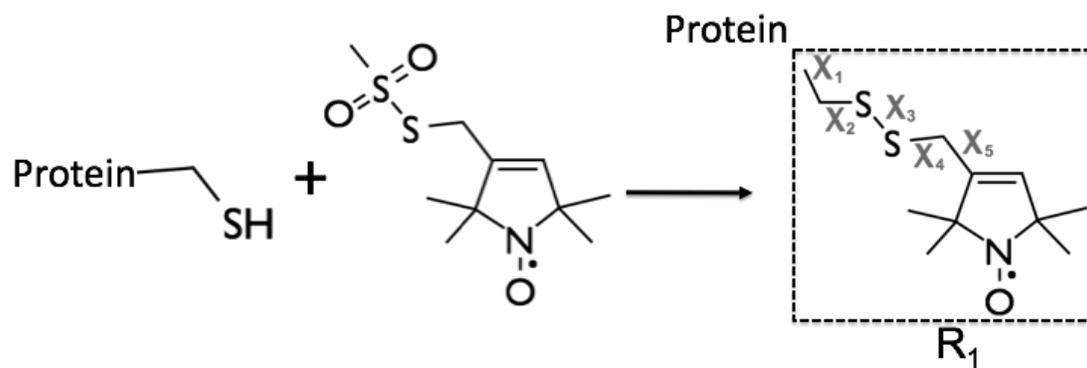


Figure 2-5. Site-directed spin-labeling scheme showing the reaction of MTSL with a free thiol group of an engineered CYS residue.

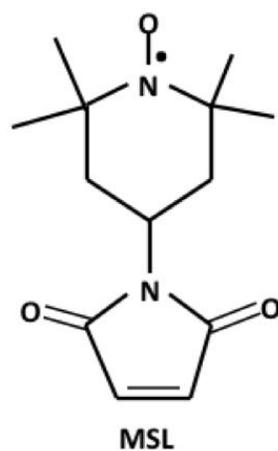


Figure 2-6. Structure of MSL spin label.

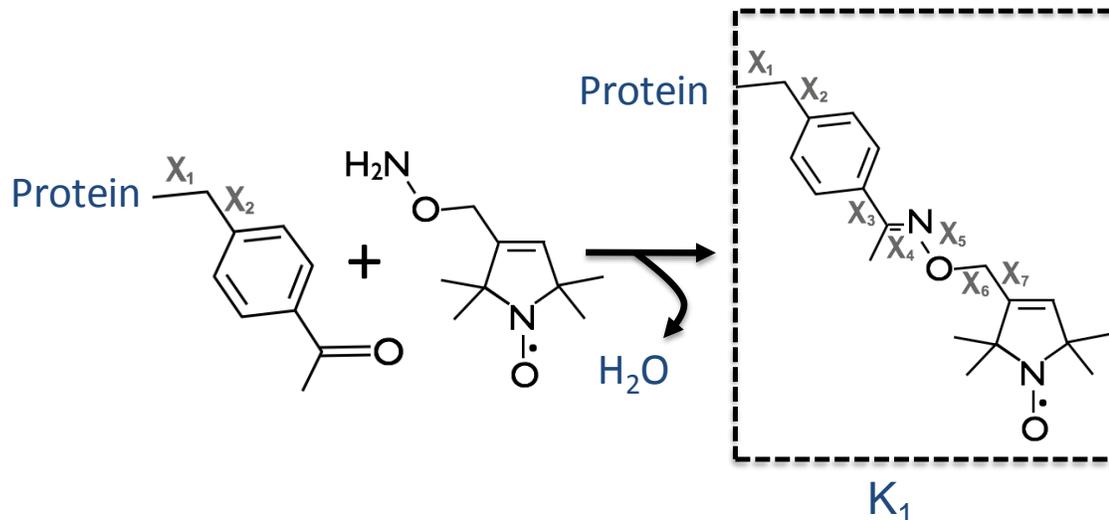


Figure 2-7. Spin-labeling scheme for the incorporation of the K1 spin label moiety.

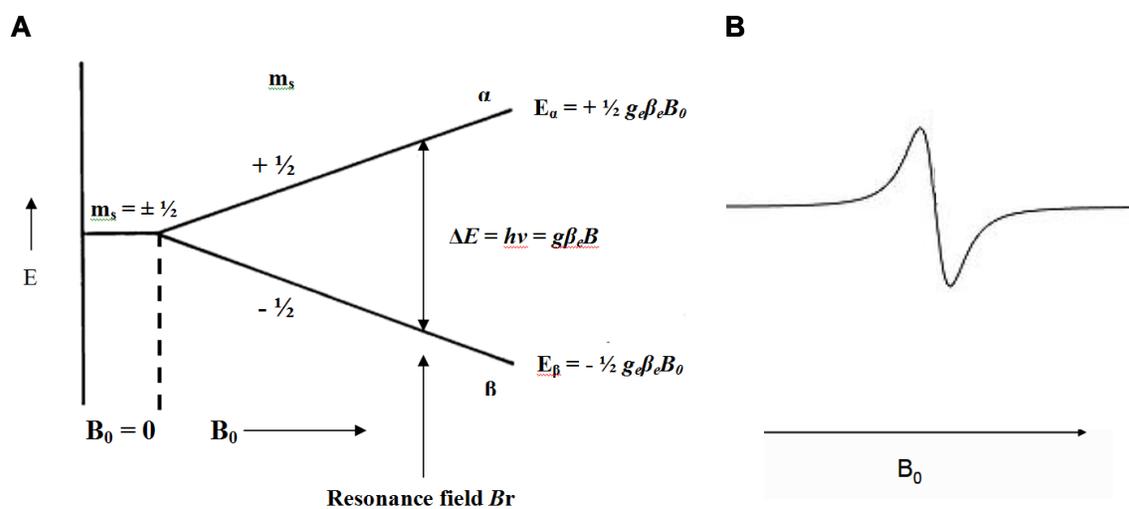
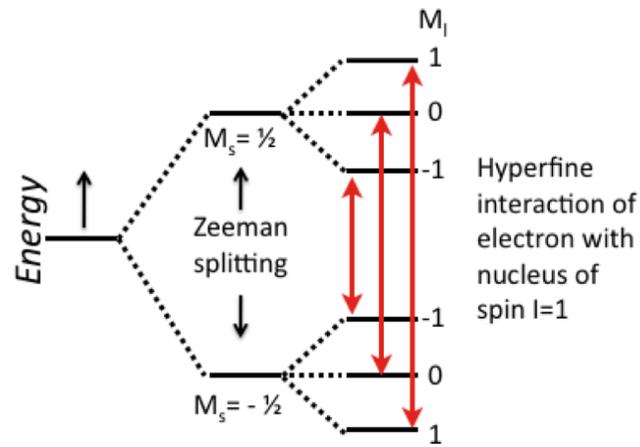


Figure 2-8. Scheme for electron transition in a magnetic field. A) Energy level diagram for a free electron in an applied magnetic field, and B) corresponding derivative of absorption spectrum.

A



B

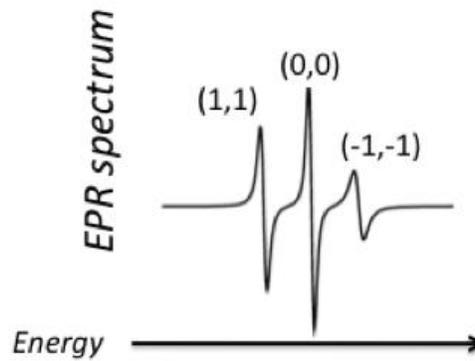


Figure 2-9. Cartoon representation of EPR. A) Energy diagram for a system with a free electron ($m_s = \frac{1}{2}$) undergoing hyperfine interaction with the nucleus of nitrogen ($m_I = 1$), and B) representative derivative EPR spectrum for a nitroxide spin label.

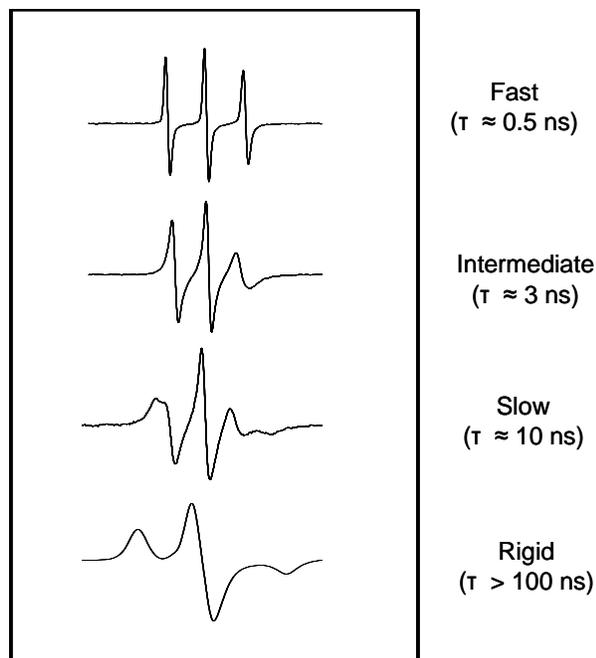


Figure 2-10. Dependence of EPR spectral line shape on motion, correlation time, and degree of anisotropy.

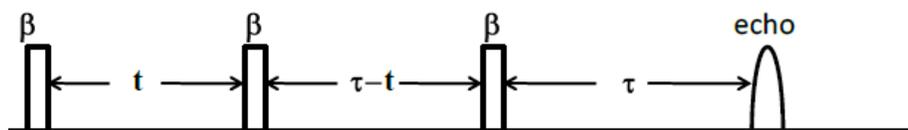


Figure 2-11. The 2 + 1 pulse DEER sequence.

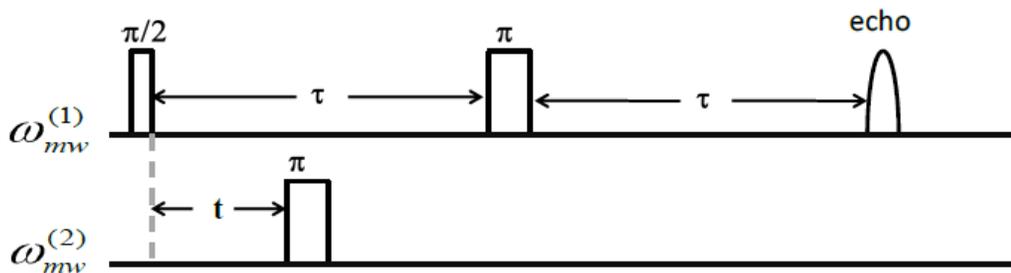


Figure 2-12. The 3-pulse DEER sequence, where time t is varied and τ is fixed.

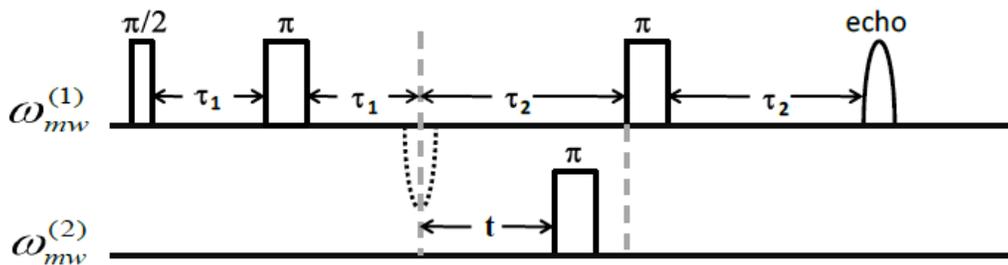


Figure 2-13. The 4-pulse DEER sequence, consisting of the observer sequence $\omega^{(1)}$ (top) with a Hahn echo sequence ($\pi/2$ - τ - π) followed by a refocusing π pulse, and a secondary pump pulse applied at frequency $\omega^{(2)}$ (bottom).

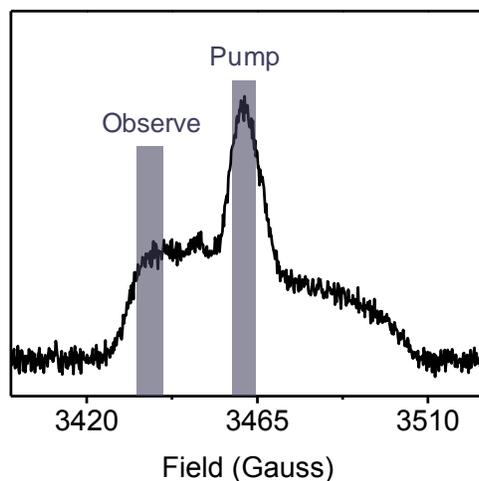


Figure 2-14. Absorption spectra for a nitroxide spin label with positions of the low field transition, approximately 26 Gauss (72 MHz) below the central resonance line, marked as the “observe” frequency, and the center field transition marked as the “pump” frequency.

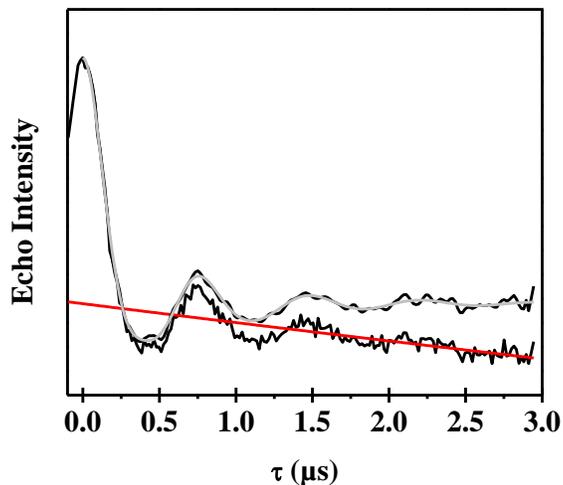


Figure 2-15. Sample dipolar evolution curve showing locations of raw dipolar modulation, background subtraction, and background-subtracted echo curve with fit.

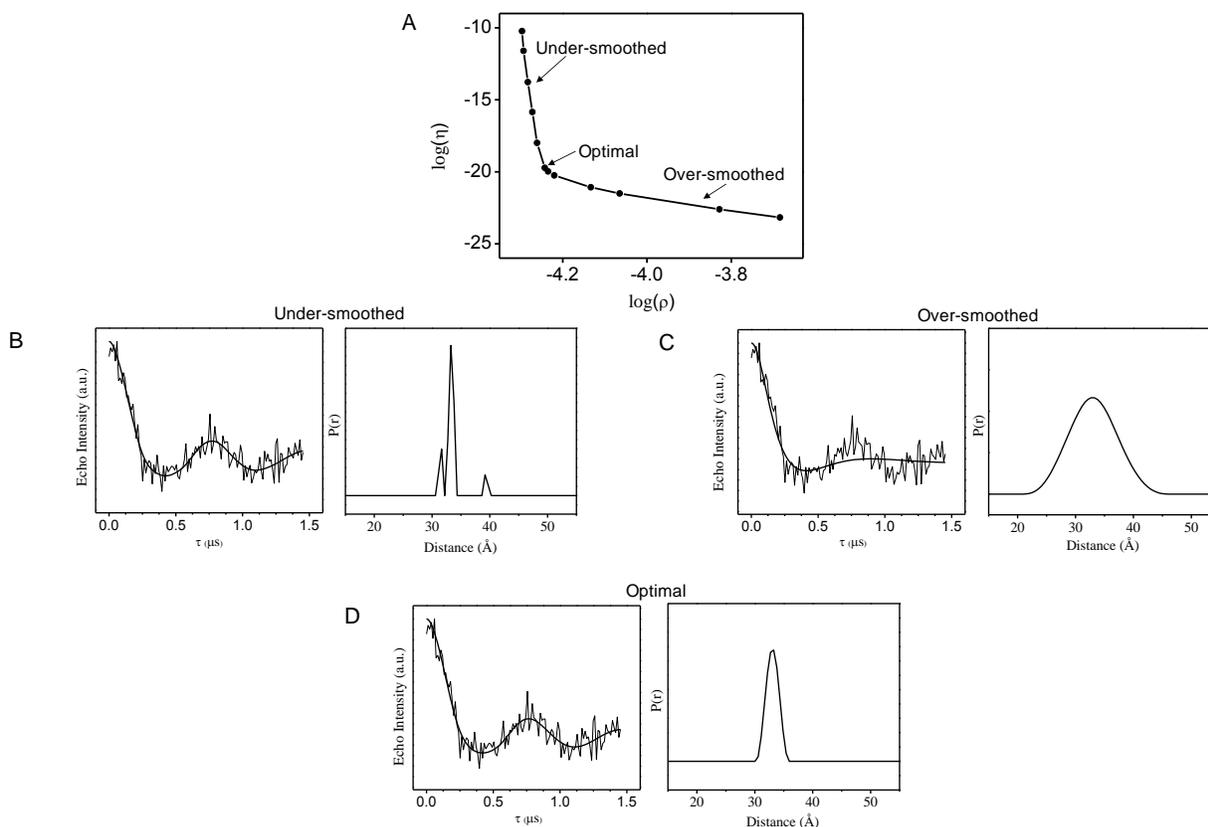


Figure 2-16. Selection of regularization parameter in TKR DEER analysis method. A) Typical L-curve showing regions of undersmoothed λ , optimal λ , and oversmoothed λ , with distance profiles and echo fits for B) undersmoothed, C) over-smoothed, and D) optimal regularization parameters.

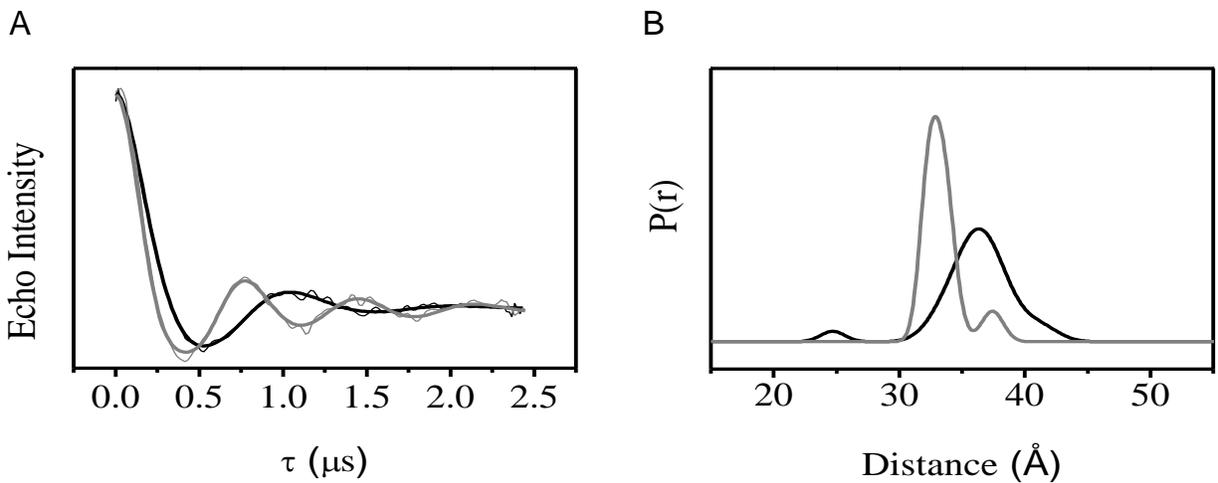


Figure 2-17. Examples of processed DEER data. Examples of A) dipolar modulated echo curves with B) corresponding distance profiles.

CHAPTER 3
SURFACE LOOP CONFORMATIONAL FLEXIBILITY OF THE GM2 ACTIVATOR
PROTEIN IS NOT ALTERED BY LIGAND BINDING BUT IS NECESSARY FOR LIPID
LIGAND EXTRACTION FROM BILAYER SURFACES

Introduction

Glycosphingolipid catabolism occurs in lysosomal compartments within the cell (Sandhoff, Kolter, 1995). GSLs are endocytosed, trafficked and sorted through early and late endosomal compartments on the way to the lysosome, where specific enzymes sequentially cleave sugar groups, eventually producing ceramide, which is finally deacylated to sphingosine (Griffiths, Hoflack et al, 1996). More than ten different enzymes and five accessory proteins are involved in this important process. One of those accessory proteins, the GM2AP, is essential for stimulating the catabolism of neuronal gangliosides by extracting ganglioside GM2 from intralysosomal vesicular membranes (Sandhoff, Kolter, 1996). As described in detail in Chapter 1, GM2 activator protein (GM2AP) is a general lipid transfer protein that functions in the lysosomal compartment of cells. GM2AP extracts ganglioside GM2 in the late endosome and presents it for hydrolysis by β -hexosaminidase A to convert ganglioside GM2 to GM3. Activator proteins often function by binding to glycosphingolipids and forming aqueous-soluble complexes thereby providing an aqueous-soluble enzyme access to a lipid generally confined to a hydrophobic environment. Mutations in GM2AP that inactivate the protein lead to the lipid storage disease known as Tay-Sachs (Sandhoff, Kolter, 1995; Gravel, Clarke, 1995). The importance of GM2AP in regulating lipid catabolism has been widely published but little is known about the mechanism that allows GM2AP to function. The focus of my research is on elucidating

key physical and structural characteristics that allow GM2AP to bind and extract its lipid substrate.

In addition to stimulating GM2 degradation, GM2AP has been proposed to act as a general lipid transporter because *in vitro*, GM2AP has been shown to bind and extract GM2 from micelles, transfer glycolipids, phospholipids and fluorescently labeled lipids (Conzelman, Burg et al., 1982; Mahuran, 1998; Ran, Fanucci, 2008), bind and inhibit PAF (Rigat, Wang et al., 1997; Rigat, Yeger et al., 2009), and present CD1 molecules to antigens (Kolter, Sanhoff, 2005; Major, Joyce et al., 2006; Zhou, Cantu et al., 2004). In order to carry out its function as a lipid transporter, GM2AP must partition with the lipid bilayer surface. A model of membrane partitioning was recently proposed based on results of sedimentation assays which showed that only a 15% fraction of GM2AP molecules remains on the bilayer surface, and therefore, implying that GM2AP is undergoing exchange with the bilayer surface (Ran, Fanucci, 2009). From fitting the results of sedimentation assays, we concluded that GM2AP establishes exchange equilibria of a minimum of four species: GM2AP in solution, GM2AP on the bilayer surface, GM2AP-lipid complex on the bilayer surface and GM2AP-lipid complex in solution. GM2AP-membrane interactions have also been studied by others using Langmuir monolayers and these data suggest that GM2AP is surface associated without deep penetration into the bilayer (Giehl, Leem et al., 1999). Recently, we also showed with site-directed spin-labeling (SDSL) and power saturation electron paramagnetic resonance (EPR) that the orientation of GM2AP on zwitterionic bilayer surfaces is such that the mobile external loops that line the opening to the lipid binding pocket are positioned at the bilayer surface (Mathias, Ran et al., 2009).

GM2AP is thought to sample three distinct conformations, although the functional importance of these conformers has not yet been defined (Wright, Li et al., 2000). Various X-ray structures with different lipid substrates reveal multiple conformations of a mobile loop region that is observed in either a “closed” or “open” conformation. These “open” and “closed” conformations suggest a functional conformation change that may occur upon lipid ligand or vesicle binding. To further probe the “binding” conformational changes of the loop region, site-directed spin labeling electron paramagnetic resonance (EPR) spectroscopy was applied to GM2AP (Lundblad, 2004). In proteins, site-directed spin labeling typically proceeds by generating site-specific cysteine amino acid substitutions for modification with paramagnetic nitroxide probes, e.g., 4-maleimide-TEMPO (Figure 3-1). Reduced cysteine residues act as nucleophiles capable of reacting with maleimide or thiosulfonate groups to form thioethers or disulfide bonds, respectively (Tipton, Carter, 2009). Cysteine residues involved in disulfide bonds do not readily react with those functional groups at pH 6.8-7.4 (Lundblad, 2004). To ensure that purified protein contained the proper connectivity of the 8 native cysteine residues and that only the introduced reporter cysteine (L126C) was labeled, a mass spectrometry based assay was developed. Because of the analytical challenge associated with verifying multiple disulfide bonds, a multitier assay was needed the full details of this experiment will be discussed in chapter 4 (Tipton, Carter, 2009). The first aim of my project is to assess the functional importance of the “mobile loop” regions seen in the crystal structure. To verify the importance of the “mobile loops” a series of cysteine (CYS) double mutants have been made where one residue in each mobile loop is substituted to a CYS. The two mutant CYS moieties

were chemically cross-linked and novel fluorescent based assays were used to monitor vesicle sedimentation, ligand extraction, and extraction kinetics of cross-linked GM2AP constructs. The results further hypothesize that restricting loop mobility slows the rate of substrate extraction.

Materials and Methods

Materials

The chemicals, reagents, and supplies were obtained from Fisher Scientific (Pittsburg, Pennsylvania) and used as received, with a few noted exceptions. Methanol and water were purchased from J.T. Baker (Philipsburg, NJ) at HPLC grade. Trypsin, endoproteinase GluC, endoproteinase Asp N, formic acid, and ammonium bicarbonate were purchased from Sigma-Aldrich (St. Louis, MO). 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) in chloroform, bis(monooleoylglycero)phosphate (BMPdi18:1 or DOBMP) in chloroform was purchased from Avanti Polar Lipids (Alabaster, AL) and used without further purification. N-(5-dimethylaminonaphthalene-1-sulfonyl)-1,2-dihexadecanoyl-sn-glycero-3-phosphoethanolamine, triethylammonium salt (DDHPE) was purchased from Molecular Probes (Eugene, OR) in the form of powder. Methanethiosulfonate spin label (MTSL) was purchased from Toronto Research Chemicals, Inc. 1,8-bis-maleimidotriethyleneglycol (BM(PEG)₂), 1,11-bis-maleimidotriethyleneglycol (BM(PEG)₃), and dithio-bismaleimidoethane (DTME) were purchased from Thermo Scientific Rockford IL.

Methods

GM2AP is sub-cloned into pET-16b vector, which is used for expression of histidine-tagged proteins. GM2AP is expressed in inclusion bodies and must be refolded following total denaturation with urea. The following sections describe, in

detail, the methods associated with the expression, purification, and spin-labeling of the various GM2AP constructs.

Expression of GM2AP constructs

GM2AP constructs are expressed in BL21 (DE3) *E. coli* as His-tagged proteins. Transformed cells are grown in 1 L sterile LB media to an $OD_{600} = 0.8-1.1$. The cells are then induced by IPTG for 8 hours at 20 °C by adding 500 μ L of 0.8 M IPTG. Following induction, the cells are pelleted using a Sorvall RC6 centrifuge at 5000 rpm for 20 minutes.

GM2AP purification protocol

The BL21 (DE3) *E. coli* pellets were resuspended in lysis buffer (0.1 M sodium chloride, 0.1 M Sodium phosphate, 1% tween). The cells were then lysed by two rounds of sonication and French press followed by centrifugation in a Eppendorf 5810 R centrifuge for 20 minutes at 10500 rpm. The isolated inclusion bodies were then resuspended in 5 M urea, homogenized, and sonicated once more. After centrifuging again in the Eppendorf 5810 R centrifuge for 20 minutes at 10500 rpm, the supernatant was loaded onto a nickel-chelating column equilibrated with 5 M urea at pH 8.0. A guanidine hydrochloride buffer (6 M guanidine hydrochloride, 0.1 M sodium phosphate, 10 mM Tris, and 5% glycerol) was used at pH 8.0, 7.0, and 4.5 to wash and elute the protein off the nickel-chelating column.

The eluted protein solution was slowly transferred into refolding buffer (50 mM tris, 10% glycerol, and 0.05% tween20) for rapid 20-fold dilution using a peristaltic pump. The refolded GM2AP mutant was then dialyzed to a pH of 7.0 and the mutant cysteine was labeled by either maleimide-biotin or maleimide-4-tempo.

A Q-column equilibrated with 25 mM tris, 2.5% glycerol, and 0.05% tween at a pH of 8.0 was used concentrate to the protein. The GM2AP mutant was then eluted from the Q-column by the addition of 150 mM sodium chloride.

The concentrated GM2AP eluted from the Q-column was loaded onto a S-200 Sephacryl column equilibrated with 25 mM ammonium bicarbonate to remove any tween20 and isolate the misfolded protein from correctly folded protein.

Spin-labeling

The substituted cysteine residue was modified with 4-maleimide-TEMPO (4MT) overnight in the dark (4 °C) in 25 mM Tris, 2.5% glycerol, and 0.05% Tween 20 buffer (pH 7.0). The spin-labeled L126C-4MT GM2AP construct was then dialyzed against 25 mM Tris, 2.5% glycerol, and 0.05% Tween 20 buffer (pH 8.0) and concentrated to 200 μ L. Concentrated spin-labeled protein was loaded onto a size-exclusion S-200 Sephacryl column (Amersham Biosciences, Piscataway, NJ) and eluted with 25 mM ammonium bicarbonate (pH 8.0) to remove the salts and detergents.

Site directed spin-labeling is used to determine if there is a conformational change in the “mobile loops” of GM2AP upon ligand binding. When spin-labeling was needed, the purified and refolded protein sample was buffer exchanged into 10 mM Tris HCl, pH 6.9 (further details of spin-labeling discussed in next section) after the final purification step. Approximately 1 mg of spin label (IASL, IAP, MTSL or MSL) was dissolved in 100 μ L ethanol, and added to approximately 40 mL of HIV-1PR in 10 mM Tris HCl, pH 6.9. The spin-labeling reaction was carried out in the dark (via wrapping reaction tube in aluminum foil) at room temperature for approximately 4-6 hours followed by 6-8 hours at 4 °C for inactive (D25N) constructs, and at 4 °C for

approximately 8-12 hours for active (D25) constructs. At this time, the sample solution was centrifuged at 12000 rpm for 20 minutes at 4 °C to remove solid impurities and aggregated proteins. The sample was then buffer exchanged into 2 mM NaOAc, pH 5, and concentrated to $OD_{280}=1.25$.

Crosslinking

The substituted double cysteine residues (I66C/L126C and I72C/S130C) were modified one of the following cross-linking reagents 1,8-bis-maleimidotriethyleneglycol (BM(PEG)₂), 1,11-bis-maleimidotriethyleneglycol (BM(PEG)₃), and dithio-bismaleimidoethane (DTME). Reactions were allowed to run overnight (4 °C) in 25 mM Tris, 2.5% glycerol, and 0.05% Tween 20 buffer (pH 7.0). Concentrated spin-labeled protein was loaded onto a size-exclusion S-200 Sephacryl column (Amherham Biosciences, Piscataway, NJ) and eluted with 25 mM Tris and 2.5% glycerol (pH 8.0) to remove any aggregated protein.

Preparation of POPC:Dansyl-DHPE liposomes

As previously reported by Ran, Fanucci, 2008, POPC in chloroform was used without further purification. POPC:dansyl-DHPE (4:1 in molar ratio) vesicles were prepared by mixing the desired amounts of both lipids (in chloroform), drying to a film by a stream of nitrogen and vacuum desiccation for 6-12 hours. POPC:R18 (19:1 in molar ratio) vesicles were prepared by mixing the desired amounts of POPC (in chloroform). For kinetic measurements, the dried lipid films were hydrated for one hour in an appropriate volume of 50 mM sodium acetate buffer, pH 4.8. Large unilamellar vesicles (LUVs) of above lipid samples were prepared by extrusion through 100 nm polycarbonate filters with 55 passes using a hand-held miniextruder. Final phospholipid

concentrations were determined on the basis of total phosphate determination by Malachite Green Phosphate Assay Kit (BioAssay Systems).

Results and Discussion

The models of GM2AP (PDB ID 1G13) from X-ray crystallography reveals that the volume of the hydrophobic cavity is roughly six times larger than the volume of a ceramide moiety (500 \AA^3). In addition, based on X-ray structure analysis of five different crystal forms, large differences in the diameter and area of the opening to the lipid binding cavity were detected and attributed to flexibility of loop regions that decorate the rim of the cavity (Wright, Li et al., 2000; Wright, Mi et al., 2004; Wright, Zhao et al., 2003; Wright, Mi et al., 2005) These regions (highlighted in Figure 3-2) are of particular interest for this study as their mobility may be related to function. The loops are located on opposing sides of the entrance to a prominent hydrophobic cleft whose width varies substantially among the different crystal structures. On one side of the cleft an “apolar loop” spanning residues V54 to W63, protrudes into solution and has a relatively stable conformation in all crystal structures examined due to its involvement in crystal lattice contacts (Wright, Zhao et al., 2003). On the opposing side, a reverse turn (S130-T133) containing W131 exhibits two alternative conformations. In structures where the cleft is wide open and bound to lipid (PDBID 1PUB), this loop is flipped out with W131 exposed to solvent (Wright, Zhao et al., 2003). The second conformation of this loop is seen in structures where the cleft is closed, and the loop is tucked in towards the interior of the cleft, burying W131 and allowing van der Waals’ contacts to exist between residues on either side of the cleft (L128, P129, L132 and I72, I66) and PDB ID 1G13 (Wright, Li et al., 2000; Wright, Mi et al., 2005). Given that this mobile loop is preceded by an extended disordered chain

segment (V122-P129), mobility of this entire region (V122 – T133) can modulate the width of this cleft, and therefore the size of the circumference of the cavity (Figure 3-2), which allows for lipid ligand to enter into the binding pocket.

The question arises as to whether the features of the protein that are seen in the crystal structure are in conformational exchange in solution and how these various conformations are related to function, as protein mobility and dynamics are often related to function (Smock, Gierasch, 2009; Clore, 2008). For the case of GM2AP, the multiple conformations in the crystal structure, which in effect modulate the size of the entrance to the cavity, may suggest that a conformational change is necessary for extracting lipid ligand from vesicles surfaces, or that the conformations will be modulated in the holo (lipid bound) or apo protein. Another possible role for the conformational flexibility is in conformational entropy to allow GM2AP to partition with the vesicle surface or to interact with hydrolases, such as HexA, in GM2 hydrolysis. To investigate the local structure and mobility of the apolar and mobile loops of GM2AP in solution, SDSL EPR was utilized, which is a powerful spectroscopic tool used to study conformational changes in proteins as well as to characterize local backbone motion (Fanucci, Cadieux et al., 2002; Fanucci, Cafiso, 2006; Hubbel, Cafiso et al., 2000; McHaourad, Lietzow et al., 1996; Columbus, Hubbell, 2004; Columbus, Kalai et al., 2001).

Thus, ten single CYS variants of GM2AP were previously generated and labeled with MTSL at the sites shown in Figure 3-3 . Six of the chosen spin-labeled sites are located in the membrane binding loops (A60R1, I66R1, I72R1, L126R1, S130R1, and N136R1) and were chosen to investigate the conformational flexibility using EPR

spectroscopy. The remaining four CYS variants of GM2AP (V54R1, L87R1, T90R1, and S115R1) were generated to probe sites thought to be either structured (α -helices and β -sheets) or loop regions of the protein. This conformational exchange process is slow on the EPR time scale such that both spectral components are resolved and the enthalpy and entropy required for the conformational change were found to be 65 kJ/mol and 215 J/mol K, respectively. Figure 3-4 shows the nitroxide EPR line shapes recorded at ambient room temperature for GM2AP in solution under basic pH and acidic pH in the presence of GM2 micelles. The EPR line shapes at each site are consistent with those expected based upon the local structure and dynamics reflected in the x-ray structures of GM2AP. This conclusion is drawn upon literature reports showing that the R1 line shape correlates with protein structural components and B-factors (Xu, Kim et al., 2008). Of the six sites in the flexible loops and disordered strand regions, A60R1, L126R1, and N136R1 have line shapes consistent with fairly mobile and solvent accessible sites on proteins (McHaourab, Lietzow et al., 1996). For example, the spectra from site L126R1 are narrow and intense, reflecting a higher degree of motional averaging. This site is located in a flexible strand with very high crystallographic B-factors. Sites I66R1, I72R1 and S130R1 have broadened EPR spectra with structure seen in the high field resonances, which indicate that these spin-labeled sites reside in more structured regions of GM2AP. In numerous X-ray structures of GM2AP, the side chains of I66 and I72 can be seen to point in towards the hydrophobic cavity, so it is likely that at these sites, the spin label motion is restricted by neighboring amino acid side chains as well as the limited space of the cleft leading to the lipid binding pocket. Further, the line shape from site S130R1 is

broadened and more reflective of a spin label attached to an alpha helical region of a protein, and some of the crystal structures of GM2AP place this residue in a helical structure (Wright, Li et al., 2000; Wright, Zhao et al., 2003).

Previous EPR spectra obtained with acidic conditions in the presence of GM2 micelles are very similar to those obtained under basic pH conditions, suggesting that little to no conformational change has occurred in the presence of ligand. In efforts to further characterize possible conformational changes of the loops upon GM2 ligand binding, two different double CYS constructs were prepared, I66R1/L126R1 and I72R1/S130R1. These sites allow for distance measurements across hydrophobic cleft between the two loops to be made. Low temperature EPR spectra were collected and analyzed for distances across the binding cleft. As can be seen in the spectra in Figure 4A, very little to no differences were detected in the spectra upon addition of ligand. The average distances obtained from analysis of the spectral line shapes using the empirical d_1/d_0 parameter (Gross, Columbus et al., 1999) are given in Table 3-1, and values are all between 18-22 Å. A slight increase of 2 Å in the average distance was seen for I66R1/L126R1 upon addition of GM2 micelles. Because the average distances are near 20 Å, which is the upper limit of distances readily detected by this method (Altenbach, Oh et al., 2001; Rabenstein, Shin, 1995), pulsed double electron-electron resonance (DEER) spectroscopy was also performed for these sites. The background subtracted DEER echo curves are shown in Figure 3-5. Unfortunately, the average distance of 20 Å is also near the lower distance cut off for accurate analysis of DEER distance distribution profiles. Nevertheless, the shapes of the echo curves are consistent with an average distance of 20-23 Å (Data in Table 3-1), and show that no

major conformational changes of these loops occur upon binding GM2 ligand. The lack of distinct oscillations in the DEER echo curve again indicates a high degree of conformational flexibility, which is discussed in more detail below.

Because the EPR line shapes and distance data reveal little to no change upon addition of GM2, two different experimental observations to ensure GM2 was bound were undertaken. In the first of these, samples where GM2 micelles were added, protein:GM2 complexes were purified by size chromatography and the presence of GM2 in the protein fraction was detected via a resorcinol assay as described previously (Ran, Fanucci, 2008; Ran, Fanucci, 2009). Second, for both A60R1 and L126R1, differential scanning calorimetry showed an increase in T_m of 0.9 °C of the thermotropic unfolding for the halo protein compared to the unliganded protein; thus, further confirming that GM2 was indeed bound to the protein. It is noteworthy that the absolute values of T_m differed for the two constructs (Data in Table 3-2), giving an indication of the relative stability of the modified proteins compared to WT (Figure 3-10).

In addition to the above-mentioned six sites in GM2AP, four additional sites were also previously chosen to investigate conformational heterogeneity. The EPR line shapes of all ten R1 labeled GM2AP constructs in basic solution with the overlaid simulations are shown in Figure 3-11. The EPR spectra were simulated using the MOMD model of Freed and colleagues (Budil, Lee et al., 1996). Using this method, two spectral components are required to adequately regenerate the experimental EPR line shapes for the sites in the flexible loops (A60R1, I66R1, I72R1, L126R1, S130R1, and N136R1). To simplify the fitting, we assumed that one component was mobile with the

value of the order parameter $C_{20} = 0$, as was done previously by others (Columbus, Hubbell, 2004). As can be seen in Figure 3-11, the two components have dramatically different line shapes, with one having narrow, isotropic like features, corresponding to a highly mobile site and the other a more broadened line shape corresponding to a site with more restricted motion. We assign the two components to a mobile and immobile spin-label conformation, consistent with the alternative conformations seen for these flexible regions in the GM2AP X-ray structures (Wright, Li et al., 2000; Wright, Mi et al., 2004; Wight, Zhao et al., 2005; Wright, Zhao et al., 2003). The relative percentages of each spectral component are also given in Figure 3-11.

To further test the hypothesis that the multiple spectral components reflect protein flexibility, four additional cysteine variants were generated within regions of the protein, which based upon the protein fold and X-ray B factors, would be predicted to have nitroxide line shapes that could be simulated with single component fits. Two of these sites, V54 and S115, are located on the back side of the cup structure in β -strands. In addition, L87 is a solvent accessible site on the short α -helix, and T90 is located in an unstructured loop. The EPR spectra obtained for these four spin-labeled sites are shown in Figure 3-11. In each case, these spectra have line shapes that are consistent with their location in structured or unstructured regions in GM2AP, and each could be adequately simulated using a single component fit and the overlay of the experimental and simulated spectra are shown in the top panel in Figure 3-11. Sample homogeneity was assessed by high resolution gel filtration and HPLC, which showed single peaks upon elution, verifying that the two components in the EPR data arise from conformational flexibility of a single R1 label and not from mislabeling or misfolding of

the disulfide bonds Figure 3-12. To support the claim that the two components in the spectral fits arise from conformational flexibility of the loops and not mislabeling or misfolding of the disulfide bonds GM2AP CYS variants were screened by mass spectrometry methods previously shown (Tipton, Carter et al., 2009). Briefly, the mass spectrometry results verified both the correct location of the spin-label and that samples were singly labeled. The isotopic patterns of the two distinct protein forms can be resolved by FTICR MS and thus, can be matched to the calculated isotopic distribution obtained from the molecular formula for the labeled and unlabeled forms of GM2AP. The full experiment for the L126R1 construct has been recently reported and will be discussed in detail in chapter 4 (Tipton, Carter et al., 2009).

To further investigate the origin of the multiple component spectra for sites located in the flexible loop regions of GM2AP, data were collected as a function of temperature over the range of 10 °C to 30 °C in 5° C increments. A similar approach has been utilized to characterize a conformational change in the transmembrane sequence of the transducer domain of *N. pharaonis* halobacterial transducer of rhodopsins II (NpHtrII) in lipid membranes (Bordignon, Klare et al., 2005). For GM2AP, L126R1 in the disordered strand and A60R1 and I66R1 adjacent to the apolar loop were selected for the variable temperature EPR experiments. If these regions of the protein are in conformational exchange, the equilibrium populations of the two components should change in a predictable manner with temperature. The EPR line shapes obtained at each temperature were then simulated for I66R1 and L126R1 using the MOMD model to extract fractions of populations of the two components. The high degree of free spin in the spectra of the A60R1 construct prevented accurate modeling

of these spectra. Representative EPR spectra and MOMD simulations of line shapes for L126R1 and I66R1 at select temperatures are shown in Figure 3-13. Changes in the EPR line shapes were seen for L126R1 with increasing temperature; however, no significant changes in the A60R1 or I66R1 line shapes were observed. The percentages of the mobile and immobile components were obtained from the line shape fittings for L126R1 and I66R1 and are plotted as a function of temperature (Figure 3-13). The plot of component percentages for I66R1 did not show a temperature dependent change, which is consistent with the EPR spectra that do not change significantly with temperature. However, it is noted that the MS results indicate under-labeling of site I66C, which suggest that spin-label incorporation at this site might interfere with WT flexibility. Keeping this fact in mind, the results indicate that site I66, at the top of the apolar loop is not undergoing conformational exchange and the two components in the spectral fit may represent alternative conformations of the spin label. This interpretation of the data is also consistent with the crystal structure, which shows little positional change of the backbone of this segment in the three monomers A, B and C (1G13) (Figure 3-2). Therefore, the origin of the two component spectra at this site may arise from different rotameric states of the spin label.

For L126R1, changes in the EPR line shapes as well as in the percentages of mobile and immobile components were observed as the temperature was varied. At low temperatures, the percent mobile and immobile components were about 10% and 90%, respectively. With increasing temperature, the percentage of mobile component increased and the percentage of the immobile component decreased until almost equal populations of the two components were seen at 25° C. The natural logarithm of the

ratio of the fractions of the two components exhibited a linear dependence with inverse temperature (Figure 3-14). Here, f_1 refers to the fraction of the mobile component and f_2 refers to the fraction of the immobile component. The data shown in Figure 3-14 were the result of two independent experiments performed on two separate samples prepared from different protein expression, refolding, spin-labeling and purification preparations. The simulations were also performed independently by different researchers. Results for the relative percentages of each component are strikingly similar. Both data sets and simulations were plotted together and the slopes of the lines were reproducible within error. The solid line shown is a linear fit of all the data taken together with values of -7.8 ± 0.8 for the slope and 25.8 ± 2.8 for the intercept, from which, the enthalpy and entropy of the conformational change were calculated to be 65 kJ/mol and 215 J/mol K, respectively. From the values of f_1 and f_2 determined from the spectral fits, the value of the equilibrium constant, K , at 298.15 K was 0.7, indicating that the activation energy for this conformational change is on the order of a few kT and that both conformations are easily accessible at physiological temperatures.

L126R1 is located in the conformationally disordered strand of GM2AP, and its position is poorly defined in all 3 monomers of the X-ray structure (1G13). In monomer A, the conformation of the loop is such that the side chain of L126 appears to point towards the protein interior while it extends out into solution in monomer C. The two spectral components determined from the MOMD simulation are consistent with these different environments of the spin label at this site. The immobile component may arise from a fraction of protein conformers having L126R1 pointing in towards the lipid

binding cavity, where the motion of the spin label can be restricted by neighboring amino acid side chains as well as by the limited space in the cleft of the cavity entrance. The mobile spectral component is consistent with the strand conformation seen in monomer C, where the spin label would extend out to solution and have a higher degree of rotational freedom.

In addition to binding its specific ligand, GM2AP has been shown *in vitro* to bind to other fatty acids (Wright, Li et al., 2005). Taking advantage of this result, a dansyl fluorescence based assay for monitoring the kinetics of lipid extraction was developed and previously reported by Ran, Fanucci 2008. The assay uses the fluorescent properties of phosphatidylethanolamine with a dansyl-labeled headgroup (dansyl-DHPE) in which, the wavelength of maximum emission and quantum yield is sensitive to the polarity of the local environment. Upon sequestration of the dansyl head group into the hydrophobic cavity of GM2AP there is an observable blue shift to 518 nm from 484 nm, corresponding to the dansyl residing in large unilamellar vesicles of POPC:dansyl-DHPE(4:1) to being bound in the hydrophobic pocket of GM2AP. A representation of this assay can be viewed in Figure 3-15. To further investigate whether loop plasticity is of functional importance the apolar and disordered loops from previously generated double CYS constructs I66C/L126C and I72C/S130C were physically tethered using three different cross-linking reagents: 1,8-bis-maleimidotriethyleneglycol (BM(PEG)₂), 1,11-bis-maleimidotriethyleneglycol (BM(PEG)₃), and dithio-bismaleimidoethane (DTME). The results (shown in Figure 3-16) dansyl-DHPE extraction assays show that merely mutating the native residues to cysteine slow the rate of ligand extraction by GM2AP. The addition the tethering

reagent further diminishes the rate of ligand extraction leading to the hypothesis the conformational plasticity is an important component in the ligand extraction mechanism of GM2AP.

To confirm that the physical tethering of loops also perturbed the extraction of the preferential biological ligand, a GM2 extraction assay was developed. The assay was used to monitor the efficiency of GM2AP at extracting GM2 from unilamellar vesicles containing PC/GM2/cholesterol/BMP (50:10:20:20 molar ratio) (Assay schematic shown in Figure 3-17)(Ran, Fanucci, 2009). We define extraction efficiency in this assay as the number of moles of GM2 divide by the number of moles of GM2AP that elute in tandem fractions from an S200 size column. The results again indicated that merely mutating the native residues in the apolar and disordered loops had an effect on the mechanism of extraction. While the physical tethering of the loops further reduced the number molecules of GM2 that GM2AP was able to extract (Data in Table 3-3).

Site-directed spin-labeling was utilized to investigate the flexible loop regions of GM2AP, an accessory protein with eight native cysteines that form four disulfide bonds essential for the stability of exposed loop regions. We were able to engineer a ninth and tenth cysteine in the protein as well as express and purify, then isolate homogeneous samples of spin labeled GM2AP. Analysis of the EPR spectral line shapes and MOMD simulations for spin labels in the disordered chain segment and both loops are consistent with the various conformations seen in X-ray structures. Multiple conformations of the mobile loop *in solution* were detected by EPR, indicating that the conformations seen in the crystal structure are present in solution and may

play a functional role in ligand binding or interactions with lipid vesicles. Although, distance measurements from EPR were not able to detect local conformational changes in the loop regions as a result of ligand binding. The holo state of R1 mutants was confirmed by differential scanning calorimetry, which observed a 0.9 °C increase in the thermotropic unfolding of GM2AP constructs with GM2 bound. Furthermore, functional assays showed that immobilization of the loops inhibited the ability of GM2AP to bind its ligand, indicating that although GM2AP may not undergo a global conformational change as a result of ligand binding, loop plasticity is essential to its role as a general lipid transfer protein.

Table 3-1. Summary of Double CYS R1 GM2AP Construct Distance Experiments

Construct	d1/d0	CW Distance	DEER Distance	Distance FWHM	Crystal Structure	Theoretical
I66R1/L126R1	0.39	18±2Å	18±2Å	12Å	18Å	18-30Å
I66R1/L126R1 +GM2	0.37	20±2Å	18±2Å	10Å	18Å	18Å-30
I72R1/S130R1	0.40	18±2Å	18±2Å	15Å	19Å	18-30Å
I72R1/S130R1 +GM2	0.38	18±2Å	18±2Å	12Å	19Å	18Å-30

Table 3-2. Differential scanning calorimetry values for GM2AP constructs

Construct	T _m
A60R1	75.0 °C
A60R1+GM2	75.9 °C
L126R1	72.7 °C
L126R1+GM2	73.8 °C

Table 3-3. GM2AP cross-linking ligand extraction data

Construct	Extraction Efficiency
WT	0.69
I66CL126C	0.40
I66CL126C BM(PEG) ₃	0.13
I72CS130C	0.51
I72CS130C BM(PEG) ₃	0.25

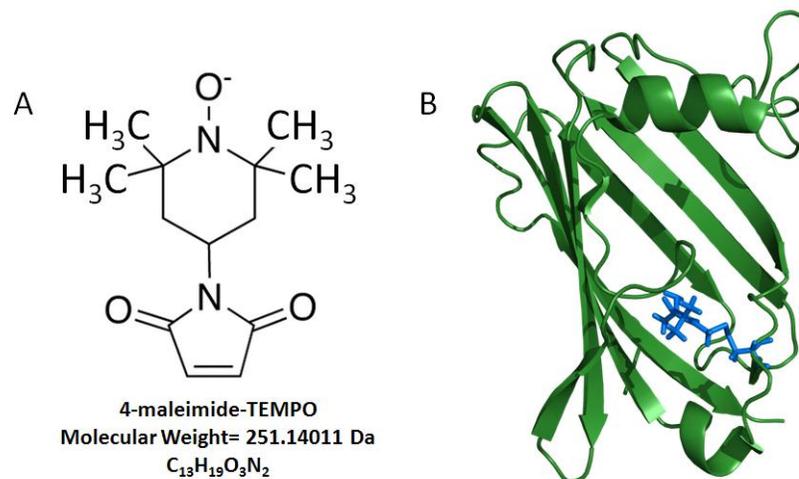


Figure 3-1. Structures of GM2AP and MTSL. A) Structure, molecular weight, and elemental composition of 4-maleimide-TEMPO. B : GM2AP crystal structure (PDB ID 1G13) with MTSL represented in blue at position L126.

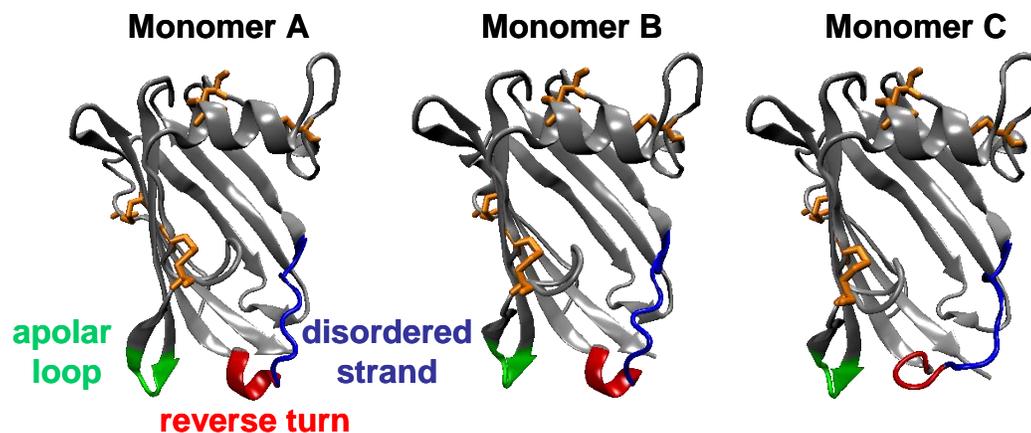


Figure 3-2. The three unique monomers (A, B, C) of apo GM2AP (PDB ID 1G13) showing the different conformations of the loop regions at the cleft entrance. The four disulfide bonds are shown in orange, the apolar loop (V54 - W63) is shown in green, the reverse turn (S130 - T133) is shown in red and disordered strand (V122 - P129) is shown in blue. Taken together, the reverse turn and disordered strand are collectively referred to within the text as a mobile loop.

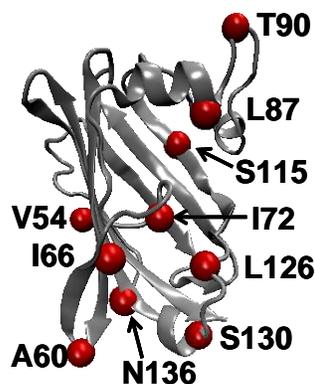


Figure 3-3. The red spheres in monomer A of GM2AP (PDB 1G13) represent the C α positions of the reporter sites chosen for SDSL EPR experiments.

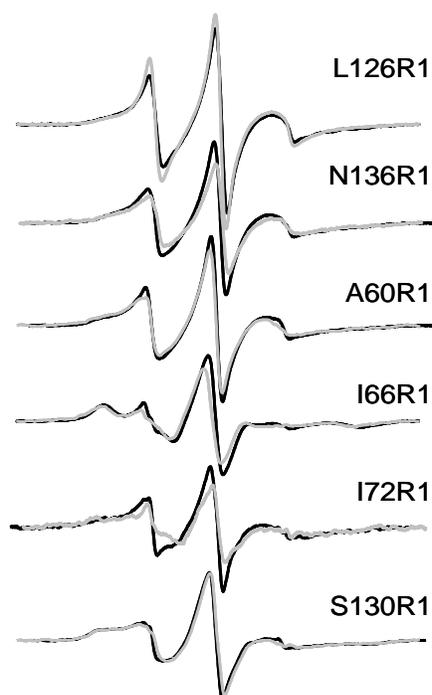


Figure 3-4. Stack plot of EPR line shapes from six sites in the flexible loops of GM2AP. Spectra were collected at ambient room temperature without temperature regulation. The black traces are for samples prepared in basic pH (8.0) whereas those in gray were collected at pH 4.8 in the presence of 4x molar excess GM2 micelles. All spectra have 100 G sweep widths. Note the spectra obtained for I72R1 in the presence of GM2 showed the most change, but this effect results from protein instability that lead to protein precipitation over time, with increased broadening of the spectral line shape.

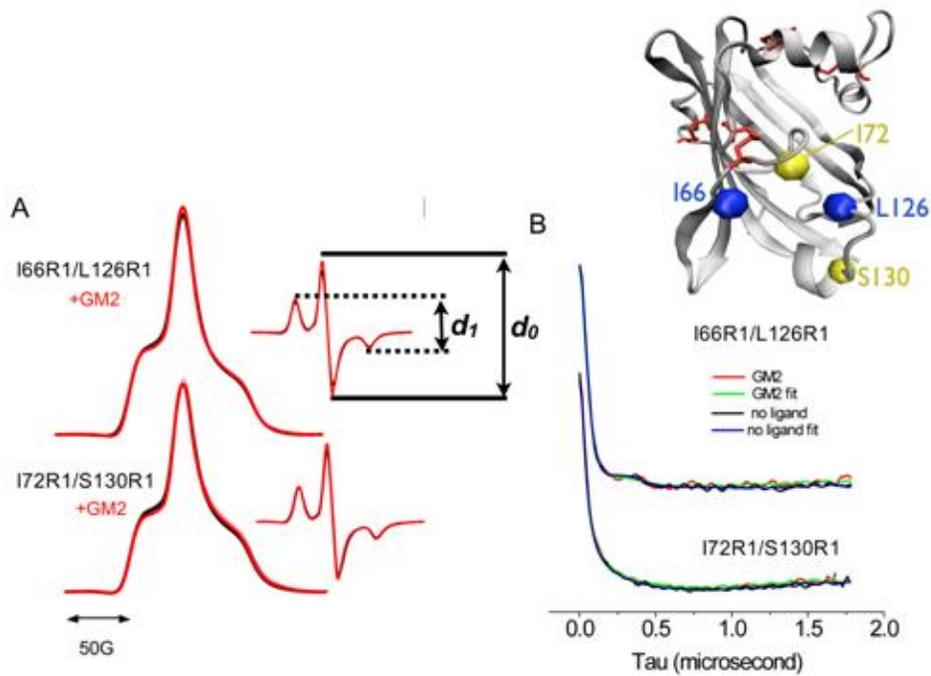


Figure 3-5. DEER data from GM2AP double mutants. A) Absorption and derivative CW EPR spectra of double CYS constructs I66R1/L126R1 and I72R1/S130R1 collected in the frozen state at -130 °C. B) Dipolar evolution spectra for double CYS constructs I66R1/L126R1 and I72R1/S130R1 collected in the frozen state at -208 °C.

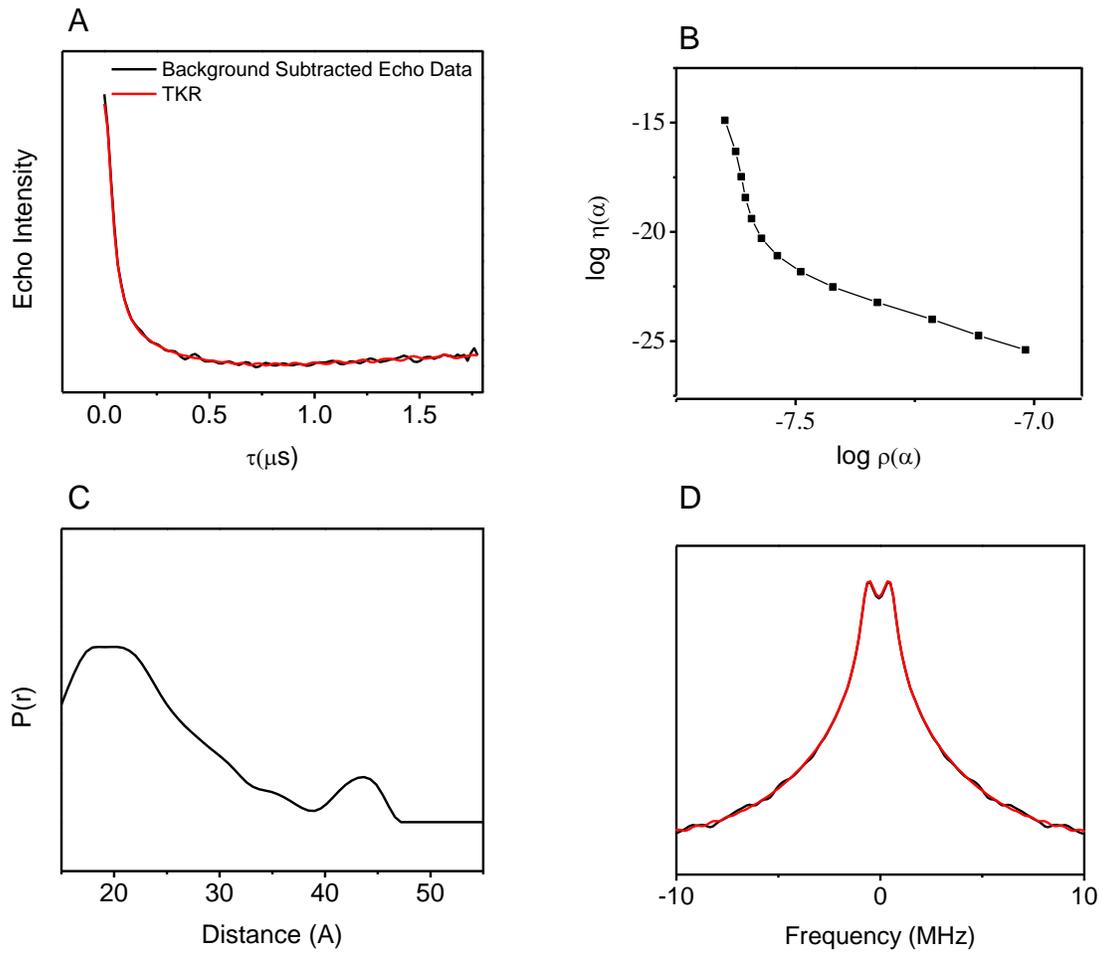


Figure 3-6. GM2AP construct I72C/S130C. A) Long pass filtered and background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red). B) L curve utilized to choose the appropriate regularization parameter ($\lambda = 5$). C) Distance profile from TKR analysis. D) Frequency domain spectrum.

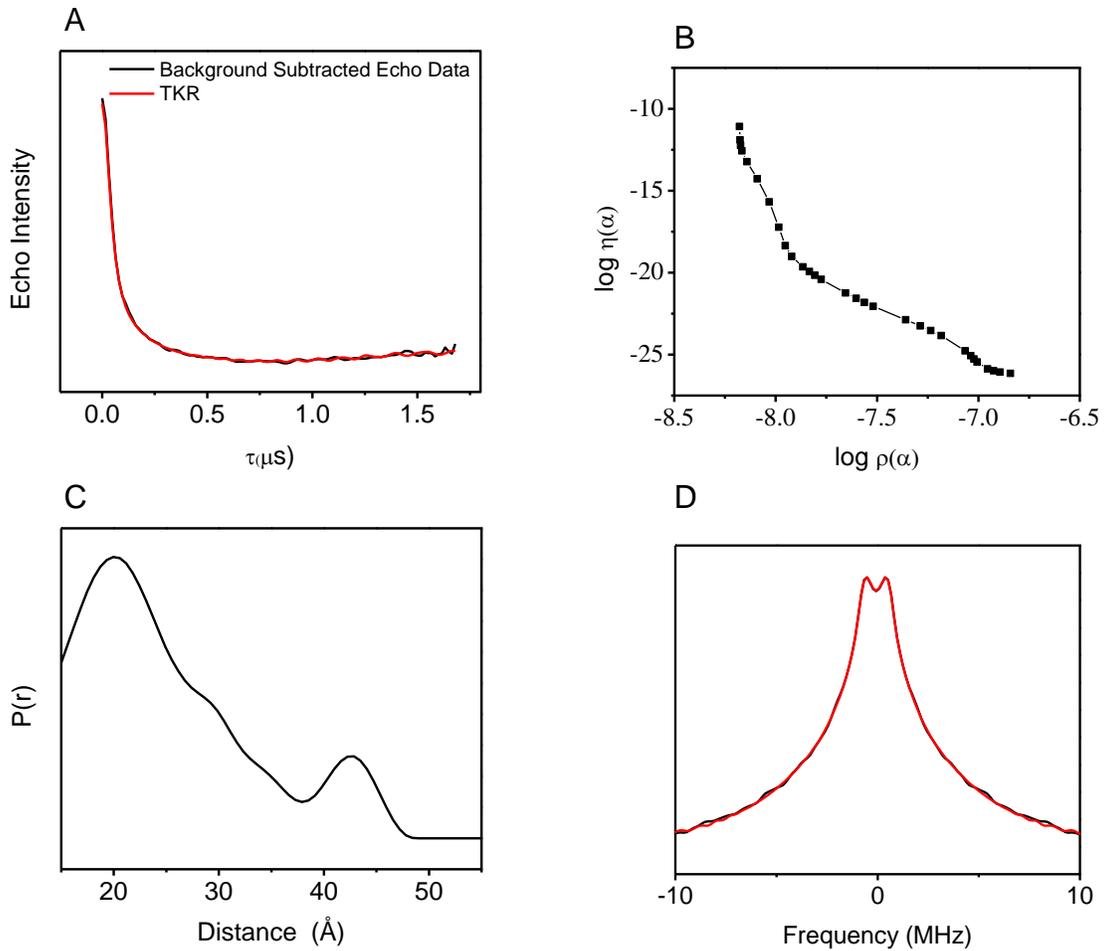


Figure 3-7. GM2AP construct I72C/S130C bound to GM2. A) Long pass filtered and background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red). B) L curve utilized to choose the appropriate regularization parameter ($\lambda = 5$). C) Distance profile from TKR analysis. D) Frequency domain spectrum.

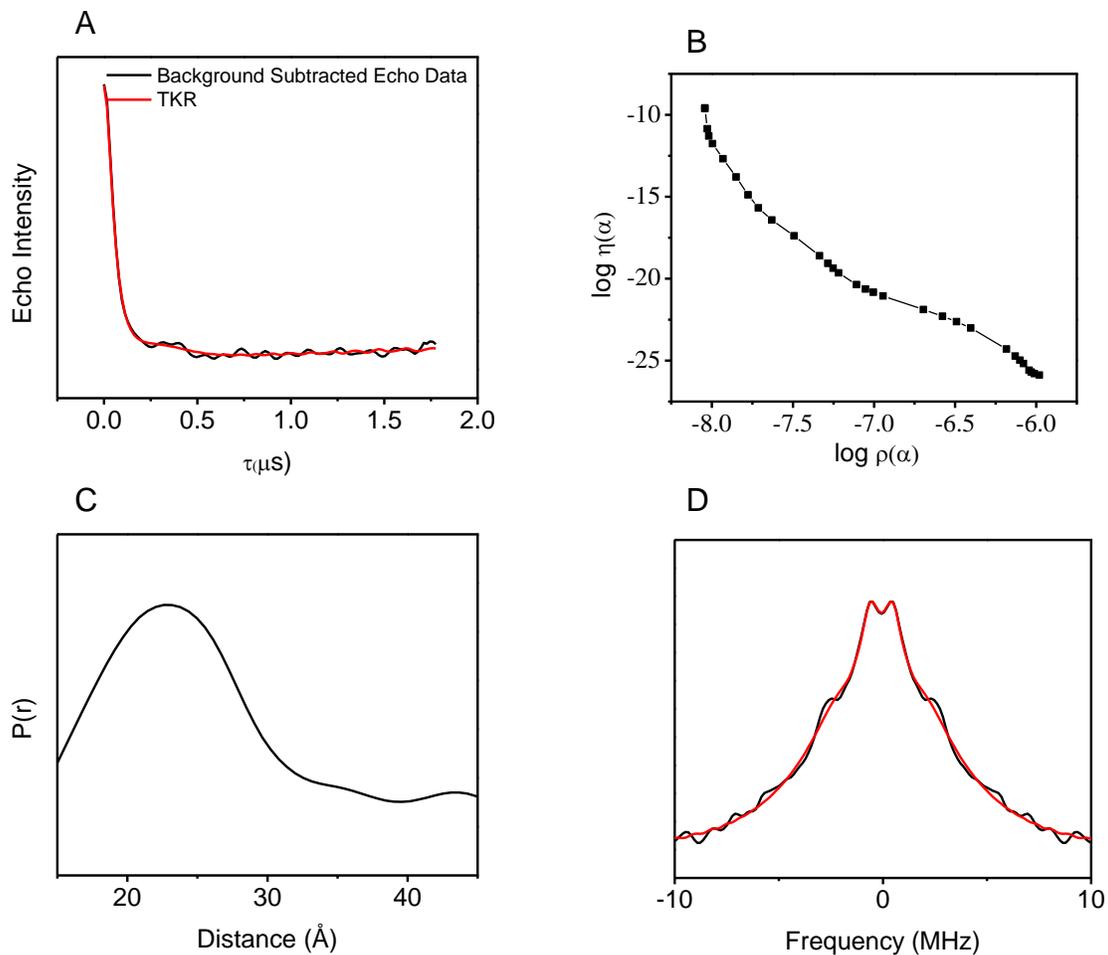


Figure 3-8. GM2AP construct I66C/S126C. A) Long pass filtered and background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red). B) L curve utilized to choose the appropriate regularization parameter ($\lambda = 5$). C) Distance profile from TKR analysis. D) Frequency domain spectrum.

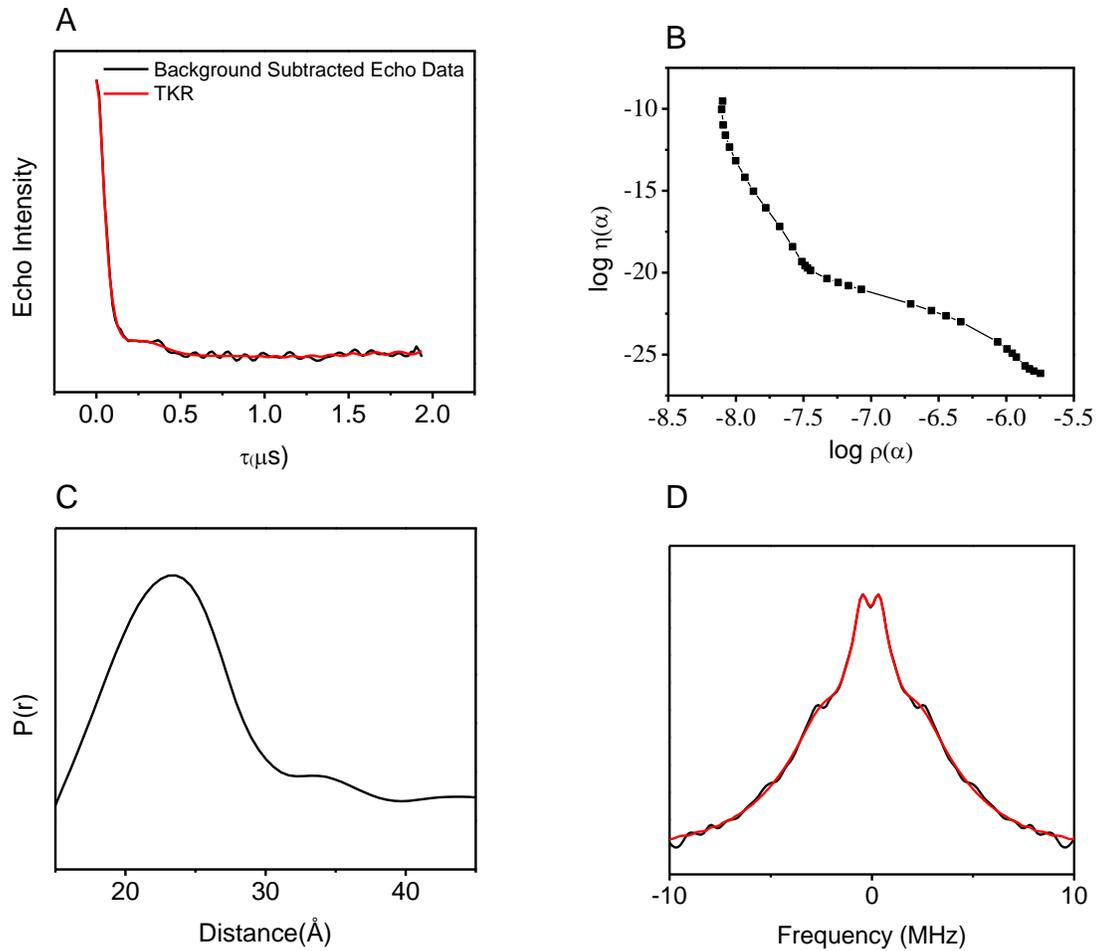


Figure 3-9. GM2AP construct I66C/L126C bound to GM2. A) Long pass filtered and background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red). B) L curve utilized to choose the appropriate regularization parameter ($\lambda = 5$). C) Distance profile from TKR analysis. D) Frequency domain spectrum.

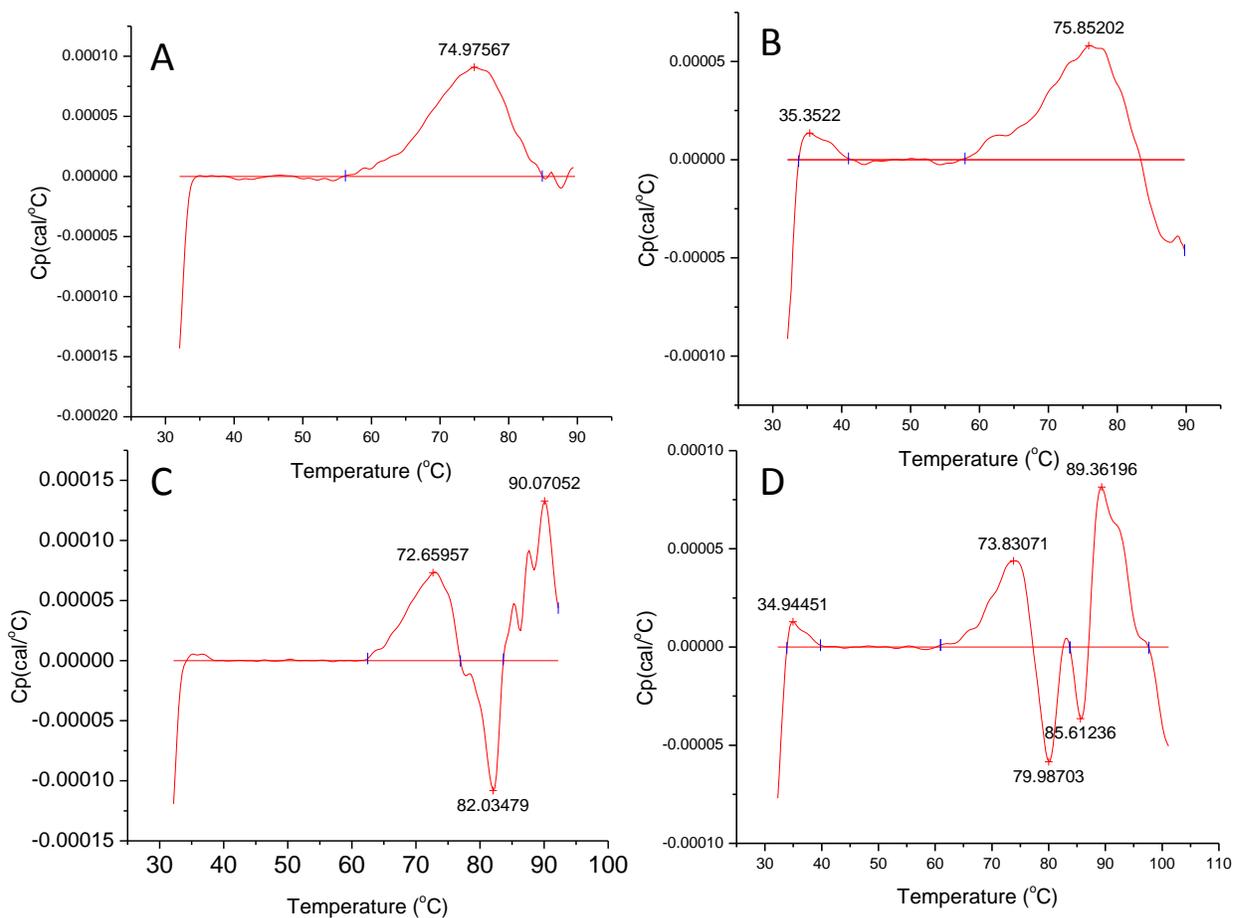


Figure 3-10. Differential scanning calorimetry for GM2AP constructs A60R1 and L126R1. A) Differential scanning calorimetry show the thermal of unfolding of A60R1. B) Differential scanning calorimetry show the thermal of unfolding of A60R1 with GM2 bound. C) Differential scanning calorimetry show the thermal of unfolding of L126R1. D) Differential scanning calorimetry show the thermal of unfolding of AL126R1 with Gm2 bound.

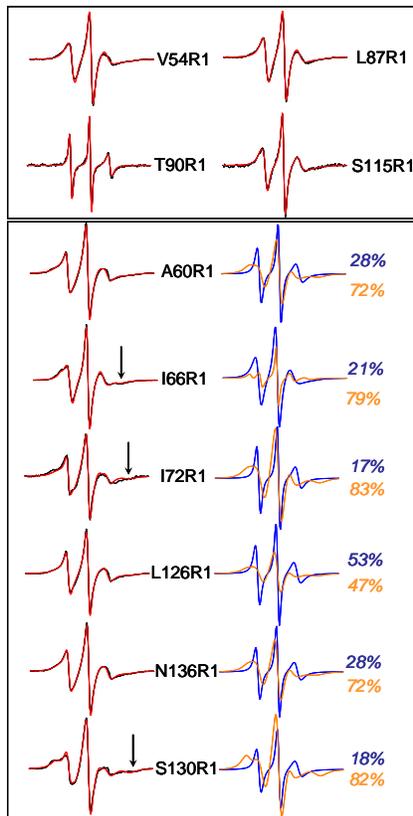


Figure 3-11. Overlay of area normalized EPR spectra (gray) and MOMD simulated spectra (black) of the ten spin labeled GM2AP constructs. All spectra have 100 G sweep widths and were collected under temperature control at 25°C. Adequate agreement between the experimental and simulated line shapes were obtained for sites V54R1, L87R1, T90R1, and S115R1 (top) using only a single spectral component. However, the spectra of the remaining sites (bottom) required two components for adequate agreement and the corresponding two spectral components are shown with relative percentages of mobile (gray) and immobile (black) components given.

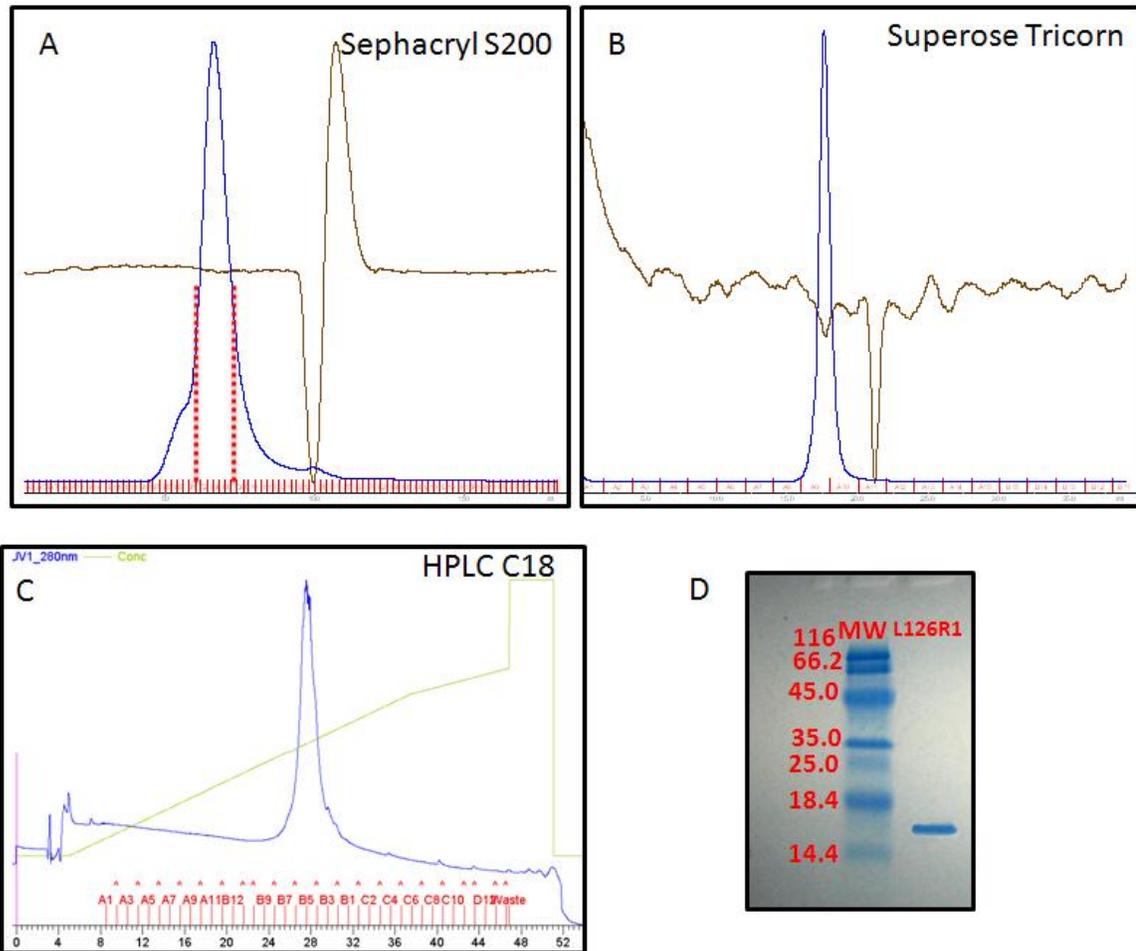


Figure 3-12. GM2AP chromatograms from protein purification. A) FPLC chromatogram for GM2AP L126R1 elution profile from S-200 sephacryl column dotted red lines represent fractions collected. B) FPLC chromatogram for GM2AP L126R1 elution profile from tricorn superose column. C) HPLC chromatogram for GM2AP L126R1 elution profile from C18 column. D) 16.5% Tris-HCL Biorad precast gel of broad range molecular weight marker in lane 1 and GM2AP L126R1.

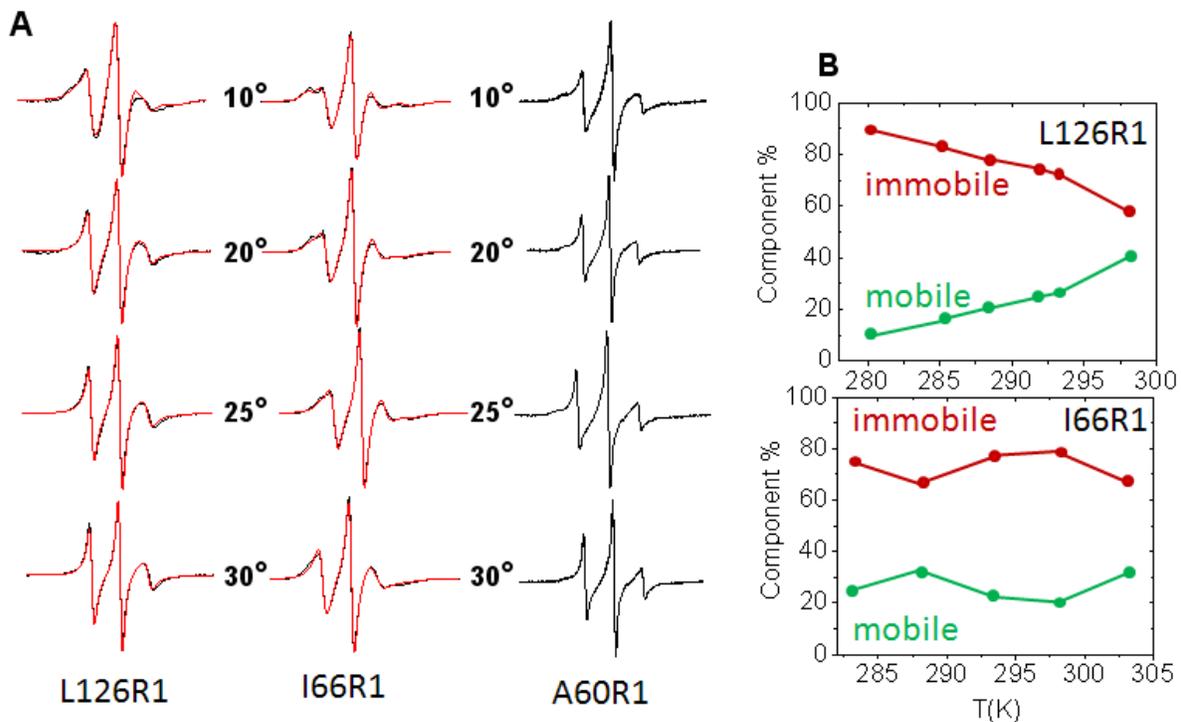


Figure 3-13. Representative EPR spectra of GM2AP L126R1 and I66R1 as a function of temperature ($^{\circ}\text{C}$). Overlain experimental spectra (black) and MOMD simulations (red) are shown. Each simulation required two spectral components for an adequate fit to the experimental data. B. Plots of the relative percentages of mobile and immobile components obtained from the spectral fits as a function of temperature (K).

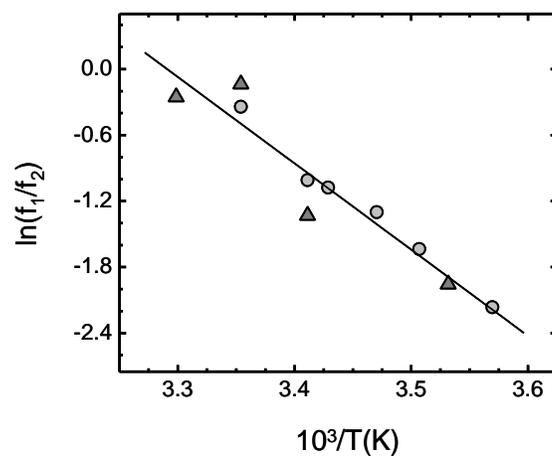
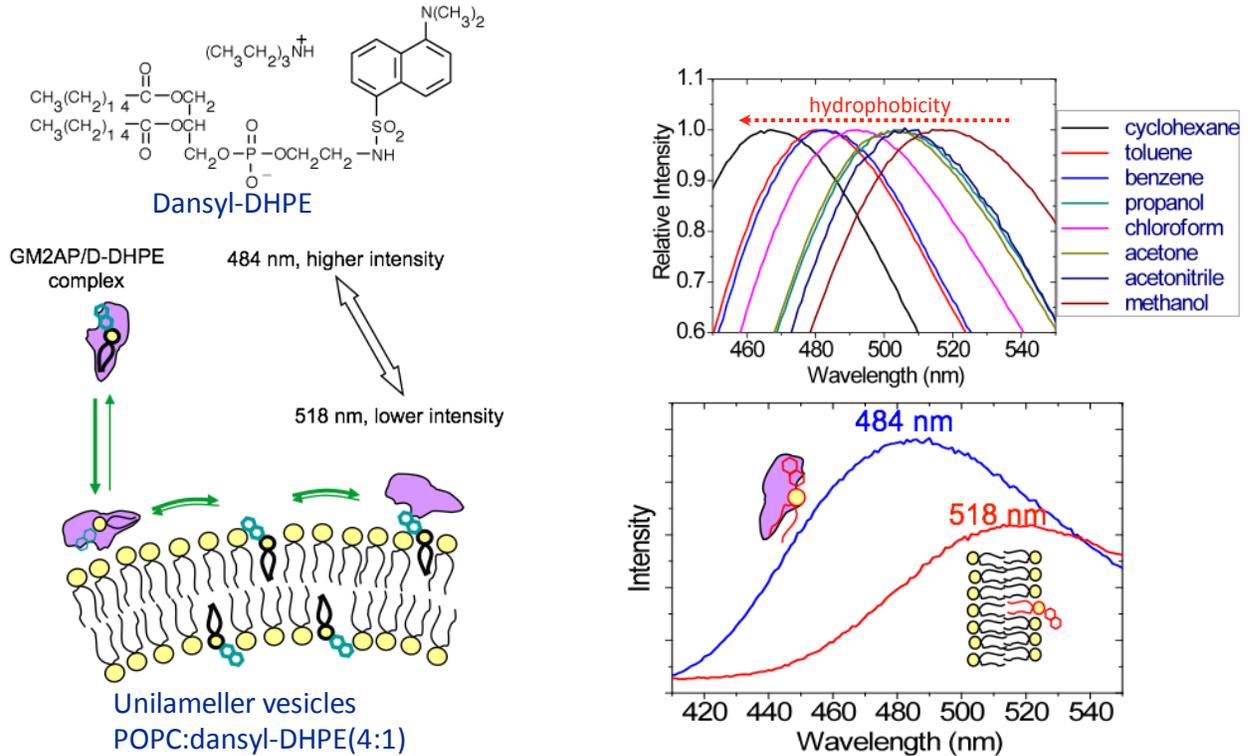


Figure 3-14. Linear regression of the ratio of the fractions of mobile and immobile L126R1 populations plotted as a function of inverse temperature. Two separate EPR data sets were collected from different protein sample preparations, and the simulations were performed independently for each set. The two data sets are indicated by triangles and circles. The solid line is the linear regression from both data sets taken together. From the slope of the line (-7.8 ± 0.8) and the intercept (25.8 ± 2.8), the enthalpy and entropy of loop motion was found to be 65 kJ/mol and 215 J/mol K, respectively. The value of the equilibrium constant K at 298.15 K was calculated to be 0.7.



11

Figure 3-15. Representative diagram of the dansyl-DHPE fluorescent lipid extraction assay. For the protein:lipid complex, the maximal emission fluorescence wavelength is near 484 nm, which is similar to dansyl-DHPE dissolved in benzene. A red-shift to 518 nm is obtained for extruded 100 nm large unilamellar vesicles of POPC:dansyl-DHPE(4:1), and is similar the emission wavelength when dansyl-DHPE is dissolved in methanol. 100 nmoles of GM2AP was prepared in 50 mM sodium acetate buffer at pH 4.8 was injected into 400 nmoles dansyl-DHPE (Ran, Fanucci, 2008).

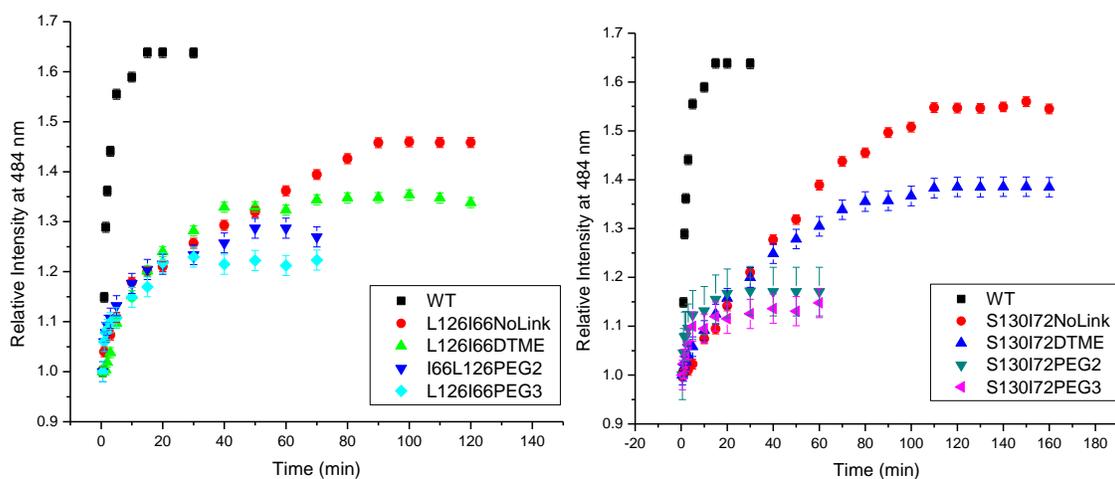


Figure 3-16. Tethering results from double mutant GM2AP constructs. A) Dansyl-DHPE extraction results for tethering the loops in I66C/L126C. B) Dansyl-DHPE extraction results for tethering the loops in I72C/S130C.

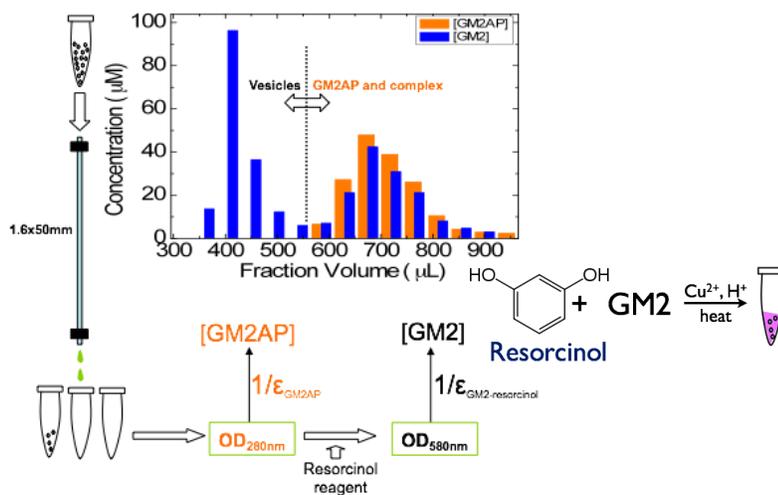


Figure 3-17. Representative diagram of the gel filtration GM2 lipid extraction assay. Gel filtration assay for GM2 extraction. The elution profile of the mixture of 7.5 nmol GM2AP with 200 nmol POPC:GM2:CHOL:BMP (50:10:20:20) vesicles in 50 mM NaOAc pH 4.8 buffer with total volume of 100 μL. The average fraction size was ~50 μL. The blue columns show the concentration of GM2 (determined from resorcinol assay) in each fraction. The black columns show the GM2AP concentration (determined from UV-VIS OD220) in each fraction. Fractions that contained vesicles were determined by light scattering at 550 nm.

CHAPTER 4
MASS SPECTROMETRIC ANALYSIS AND CONFIRMATION OF DISULFIDE
BONDING AND REPORTER SITE LABELING EFFICIENCY IN GM2AP

Introduction

Much of the work described in this chapter was published in an *Analytical Chemistry* article entitled “Sequential Proteolysis and High-Field FTICR MS To Determine Disulfide Connectivity and 4-Maleimide TEMPO Spin-Label Location in L126C GM2 Activator Protein,” by Jeremiah D. Tipton, Jeffrey D. Carter, Jordan D. Mathias, Mark R. Emmett, Gail E. Fanucci, and Alan G. Marshall (Tipton, Carter et al. 2009). Jeremiah Tipton was a post-doctoral fellow at the National High Magnetic Field Laboratory (NHMFL), and Mark Emmett and Alan Marshall are at the Department of Chemistry and Biochemistry at Florida State University in Tallahassee. The work was supported by NIH (GM-78359), NIH (R01GM077232), NSF Division of Materials Research through DMR-0654118, and the State of Florida (Tipton, Carter et al. 2009).

To validate the EPR results that conformational heterogeneity observed in EPR spectra arose from multiple conformations of GMAP not mis-labeling, a mass spectrometry protocol was developed to ensure that spin-label reactions only modified the mutant reporter cysteine (Cys). Specifically, this assay will be used to demonstrate that the spin label has attached to only the engineered reporter Cys that was introduced into the primary sequence via mutagenesis. Electrospray Ionization combined with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (ESI-FT-ICR MS) with a 14.5 T superconducting magnet was chosen to ensure proper labeling and disulfide bond formation in GM2AP.

The goal was to develop a novel mass spectrometry (MS) method to isolate Cys-residues in order to determine whether the Cys moieties were in disulfide bonds or accessible to modification (i.e. biologically active). This aim arose as a product of many difficulties experienced in obtaining preliminary mass spectra due to poor ionization efficiency of particular peptides of interest. For a series of GM2AP single Cys mutants a ninth CYS introduced into various sites of the protein to serve as a Cys reporter site for modification. Through a collaboration with Alan Marshall at the National High Field Magnetic Laboratory, (Tallahassee, FL) we were able to determine that the ninth reporter Cys residue was the only Cys residue modified by our labeling reagents and confirmed that the other eight Cys maintained the naturally occurring disulfide bonds.

Methods

Direct Infusion and Reversed-Phase Capillary Liquid Chromatography

Direct infusion of the intact proteins was performed with an Advion BioSystems Nanomate (Ithaca, NY). The intact proteins were diluted to 1-5 pmol/ μ L in H₂O/methanol/formic acid, 43/50/2 (v/v/v). Digests were separated by reversed-phase nanoliquid chromatography (10 μ L sample) with a C18 capillary column (New Objective, Woburn, MA] 5 cm 0.75 μ m, 15 μ m i.d spray tip). An Eksigent NanoLC (Dublin, CA) was used to deliver a 35 min gradient (5 to 95% B) at 400 nL/min with solution A as 0.5% formic acid (v/v) in 5% aqueous methanol and solution B as 0.5% formic acid (v/v) in 95% aqueous methanol.

Mass Spectrometry

Mass spectrometry was performed with either a modified hybrid linear quadrupole ion trap FTICR mass spectrometer (LTQ-FT, Thermo Fisher Corp., Bremen, Germany) equipped with an actively shielded 14.5 T superconducting magnet (Magnex, Oxford,

U.K.) or a custom-built 9.4 T FTICR MS (Senko, Hendrickson et al., 1996; Hakansson, Chalmers et al., 2003; Marshall, Guan, 1996). The protease-treated samples were screened by online LC/MS with the 14.5 T LTQ FTICR instrument operated in top-5 data-dependent mode (high resolution FTICR, low resolution LTQ CID (collision-induced dissociation) MS/MS. For each precursor ion measurement, 1 million charges were accumulated in the LTQ prior to transfer (1 ms transfer period) through three octopole ions guides (2.2 MHz, 250 Vp-p) to a capacitively coupled open cylindrical ICR cell for analysis (Beu, Laude et al., 1992). The robust automatic gain control for transferring the same number of ions for each FTICR MS scan yielded less than 0.500 ppm rms mass error with external calibration (Schwartz, Senko et al., 2002). Each of the five most abundant ions was collisionally dissociated in the LTQ for low resolution MS/MS (3 microscans, 10,000 target ions, 2.0 Da isolation width, 35% relative collision energy, 0.250 activation q, 30 ms activation period, and dynamic exclusion list size of 60 for 1 min). Data were collected with Xcalibur 2.0 software (Thermo Fisher). Following the *identification* of disulfide-containing peptides from the first pass low resolution LTQ MS/MS analysis (as described in the Data Analysis and Informatics section), the samples were reanalyzed by nano-LC, selected-ion high resolution MS/MS to confirm charge state and fragment ion assignments.

The second experiment is needed for verification of disulfide connectivity. The scan type for the high resolution MS/MS was set at 7 to 10 scans per segment. The first scan was the same as the precursor ion scan found in the top-5 data-dependent experiment. The next 2 to 9 scans were set to select relevant ions (disulfide connected peptide fragments) for high resolution MS/MS. The 2 or 3 most abundant charge states

(for relevant disulfide connected proteolytic fragment targets) were selected for high resolution MS/MS analysis because different charge states are known to fragment differently with CID (Tangm Thibault et al., 1993; Griffiths, Jonsson, 2001; Tsapraili, Somogyi et al., 1999). A selected-ion high resolution MS/MS scan consisted of a target ion accumulation target of 50,000 ions with a 5 Da mass selection window, CID fragmentation in the LTQ (35.0 scaled collision energy, 0.250 activation q, and 30 ms activation period), and transfer to the cell for measurement. Low resolution MS/MS spectra exhibiting ambiguous fragment assignments were easily assigned by high resolution MS/MS based on charge state validation and mass errors less than 0.500 ppm. High resolution MS/MS data were collected with both Xcalibur 2.0 and MIDAS software. Electron capture dissociation (ECD) was performed with a custom-built 9.4 T FTICR and is known to specifically break disulfide bonds more readily than the polypeptide backbone (Zubarec, Krunger et al., 1999; Zubarev, Krunger et al., 2004; Uggerud, 2004).. Thus, disulfide bonds not verified by high resolution CID MS/MS were analyzed by ECD FTICR MS. Ions were selected with a quadrupole mass filter and accumulated in a second octopole for 1 s (Senko, Hendrickson et al., 1997). The ions were then transferred through multiple ion guides and trapped in an open-ended cylindrical Penning trap by gated trapping for ECD analysis (Tsybin, Hendrickson et al., 2006). Accurate mass and time-tag knowledge of disulfide peptides from the top-5 data-dependent analysis of the different digest facilitated the use of “peak parking” with the Eksigent nano-LC (Pasa-Tolic, Masselon, 2004). Namely, at the elution time for the peptide of interest, the flow rate was reduced from 400 nL/min to 50 nL/min so that 25 ECD spectra could be signal-averaged. Data acquisition was performed by use of a

Predator data station and data analysis by MIDAS 3.4 software (Blakney, Robinson et al., 1990).

Sequential Digestion with Trypsin, GluC, and AspN

The solvent-exchanged sample was diluted to 50 pmol/ μ L (i.e., 50 μ M), with a final volume of 160 μ L. To that solution, 100 ng of trypsin (final substrate:trypsin ratio 1000:1) was added and incubated overnight at 36 °C. Next, a 100 μ L aliquot of the trypsin digest was incubated overnight with 200 ng of endoproteinase GluC (substrate:GluC, 400:1) at 36 °C. Finally, 50 μ L of the trypsin/Glu C digested solution was incubated with 100 ng of endoproteinase AspN (substrate:AspN, 400:1). Each sample was diluted to 1 pmol/ μ L prior to LC MS analysis.

Data Analysis and Informatics

Peaks from the Xcaliber files were extracted with a customized peak picking algorithm, thresholded at 10% of the maximum peak magnitude for low resolution MS/MS and 1% for high resolution MS. The resulting files were searched with MASCOT (Matrix Science, Cambridge, UK) against a custom-created database containing the wild-type and L126C mutant sequences. To aid in identification of the MSL labeled Cys, the mass of the MSL label (252.14739) was added to the MASCOT modification database. Results from MASCOT were visualized and validated with Scaffold (Proteomics Software, Portland, OR) software. After *identification* of the MSL labeled peptide, the remaining disulfide peptides from each digest were *identified* by comparing high-resolution MS precursor ion measurements to values calculated from the elemental composition of the assigned peptides. High-resolution MS/MS and ECD fragment ion assignments were performed by hand.

Results and Discussion

Mass spectrometry enables direct identification of posttranslational modifications, including disulfide bonds (Tipton, Carter et al. 2009; Bean, Carr et al., 1992; Badock, Raida et al., 1998; Gorman, Wallis, 2002; Xu, Zhang et al., 2008; Yen, Joshi et al., 2002; Nair, Nilson et al., 2006; Popolo, Ragni et al., 2008). Because of the analytical challenge associated with verifying multiple disulfide bonds, a multitier assay was developed (Tipton, Carter 2009). First, the expressed and purified GM2AP constructs (Protein gel shown in Figure 4-1) were analyzed to determine the mass of the intact protein by direct infusion ESI 14.5 T LTQ FTICR MS (Figure 4-2). A 240 Da mass shift was observed between the wild-type and spin labeled constructs, corresponding to the addition of one MSL label at the engineered reporter Cys (L126-MSL). Also, a 480 Da shift was observed in the double mutant I66-MSL/L126-MSL. Although, initial intact protein experiments observed that the I66C construct provided a poor labeling reporter site, which is most likely due to the Cys side chain residing in the hydrophobic cleft. The spectra also show several salt and oxidation adducts, both common for intact protein analysis. A second criterion for identification of the correct construct is presented in Figure 4-3, top. If there are less than four disulfide bonds, the 16+ charge state isotopic distribution, calculated from the molecular formula for the L126C-MSL GM2AP construct, will decrease by $0.12598 \text{ Da} \left(\frac{2(1.007825)}{16} \right)$ for each addition of 2 hydrogens. Thus, a positive match between the calculated and experimental isotopic distributions is first checked before enzymatically digested samples are analyzed. For this case, Figure 4-3, bottom illustrates an excellent agreement between the measured and experimental isotopic distributions for the 16+ charge state isotopic distribution of the L126C-MSL GM2AP construct. Thus, mass analysis of intact L126C-MSL GM2AP

unambiguously confirms the presence of four disulfide bonds and a single MSL modification but not the disulfide bonding connectivity.

Following the identification of the intact protein a systematic approach was used to verify proper folding and labeling of a recombinantly overexpressed and MSL labeled L126C GM2AP construct. A sequential proteolytic digestion scheme was used, and include the enzymes, trypsin, endoproteinase GluC, and endoproteinase AspN. Next, a combination of standard proteomic techniques and multiple scan functions with high resolution MS and/or low/high resolution MS/MS were performed on a modified 14.5 T LTQ FTICR MS. FTICR MS provides high mass accuracy and allows for quick peptide and fragment ion assignments with rms mass errors of less than 0.5 ppm (Schaub, Hendrickson et al., 2008). Tryptic digestions were analyzed with LC high-resolution MS and low-resolution MS/MS. The low resolution MS/MS was chosen at first because of the slower duty cycle associated with high resolution MS/MS. Figure 4-4 (top) presents the sequence of L126C GM2AP, expected disulfide connectivity, location of the reporter Cys, and sites of proteolytic cleavage for different proteases. MASCOT/Scaffold analysis identifies the tryptic digest fragments shown in red in Figure 4-4, bottom but cannot identify the two overlapping disulfide-containing peptides (Figure 4-4, bottom). MASCOT is typically used to analyze proteomic data sets in which the protein sample from cell lysate (including proteins containing disulfide bonds) have been reduced with dithiothreitol and alkylated with iodoacetamide. Thus, in the previous mentioned case peptides that contain Cys will have reacted with iodoacetamide making the disulfide connectivity pattern ambiguous. This experiment

was designed to retain the disulfide connectivity. Thus, the current version of MASCOT is not useful to identify peptides that are linked by disulfide bonds.

However, based on the known primary sequence, accurate mass measurement of the tryptic digest peptide precursor ions, and sequence tags from subsequent MS/MS, manual analysis revealed all of peptides containing *one or more* disulfide linkages, thus providing 100% sequence coverage. In addition, from Figure 4-4 (top), the connectivity for tryptic peptide Cys-1 (linking C8 to C152), containing a *single* disulfide bond (see Figure 4-5) could be uniquely determined. Thus, from the predicted sequence and the results of the tryptic digest, the expected tryptic disulfide peptide sequences and masses could be generated. To accurately *verify* the disulfide bound peptide fragments generated the need for the sequential digestion platform.

Furthermore, high mass accuracy reduces the number of possible matches for a measured peptide to the possibilities found in a database (or protein) of known sequences. For example, if the tryptic peptide IESVLSSSGK, mass 1005.5342 Da, is searched against the possible peptide fragments from the sequence of GM2AP at 1, 20, and 2000 ppm mass error, the numbers of matched peptide sequences are 1, 4, and 7, respectively. Thus, high mass accuracy yields higher confidence in assignments by reducing the number of possible peptide matches.

The peptide containing the MSL spin label was observed and confirmed with a mass error of 0.340 ppm, Figure 4-5 shows the observed b- and y-ion series for the tryptic peptide containing the MSL reporter Cys (in green). In contrast to other labile post-translational modifications, the data illustrates that the MSL thioether bond does not break during CID fragmentation, because the b9-15 and y13-14 ion series retain

the mass associated with MSL addition (Cooper, Hakansson et al., 2005; Syka, Coon et al., 2004; Beausoleil, Jedrychowski, 2004). Furthermore, if a population of the expressed protein were not properly folded, other non-reporter Cys residues would be modified by the MSL label. This is not the case because other MSL modified non-reporter Cys were not observed.

Furthermore, verification of disulfide connectivity was performed from sequential trypsin, Glu-C, and Asp-N enzymatic digestions. From the tryptic peptide assignments and the predicted subsequent Glu-C and AspN proteolytic cleavage sites, the remaining disulfide-connected peptides were generated and denoted as Cys-2a, Cys-2b, and Cys-2c (see Figure 4-6). Charge state m/z values are also calculated for nano-LC selected-ion high resolution MS/MS. Figure 4-7 presents the fragment ions from the disulfide-containing peptides identified by high mass accuracy MS and CID MS/MS. From the trypsin digest, Cys-1 (Figure 4-7, top; Table 4-1) resulted in b- and y-ions that retained the mass increase associated with the second, shorter LGCIK sequence. High resolution MS/MS provided unambiguous assignment of the fragment ions for Cys-1; i.e., at low resolution/mass accuracy in the LTQ many of the fragments had assignment ambiguities. From the trypsin/Glu C digestion, Cys-2a was verified (Figure 4-7, bottom; Table 4-2) with a tripeptide sequence tag. The Cys precursor ion mass measured to within 0.340 ppm mass error (Table 4-3) includes the loss of 2 hydrogens for the disulfide bond formed between C68 and C75. As for Cys-2 and Cys-2b, high resolution CID MS/MS provided a sequence tag for identification but did not reveal the disulfide bond connectivity. Cys-2 and Cys-2b are essentially “cyclic” peptides and are thus expected to fragment poorly by low-energy CID (Bean, Carr et al., 1992; Badcock,

Raida et al., 1998). Therefore, those peptides were fragmented by ECD with a custom-built 9.4 T FTICR MS. Both peptides resulted in only a charge-reduced species, not fragmentation. The trypsin/Glu C sample was therefore digested with Asp N to release Cys-2c (Figure 4-6). Once again, high resolution MS/MS sequencing could not verify the final two disulfide bonds; however, peptide identity was confirmed.

Figure 4-8 presents the spectrum after “peak parking” ECD FTICR MS. Once peak parking is initiated, i.e., flow rate slowed from 200 nL/min to 50 nL/min, the separation will be less than optimal for the remainder of the analysis. However, the ECD experiment was designed for analysis of only Cys-2c, thus the loss in separation efficiency for the remainder of the gradient was irrelevant. ECD is known to specifically fragment disulfide bonds (Zubarec, Krunger et al., 1999; Zubarev, Krunger et al., 2004; Uggerud, 2004). In this case, ECD specifically broke the C94-C105 disulfide bond to generate Frag2 (Figure 4-8) and EPC⁹⁴PEPLR. ECD analysis thus provided the final piece of evidence to verify the last two disulfide bonds. Finally, Table 3 summarizes the tryptic peptides identified by MASCOT, along with the manually identified disulfide bonds. The data acquired from all three sequential digests by accurate mass measurement MS and MS/MS sequence tags resulted in quick identification of all four disulfide peptides as well as the sequence location of the MSL-Cys. The result is a high-confidence MS assay that confirms the 4 native disulfide bonds are intact in GM2AP-MSL labeled constructs and that only the reporter Cys is labeled with the paramagnetic spin label.

Table 4-1. Peak list and mass errors from high-resolution CID MS/MS of Cys-1

Ion	Charge	Experimental (m/z)	Experimental Mass (Da)	Calculated Mass	Error (ppm)
b ₈	1	978.3773	977.3700	977.3996	0.466
b ₁₁	2	920.8761	1839.7377	1839.7383	-0.348
b ₁₁ -H ₂ O	2	911.8707	1821.7269	1821.7277	-0.478
b ₁₂	2	985.3975	1968.7805	1968.7809	-0.223
b ₁₄	2	1077.9551	2153.8957	2153.8973	-0.761
b ₁₅	2	1135.4694	2268.9243	2268.9242	0.026
b ₁₅	3	757.3150	2268.9232	3368.9242	-0.447
b ₁₅ -H ₂ O	3	751.3115	2250.9127	2250.9136	-0.420
y ₅	1	555.3612	544.3539	554.3535	0.821
y ₉	2	492.7772	983.5399	983.5395	0.386
y ₉ -H ₂ O	2	483.7719	965.5295	965.5395	-0.212
y ₁₂	2	923.9610	1845.9081	1845.9082	-0.347
y ₁₂	3	616.3098	1845.9081	1845.9082	-0.279
y ₁₃	2	981.4744	1960.9351	1960.9351	-0.428
y ₁₄	2	1074.5148	2147.0151	2147.0144	0.307
y ₁₅	2	1118.0305	2234.0465	2234.0464	0.027
y ₁₇	2	1235.0806	2468.1467	2468.1469	-0.097
y ₁₈	3	852.7333	2555.1781	2555.1789	-0.319
Y ₁₉	3	896.4133	2686.2181	2686.2194	-0.490

Table 4-2. Peak list and mass errors from high-resolution CID MS/MS of Cys-2a

Ion	Charge	Experimental (m/z)	Experimental Mass (Da)	Calculated Mass	Error (ppm)
M-2H ₂ O	1	705.7891	1409.5637	1409.5631	0.390
b ₁₀	1	1051.4226	1050.4153	1050.4150	0.271
b ₁₀ -H ₂ O	1	1033.4120	1032.4047	1032.4044	0.286
b ₁₁	1	1152.4700	1151.4627	1151.4627	0.004
b ₁₁ -H ₂ O	1	1134.4594	1133.4521	1133.4521	0.013
b ₁₂	1	1299.5388	1298.5316	1298.5311	0.358
b ₁₂ -H ₂ O	1	1281.5282	1280.5209	1280.5205	0.308

Table 4-3. Peak list and mass errors for relevant peptides in GM2AP L126C-MSL

Experimental (m/z)	Experimental Mass (Da)	Calculated Mass	Error (ppm)	Peptide
568.2611	2269.0151	2269.0148	0.138	GHHHHHHHHHHSSGHIEGR
670.3883	669.3810	669.3810	0.021	DPAVIR
1601.8937	3201.7728	3201.7747	-0.596	SLTLEPDPIVVPGNVTLSSVGS TSVPLSSPK
408.2472	814.4799	814.4800	-0.110	DLVLEK
458.2685	914.5224	914.5225	-0.084	EVAGLWIK
447.7318	893.4490	893.4494	-0.456	EGTYSLPK
855.7492	2564.2256	2564.2257	-0.014	SEFVVPDC*ELPSWLTTGNYR
503.7744	1005.5342	1005.5342	0.015	IESVLSSSGK
706.8267	2823.2777	2923.2777	0.000	Cys-1
845.8801	5069.2368	5069.2374	-0.120	Cys-2
723.7993	2562.2109	2562.2100	0.340	Cys-2a
911.4231	3641.6632	3641.6635	-0.126	Cys-2b
869.3866	2605.1378	2605.1378	-0.002	Cys-2

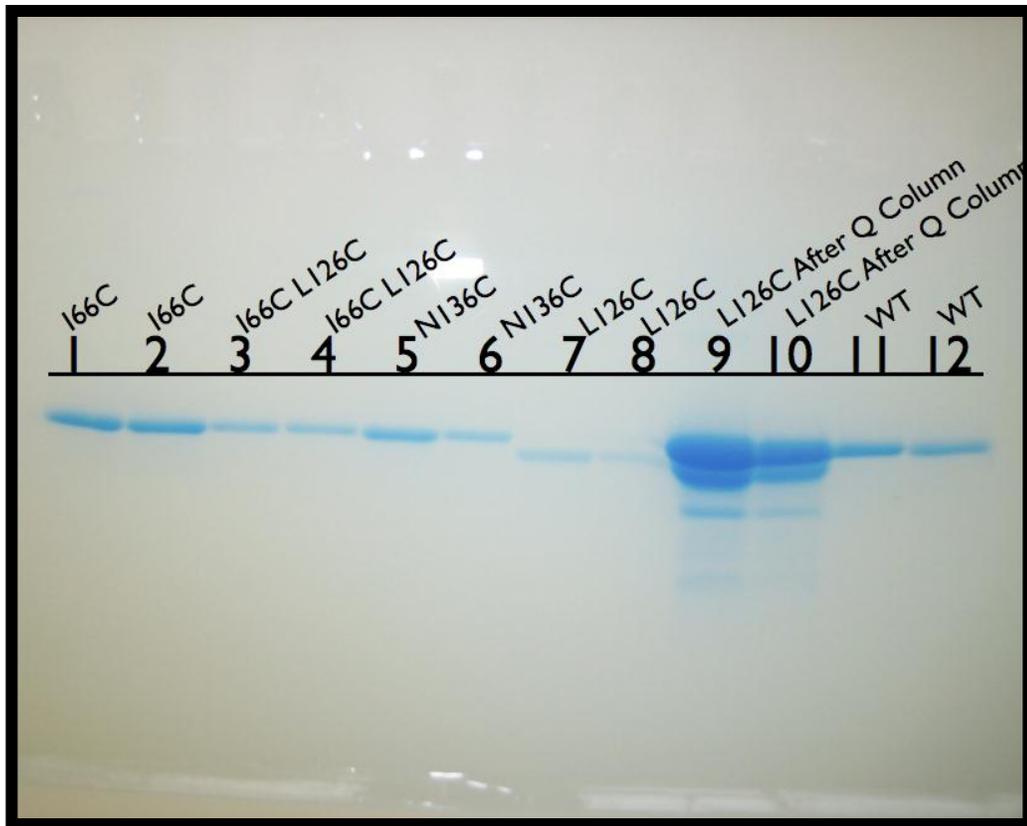


Figure 4-1. 16.5% Tris-HCL Biorad pre-cast gel of GM2AP constructs purified for MS analysis. Lanes 1-8 and 11-12 show GM2AP constructs after purification by sephacryl S-200 column. Lanes 9-10 show pre-sephacryl protein sample after concentration by an anion exchange Q-column.

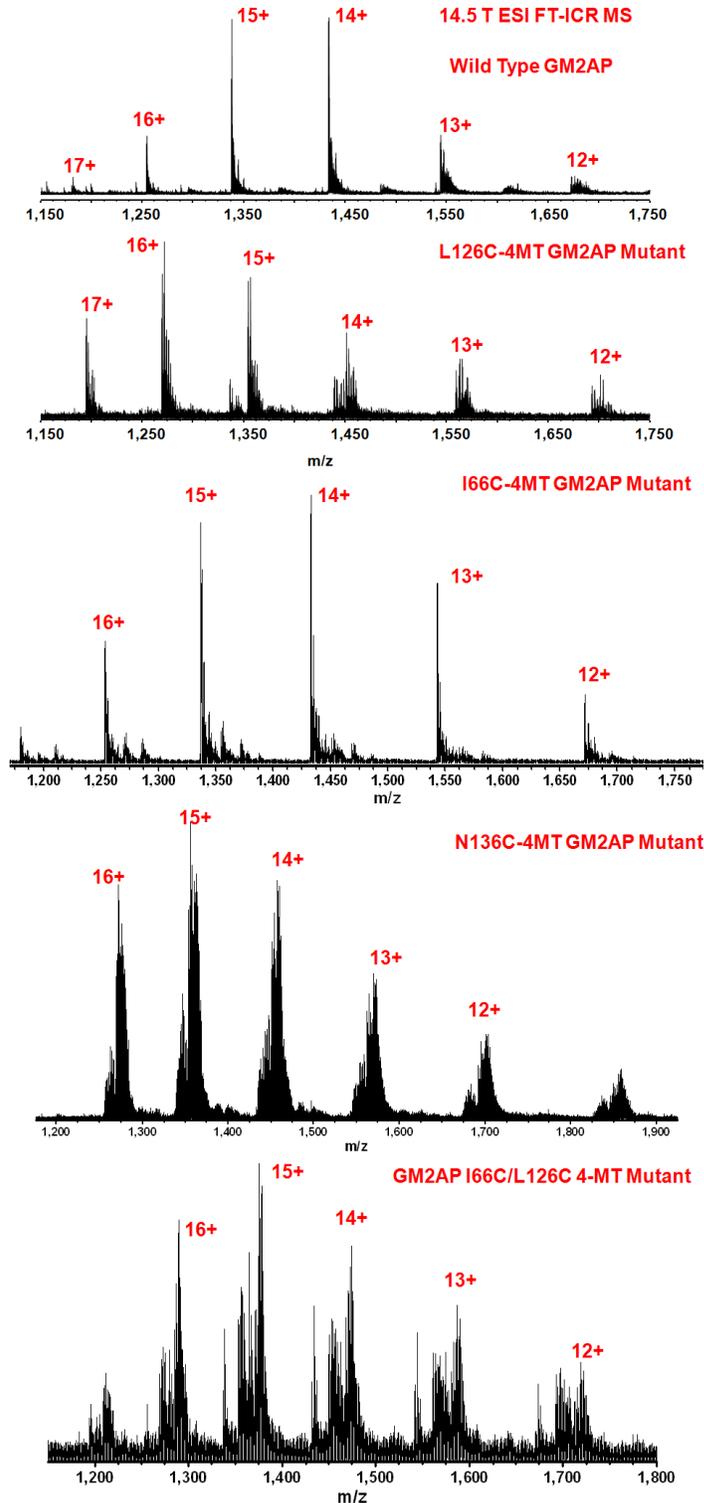
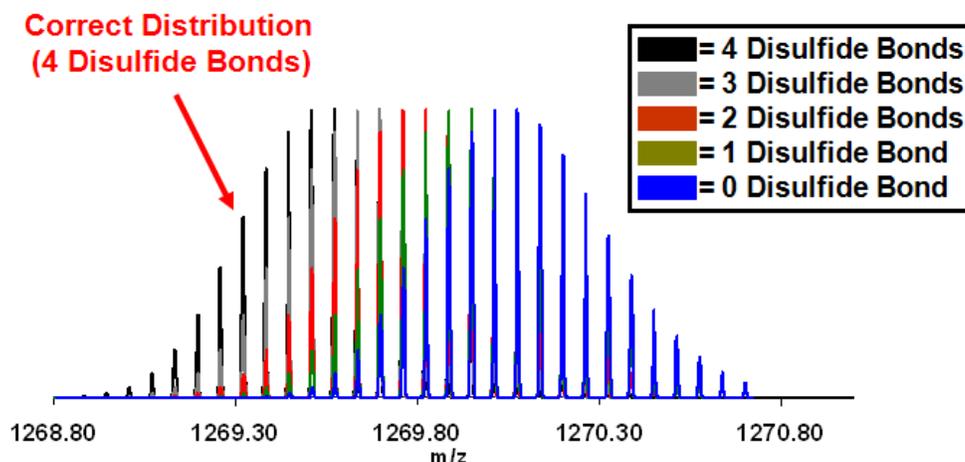


Figure 4-2. Broadband 14.5 T electrospray positive-ion of intact wild-type, L126C-MSL, I66C-MSL, N136C-MSL, I66C/L126C-MSL respectively. The recombinant CYS GM2AP constructs were labeled with 4-maleimide-TEMPO. The shift in the m/z for the observed charge states indicates labeling of the protein. Adapted from Tipton, Carter et al. 2009.

Calculated Isotopic Distributions for L126C GM2AP (16+ Charge State)



Isotopic Distribution for L126C-4MT GM2AP with 4 Disulfide Bonds

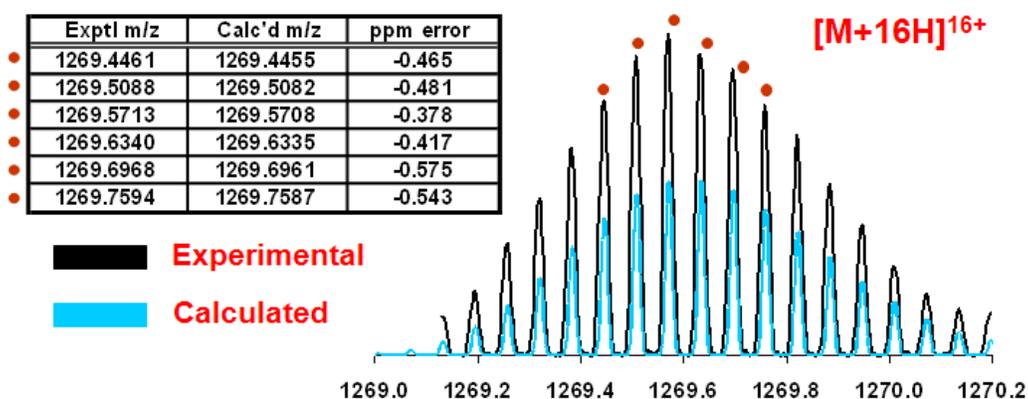


Figure 4-3. Parent ion FTICR data for GM2AP-MSL. Top: Calculated isotopic distributions corresponding to 0, 1, 2, 3, and 4 disulfide bonds for L126C-MSL GM2AP $[M+16H]^{16+}$. The calculated m/z distribution decreases by $(2)(1.007825)/16 = 0.1260$ for each additional disulfide linkage. Bottom: Calculated and experimental distribution confirms that the protein has the correct amino acid sequence, with four disulfide bonds, and a single MSL label on a cysteine residue. The five highest-magnitude in the isotopic distribution were chosen for comparison to calculated masses (red dots). Adapted from Tipton, Carter et al. 2009.

L126C GM2AP



L126C-4MT GM2AP Sequence Coverage after Solution Trypsin Digestion and LC/MS

GHHHHHHHHHSSGHIE**GR**H**MSSFSW**D**M**C**D**E**G**K**D**P****
AVIRSLTLEP**DPIVVPGNVTL**S**VVGST**S**VPLSS**P**L**K**V**D**L**
VLEKEVAGLWIKIPC**TDYIG**S****C**T**F**E**H**F**C**D**V**L**D**M**L**I**P**T**G**E**
PC**P**E**P**L**R**T**Y**G**L**P****C**H**C**P**F**K**E**G**T**Y**S**L**P**K**S**E**F**V**V**P**D****C**¹²⁶**
EL**P**S**W**L**T**T**G**N**Y**R**I**E**S**V**L**S**S**S**G**K**R**L**G**C**I**K**I**A**A**S**L**K**G**I****

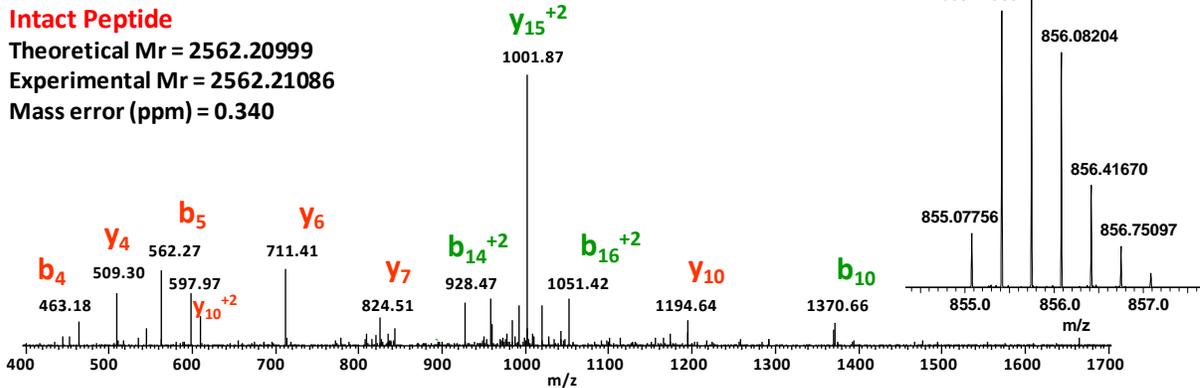
— **MASCOT Identification**
— **Manual Identification**

Figure 4-4. Protease digestion scheme for GM2AP L126-MSL. Top : L126C GM2AP amino acid sequence with the location of reported disulfide bonds and predicted cleavages for trypsin (blue :K,R), Glu C(dark green :E), and AspN (orange :D). The location of the L126C amino acid substitution is shown in red. Bottom : Sequence coverage for the L126C-MSL Gm2AP construct after trypsin digestion and nano-LC analysis. The green sequence was identified by MASCOT ; the blue sequence was identified by hand and corresponds to the peptides still containing disulfide bonds. Adapted from Tipton, Carter et al. 2009.

GM2AP L126C mutant - MSL labeled peptide – MS/MS

Intact Peptide

Theoretical Mr = 2562.20999
 Experimental Mr = 2562.21086
 Mass error (ppm) = 0.340



Current verified fragments

Green label = Fragment + Label

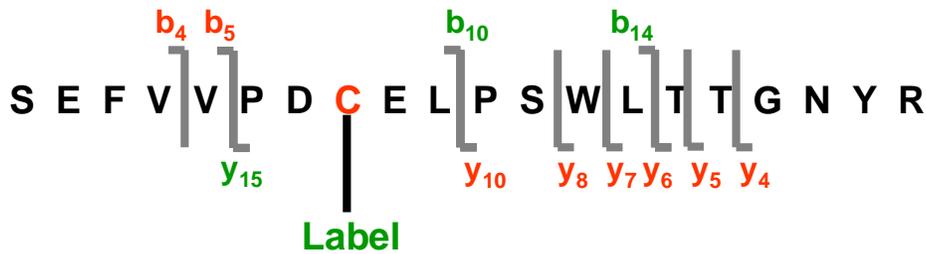


Figure 4-5. FTICR results for GM2AP L126-MSL peptide. Top left: identified peptide containing MSL label from L126C-MSL tryptic digests for CID MS/MS. Bottom: Identified fragments by low resolution CID MS/MS fragments in green contain the MSL label. Adapted from Tipton, Carter et al. 2009.

Disulfide Bonds Verified by High Mass Accuracy CID MS/MS

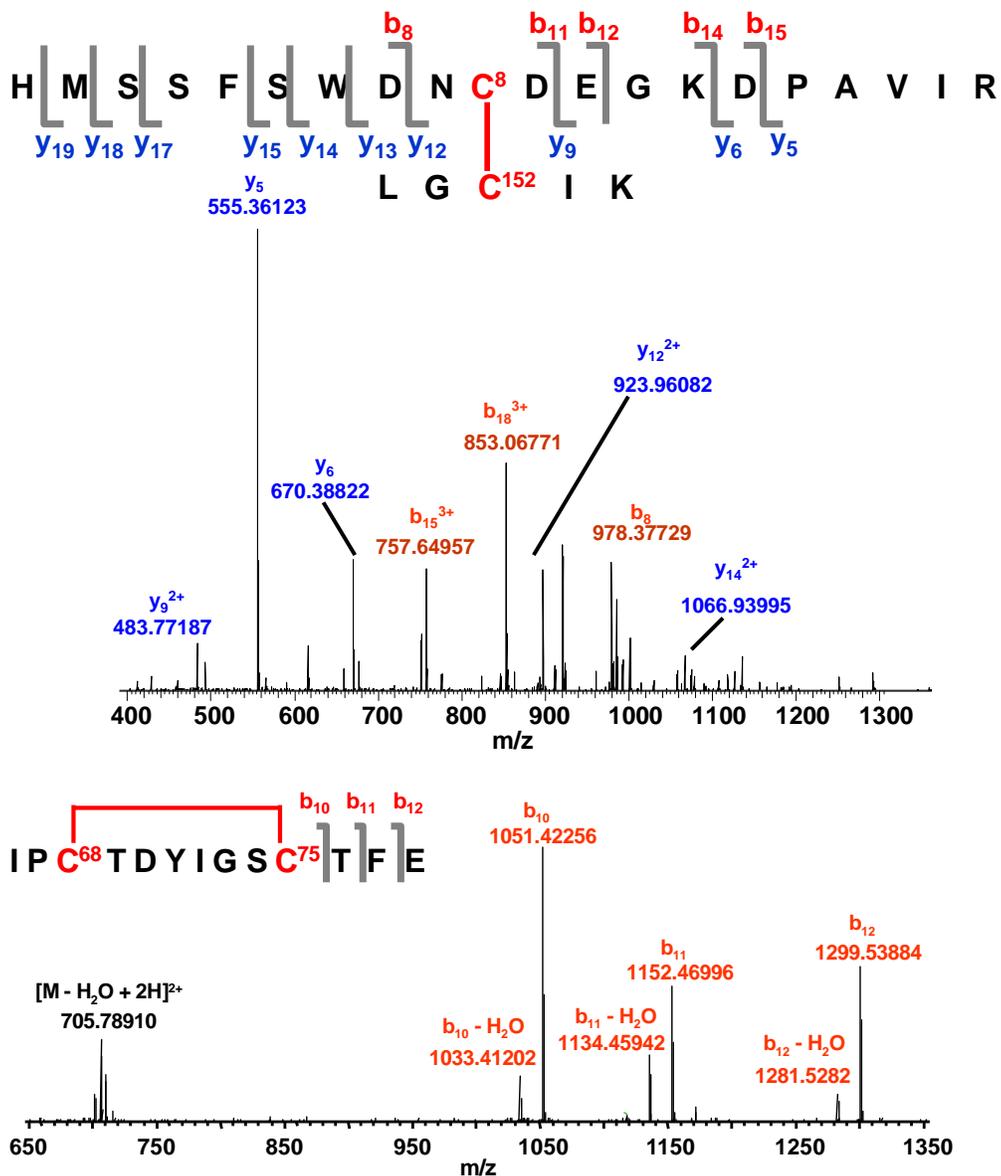


Figure 4-7. CID MS/MS disulfide identification. Top: Disulfide bonds that were verified through CID MS/MS analysis. Top : high resolution MS/MS of Cys-1. Several large peptide fragment masses verify the peptide sequence. Bottom: high-resolution MS/MS of Cys-2a. The calculated mass for the intact peptide matches the experimental mass corresponding to one disulfide bond (Table 1). A sequence tag of three amino acids verifies the identity of the disulfide-containing peptide. As expected, the cyclic peptide does not yield fragments. Adapted from Tipton, Carter et al. 2009.



Ion	Exptl m/z	Calc'd m/z	Error (ppm)
Frag2	1667.7072	1667.7046	1.6
Frag2 - H ₂ O	1649.7037	1649.6941	4.8
Frag2 - H ₂ O - NH ₄	1632.6663	1632.667	-0.4
EPCPEPLR	939.4465	939.4478	-1.4
EPCPEPLR - H ₂ O	922.4451	922.4465	-1.5

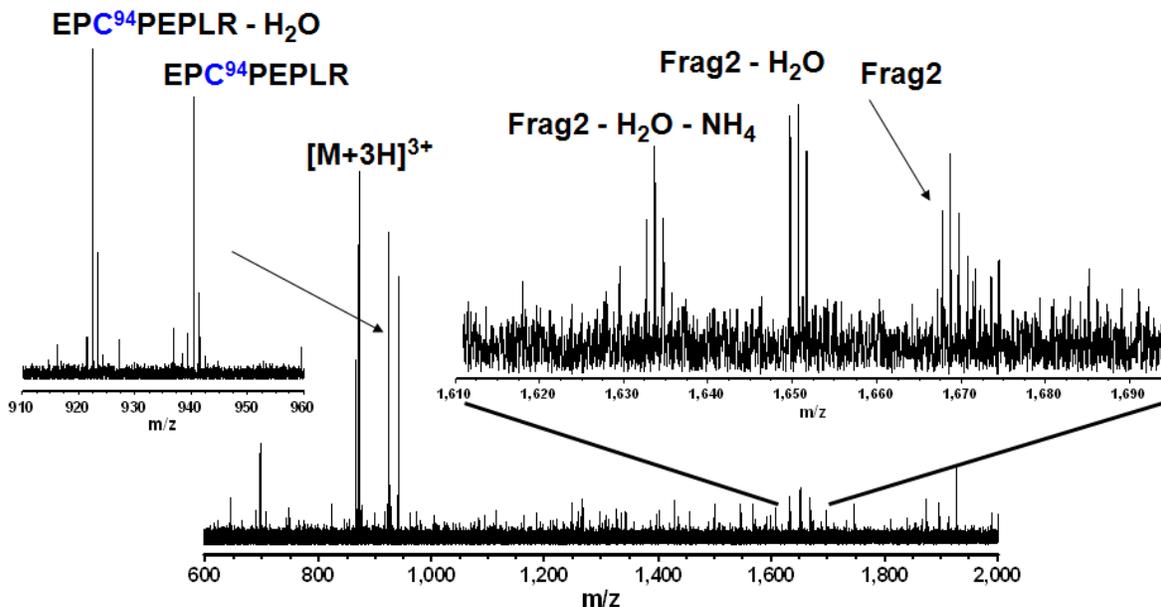


Figure 4-8. ECD MS/MS disulfide identification. Peak-parked nano-LC high resolution ECD MS/MS (25 scans, 3 min). ECD typically breaks disulfide bonds prior to amino acid backbone fragmentation. Two fragments corresponding to the specific fragmentation of the C⁹⁴-C¹⁰⁵ disulfide bond were observed. Adapted from Tipton, Carter et al. 2009.

CHAPTER 5 DOUBLE ELECTRON-ELECTRON RESONANCE STUDIES OF HIV-1 PROTEASE

Introduction and Previous Work

HIV-1 protease (HIV-1PR) is essential for the maturation of the retrovirus that continues HIV replication, infection, and progression to AIDS. The aspartic protease homodimer cleaves polyproteins *gag* and *gag-pol* into functional proteins. The active site pocket houses catalytic D25 residues, and is physically mediated by the conformational changes of the β -hairpin “flap” region. HIV-1PR is a target of antiretroviral drugs (ARVs) that slow the rate of viral maturation (Ashorn, McQuade et al., 1990). The ribbon diagram in Figure 5-1 details critical regions of HIV-1PR (PDB ID 2PBX), including the active site pocket, the flap and hinge regions, and the sites L63 and A71.

The double electron-electron resonance (DEER) technique, in combination with site-directed spin-labeling (SDSL), is an excellent tool for examining changes in the conformational ensemble of the HIV-1 protease flap region (Galiano, Bonora et al., 2007; Galiano, Ding et al., 2009, Kear, Blackburn et al., 2009; Torbeev, Raghuraman et al., 2009; Pannier, Velt et al., 2000). DEER measures the strength of the dipolar coupling between two electron spins (here, spin labels) spaced at distances between 20-60 Å, where the coupling is proportional to r^{-3} , with r representing the distance between the spins (Jeschke, Chechik et al., 2006; Ding, Layen et al., 2008). Conformers designated as wide-open, semi-open, closed, and tucked/curled have been identified with SDSL DEER and X-ray crystallography techniques and modeled using molecular dynamic simulations (Kear, Blackburn et al., 2009; Torbeev, Raghuraman et al., 2009; Freedberg, Ishima et al., 2002; Hornak, Okur et al., 2006,

Ishima, Freedberg, 1999; Todd, and Freire, 1999; Todd, Luque et al., 2000; Todd, Semo et al., 1998; Velazquez-Campoy, Todd et al., 2000; Goodenow, Bloom et al., 2002) .

Initial work on DEER measurements of HIV-1PR was performed by former Fanucci group member Dr. Luis Galiano (Ph.D. 2008). At that time, we did not have the capability here at the University of Florida (UF) to collect DEER data, and Galiano's experiments were performed in the lab of Peter Fajer at the National High Magnetic Field Lab (NHMFL) with the help of post-doctoral research fellow Marco Bonora. This original data was collected at X-band on a Brüker EleXsys E580/E680 equipped with the ER 4118X-MD5 Dielectric Ring Resonator. Dr. Galiano's work demonstrated the success of the DEER methodology when applied to the study of flap conformations in Subtype B_{SI} HIV-1PR. These results showed clear differences in the most probable distance between the flaps and the flexibility of the flaps in the presence and absence of the protease inhibitor Ritonavir (Galiano, Bonora et al. 2007).

The dipolar modulated echo data for the apo construct (blue) is noticeably different than that of the protease in the presence of protease inhibitor Ritonavir (red), as shown in Figure 5-2A. The resulting distance distribution profiles (Figure 5-2B) report very different flap conformations and flexibility. The average distance between the labeled sites on each flap shifted by about 3 Å (from 36 Å to 33 Å) from open to closed upon binding Ritonavir. The breadth of the distance distribution profile decreased dramatically upon addition of the inhibitor, indicating that the flexibility of the flaps decreased as well, demonstrating a structural mechanism of inhibitor resistance.

Steve Kent and co-workers obtained active and inactive (D25N) K55MTSL-labeled HIV-1 protease constructs via total chemical synthesis and, using DEER, reported interflap distances in the presence of three peptidomimetic inhibitors. Each of The inhibitors, called MVT-101, KVS-1, and JG-365, capture different stages of the peptide bond hydrolysis reaction, where MVT-101 locks the protease in an early transition-like state, KVS-1 mimics the tetrahedral intermediate in the reaction, and JG-365 mimics a later transition-like state (Baca and Kent 1993; Torbeev, Mandal et al. 2008). The respective distance distribution profiles suggested that the flaps adopted different catalytic properties throughout the course of the catalytic reaction, with the flaps predominantly closed and with minimal flexibility during the early stages of the reaction, with a more open conformation with increased flexibility aiding in product release as the reaction proceeds (Torbeev, Raghuraman et al. 2008).

Dr. Luis Galiano and Dr. Ralph Weber, Senior Applications Scientist for Brüker's EPR division, later focused on understanding the role of secondary polymorphisms on acquired drug-resistance by the protease. V6 and MDR769 are patient isolates where V6 is from a pediatric patient treated with RTV and MDR769 was isolated from a patient previously treated with IDV, NFV, SQV, and APV. Findings showed that mutations that arise in response to PI treatment alter the flap conformations of the apoenzyme, thus affecting the conformational ensemble of the protease. In order to provide structural insight into the experimental data,

MD simulations were performed in the lab of Carlos Simmerling, and the structural data obtained was in excellent agreement with the previous experimental data. The Subtype B flap region adapts an average conformation described as semi-

open conformation by X-ray crystallographic studies. MDR769 adopts a more open average conformation and V6 a more closed average conformation with respect to Subtype B. These results, featured in an issue of *Chemical and Engineering News* (Drahl 2009), were important because they demonstrated that drug-induced polymorphisms affect flap conformations and flexibility, a likely contributor to drug resistance.

Dr. Mandy E. Blackburn compared the distance distribution profiles of HIV-1PR in the apo form with those in the presence of nine separate FDA-approved protease inhibitors and a substrate mimic (Galiano, Blackburn et al. 2009), leading to a publication found among the top ten most accessed *Biochemistry* rapid reports in July-September 2009. HIV-1PR distance profiles were split into two groups by inhibitor, those that showed a “strong/moderate” affect on flap closing, and those that revealed a “weak” affect. Inhibitors defined as having strong interactions must show at least 70% of the conformational ensemble in the closed flap conformation, and include all but IDV, NFV, and ATV, which were placed in the “moderate/weak” affect category because the flaps are predominantly found in the semi-open conformation. Moderate affects were seen for ATV, where approximately 40% of the conformational ensemble is in the closed conformation, while IDV and NFV had less than 20% of the conformational ensemble in the closed conformation indicating a “weak affect (Galiano, Blackburn et al. 2009). No strong correlation seems to exist between the relative percent closed conformation and the K_i or K_D values. The DEER data does, however, seem to be in agreement with the number of non-water mediated hydrogen bonds between the inhibitor and the protease.

Dr. Jamie L. Kear focused on examining conformational sampling in the apoenzyme of various subtypes and patient isolates to determine how naturally evolved and drug-induced polymorphisms alter the effect of inhibitors on the conformational ensemble and flap flexibility of non-B subtypes and drug-resistant patient isolates from protease inhibitor exposed patients and found that the flap conformational ensembles differed greatly among the apo protease constructs. These results demonstrated that natural and drug-induced polymorphisms in the amino acid sequence of various subtypes and patient isolates, whether because of naturally occurring amino acid substitutions or drug-pressure selected mutations alter the average flap conformations and flexibility of the flaps. These results were featured in an issue of *AIDS Weekly* (30 Nov. 2009).

Here, SDSL DEER has been used to further the examination of HIV-1PR flap conformational ensembles. DEER data are generally collected at cryogenic temperatures due to a phase memory time that is too short for detection of the spin echo at room temperature. The question is often raised on whether the conformations trapped at 65K accurately represent the thermodynamic conformations sampled. Distances between spin-labeled sites in the flap region of an apo, and CA-p2, Indinavir (IDV) and Tipranavir (TPV)-bound HIV-1PR Subtype C construct were measured following four different freezing conditions: a slow -20 °C freeze, an intermediate liquid nitrogen freezing freeze, and rapid freezes from both 25 °C and 37 °C using isopentane in combination with liquid nitrogen (Figure 5-3).

Additionally studies were done to determine the conformational ensembles for four protease constructs, namely PRE and POST, and L63P and A71V secondary

mutants. PRE and POST were derived from a pediatric subject infected via maternal transmission with HIV (Ho, Coman et al., 2008) before and after PI treatment; respectively. The *gag-pol* alleles were isolated from serial blood samples obtained over 7 years, starting before therapy initiation (PRE) and after the development of multiple drug resistance following 77 weeks of PI therapy with Ritonavir (RTV) and an additional 16 weeks with Indinavir (IDV) (POST) (Barrie, Perez et al., 1996). PRE contains polymorphisms caused by the normal process of genetic drift. Results indicate that PRE adopts a predominately closed flap conformation, but that the emergence of drug-pressure selected mutations after (POST) causes the conformational sampling to favor a semi-open population. Also investigated were the effects of two secondary mutations (away from the active site cavity, Figure 5-4), L63P and A71V on the conformational ensemble. L63P is hypothesized to alter the conformational sampling resulting in a more closed state in the PRE construct. A71V, a common secondary mutation seen in multi-drug resistant constructs (Clemente, Hemraini et al., 2003), is thought to induce a more closed conformation.

A minimal amount of early work was done to incorporate and site-directed spin-label an unnatural amino acid into HIV-1PR at position 55 of the flaps. This method differs from the traditional approach to site-directed spin-labeling, in which a thiol-reactive spin label is attached to a native or non-native CYS residue. Previous work from the lab of Wayne Hubbel reported an orthogonal labeling strategy that does not label any functional group found within the 20 naturally occurring amino acids (Fleissner, Brustad et al., 2009). In this method, Fleissner utilized the genetically encoded unnatural amino acid p-acetyl-L-phenylalanine (p-AcPhe) (Figure 5-5) and

reacted it with a hydroxylamine reagent to yield a nitroxide side chain termed K1 (Figure 5-6). They were able to demonstrate success using seven different mutants of T4 lysozyme, each containing a single p-AcPhe at a solvent-exposed helix site. It was shown that the EPR spectra of the K1 mutants had higher nitroxide mobilities than the spectra of mutants containing the R1 side chain. Here we were successful in expressing HIV-1 protease K55p-acetyl-L-phenylalanine in large-scale inclusion bodies, purifying, refolding (with mediocre success) in acidic solution, and spin labeling. A low-quality DEER dataset was obtained, and the results will be discussed here.

Materials and Methods

Materials

The chemicals, reagents, and supplies were obtained from Fisher Scientific (Pittsburg, Pennsylvania) and used as received, with a few noted exceptions. pET23 plasmid DNA was purchased from Novagen (Gibbstown, New Jersey). HiTrap Q HP Anion Exchange column, HiPrep 16/60 Sephacryl S-200 High Resolution Size Exclusion column was purchased from GE Biosciences (formerly Amersham, Pittsburg, Pennsylvania). 4-Maleimido-2,2,6,6-tetramethyl-1-piperidinyloxy (4-Maleimido-TEMPO, MSL) was purchased from Sigma-Aldrich (St. Louis, MO). (1-Oxyl-2,2,5,5-tetramethyl- Δ^3 -pyrroline-3-methyl) Methanethiosulfonate spin label (MTSL) was purchased from Toronto Research Chemicals, Inc. (North York, Ontario, Canada). 0.60 I.D. x 0.84 O.D. capillary tubes were purchased from Fiber Optic Center (New Bedford, Massachusetts). BL21*(DE3) pLysS *E. coli* cells were purchased from Invitrogen (Carlsbad, California). NaOAc ($C_2D_3O_2$), D_2O , d_8 -glycerol were purchased from Cambridge Isotope Labs (Andover, MA). Ritonavir, Indinavir, Tipranavir, Darunavir,

Amprenavir, Atazanavir, Nelfinivir, Saquinavir, and Lopinavir were generously received from the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH (Bethesda, Maryland) (NIH). The non-hydrolysable substrate mimic, CA-p2 (H-Arg-Val-Leu-r-Phe-Glu-Ala-Nle-NH₂ (R-V-L-r-F-E-A-Nle-NH₂, r = reduced) was synthesized and purchased from the University of Florida Protein Chemistry Core Facility (Gainesville, Florida).

Methods

Expression of HIV-1 protease

pET23a vectors with HIV-1PR genes were transformed separately into *E. coli* strain BL21*(DE3)pLysS via standard heat-shock methodology. The transformed cells were inoculated in 5 mL sterile Luria Bertani (LB) media (Table 3-9) and grown at 37 °C with shaking at 250 rpm to an optical density (OD₆₀₀) of approximately 0.60, then transferred to 1 L sterile LB media and grown to an OD₆₀₀ of approximately 1.0 with shaking at approximately 200 RPM. Cells were then induced using 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG). The culture was then incubated with shaking at 250 rpm for 5-6 hours at 37°C. Cells were harvested by centrifugation for 15 minutes at 8500 g using a Sorvall RC6 floor-model centrifuge with SLA-3000 rotor at 4 °C and supernatant was discarded.

Purification of HIV-1 protease

Prior to protein purification, 250 mL of 9 M and 150 mL of 1 M urea buffers were freshly prepared in separate containers, and 1-2 g AG 501-X8 (D) resin (20-50 mesh) was added to each of the urea buffers. These solutions were placed on a heated (30 °C) stir plate and mixed by stir bar for approximately 2 hours to dissolve the urea, and the resin was removed by filtration. All buffer exchange steps described in the

following sections were carried out using a 5 mL HiTrap Desalting column from GE Healthcare (packed with Sephadex G25), which is first washed successively with 3 – 4 column volumes (15 – 20 mL) of nanopure water (nH₂O), 1 M NaCl, nH₂O, 0.5 M NaOH, nH₂O, and then equilibrated in the desired buffer.

Pelleted cells from 1 L growth were resuspended in 30 mL resuspension buffer, and 23.4 μ L of β -Mercaptoethanol (BME) was added to the reaction and mixed well by swirling. The approximate weight of the wet pellet was generally 5-6 grams, but varied from purification to purification. In order to lyse the cells and release the cellular contents, the sample was sonicated for 2 minutes using a Fisher brand tip sonicator at approximately 25 watts output power. Sonication was always performed in cycles of 5 seconds on followed by 5 seconds off to avoid shearing the proteins present in the lysate. The sample was then passed 3 times through a 35 mL French pressure cell (Thermo Scientific, Waltham, Massachusetts) operating at approximately 1200 pounds per square inch (psi).

Next, the lysed cells were centrifuged for 30 minutes at 18500 x g and 4 °C using the Eppendorf 5810R centrifuge with F34-6-38 rotor to collect cell debris and protease-containing inclusion bodies, and the supernatant was discarded. The inclusion body-containing pellet was resuspended, homogenized (using a 50 mL Dounce Tissue Homogenizer) and sonicated in 40 mL fresh wash buffer #1, then centrifuged for 30 minutes at 18500 x g and 4 °C, and the supernatant was discarded. This process was repeating with 40 mL wash buffer #2 and 40 mL wash buffer #3. Each of these steps functioned to isolate and wash the inclusion bodies, removing non-target proteins and remaining cellular components. In order to solubilize the inclusion bodies, the pellet

was then resuspended, homogenized and sonicated in 30 mL inclusion body resuspension buffer containing 9 M urea, then centrifuged at 18500 x g and 4 °C for 30 minutes. The supernatant, which contained solubilized target proteins in unfolded, monomeric form, was collected. Because purification proceeds by anion exchange chromatography, the pH of the inclusion body resuspension buffer needs must be adjusted according to the specific isoelectric point (pI) of the protein being purified. A 5 mL HiTrap Q HP Anion Exchange column was equilibrated with inclusion body resuspension buffer on an Akta Prime liquid chromatography system, and the supernatant containing solubilized HIV-1PR monomers was applied to the column at a rate of 5 mL/min. Fraction collection was started immediately and flow through was collected in 4 mL fractions, (specific fraction numbers depend on the results of the chromatogram). Fractions containing HIV-1PR were collected and pooled into a clean 50 mL Eppendorf tube (approximately 32 mL from 8 separate 4 mL fractions), then acidified to pH 5 (typically with approximately 39.3 µL formic acid) and stored for approximately 12 hours at 4 °C to allow some contaminants to precipitate, after which time the protease sample was decanted into a 50 mL polypropylene Eppendorf centrifuge tube and centrifuged for 30 minutes at 12000 rpm and 6 °C to separate precipitated contaminants. 300 mL 10 mM formic acid solution was prepared in a clean beaker and cooled on ice to approximately 0 °C. HIV-1PR was refolded by doing a 10-fold stepwise dilution on ice using a peristaltic pump for approximately 2 hours, and the pH was subsequently adjusted to approximately 3.8 (typically by adding about 1 mL 2.5 M sodium acetate). The solution temperature was brought to approximately 30 °C and the pH was adjusted to 5 (typically by adding about 3 mL 2.5 M sodium

acetate). After approximately 20 minutes of wait time, the solution was moved to balanced centrifuge tubes and centrifuged for 20 minutes at 18500 x g and 23 °C to remove contaminants that precipitated during refolding. The sample was concentrated to $OD_{280} = 0.5$ using an Amicon 100 mL equipped with a Millipore 10,000 Da MW cut-off polyethersulfone membrane. If further purification was required, the protease sample was buffer exchanged into 50 mM NaOAc, pH 5, and 5 mL of the concentrated protein sample was loaded on to a HiPrep 16/60 Sephacryl S-200 equilibrated with the same buffer and run at 0.5 mL/min and 2 mL fractions containing HIV-1PR were collected (fraction numbers vary according to chromatogram). The result is >95% pure HIV-1 Protease. No protease inhibitors were added to the lysate at any time because the target protein is a protease.

Incorporation of the unnatural amino acid

The plasmid pSUPAR3 (Figure 5-7) was co-transformed into BL21(DE3) with the expression plasmid containing the HIV-1PR gene and subsequently plated onto media containing ampicillin and chloramphenicol. The transformed cells were inoculated in 5 mL sterile Luria Bertani (LB) media and grown at 37 °C with shaking at 250 rpm to an optical density (OD_{600}) of approximately 0.60, then transferred to 1 L sterile LB media and grown to an OD_{600} of approximately 1.0 with shaking at approximately 200 RPM. Cells were then induced using IPTG and arabinose. The culture was then incubated with shaking at 250 rpm for 5-6 hours at 37°C. Cells were harvested by centrifugation for 15 minutes at 8500 g using a Sorvall RC6 floor-model centrifuge with SLA-3000 rotor at 4 °C and supernatant was discarded. Purification was carried out as described above. Protein containing *p*-acetylphenylalanine was subsequently reacted with a ketone-specific reagent (HO-4120) to produce ketoxime-linked spin label side chain.

DEER experiments

Pulsed EPR data were collected on a Brüker EleXsys E580 EPR spectrometer (Billerica, MA) equipped with the ER 4118X-MD-5 dielectric ring resonator at a temperature of 65 K (cooled via liquid helium) and 4-pulse DEER sequence, described in Chapter 2. DEER data collection was typically preceded by a series of preliminary experiments, also described in Chapter 2, designed to accurately determine the center field, T_m , the d_0 (time in ns at which the echo begins) and pulse gate (the breadth of the echo in ns), and appropriate positions for the observer sequence and the pump pulse. The “observer sequence” is applied at the low field resonance approximately 26 Gauss below the “pump pulse” at central resonance.

DEER data analysis

Data analysis was carried out as described previously by Blackburn et al. (Galiano, Blackburn et al. 2009). The raw experimental data was processed via Tikhonov Regularization using DeerAnalysis2008 software, available at <http://www.epr.ethz.ch/software/index>, and then the distance profiles were reconstructed with a series of Gaussian functions via an in-house Matlab based program called DeerSim.

Population validation

Population validation was performed as discussed previously by Blackburn et al. (Galiano, Blackburn et al. 2009). The validity of each population with a relative percentage of 10% or below was tested via population suppression (discussed in detail in Chapter 2) using DeerSim and DeerAnalysis2008.

Results and Discussion

Effect of Sample Freezing Conditions on the Conformational Ensemble of HIV-1PR

Double electron-electron resonance (DEER) is a pulsed electron paramagnetic resonance (EPR) spectroscopy technique also referred to as pulsed electron double resonance (PELDOR). Used in conjunction with site-directed spin-labeling (SDSL), DEER has been utilized to examine the conformational ensembles of the flaps in HIV-1PR under various conditions (Galiano, Bonora et al. 2007; Blackburn, Veloro et al. 2009; Galiano, Blackburn et al. 2009; Galiano, Ding et al. 2009; Kear, Blackburn et al. 2009). Distance measurements by SDSL DEER are based on the strength of the dipolar coupling of the unpaired spins (Pannier, Veit et al. 2000; Jeschke, Chechik et al. 2006), and can yield distances of 20-60 Å between spin labels. Results have provided detailed information regarding flap conformations sampled by the protease, with conformers designated as wide-open, semi-open, closed, and tucked or curled. These states were previously modeled using molecular dynamic simulations (Ding, Layten et al. 2008).

DEER experiments are typically run at cryogenic temperatures; particularly, ours are collected at 65 K. Prior to inserting the sample into the temperature-controlled cryostat, the sample must be frozen to a glass state. Most commonly, samples are frozen by submerging the sample tube in liquid nitrogen (LN₂) until a glass state is achieved. Work here was designed to determine what, if any, effects are created by the method of freezing and speed at which sample freezing takes place. Four different freezing techniques were analyzed, all four using LN₂ as the cooling agent. A room temperature sample was frozen by direct immersion for approximately 30 seconds in

LN₂. A room temperature sample was first frozen at -20 °C, followed by immersion in LN₂. A more rapid freezing method than those previously described can be achieved by using isopentane cooled by immersion into liquid nitrogen as the medium for freezing. Isopentane is desirable for this purpose because it is a liquid at room temperature, it is less volatile with a higher boiling point than LN₂, and it conducts heat very well. Accordingly, both a room temperature sample and a sample equilibrated to biological temperature (37 °C) were immersed in isopentane suspended in LN₂ as the method of freezing the sample prior to data collection. For all samples, experimental dipolar modulated echo curves were analyzed via Tikhonov regularization (TKR) with DeerAnalysis2008, (Jeschke, Chechik et al. 2006) and populations assigned to flap conformations of curled/tucked, closed, semi-open, and wide-open were obtained.

In this case, we used a subtype C construct modified with two EPR-active spin labels incorporated into the flaps at the aqueous exposed sites K55C and K55C' (Figure 5-1). The details of the subtype C construct are given in Figure 5-8. The conformational ensembles were determined for the apo protease and the protease bound to indinavir, tipranavir, and the non-hydrolyzable substrate mimic called CA-p2. The conformational ensemble for apo subtype C HIV-1PR is reported in the literature, with the average distance and relative percentage for each population as 29.7 Å and 13%, 33.3 Å and 8%, 36.7 Å and 52%, and 40.2 Å and 27% and assigned to conformations where the flaps are tucked/curled, closed, semi-open, and wide open, respectively (Kear, Blackburn et al. 2009); to date, no conformational ensemble is reported for subtype C with CA-p2, IDV, or TPV.

The distance parameters for HIV-1PR subtype C (apo) under various freezing conditions are shown in Table 5-1 and Figures 5-9 – 5-13; subtype C (IDV) are shown in Table 5-2 and Figures 5-14 – 5-18; subtype C (TPV) are shown in Table 5-3 and Figures 5-19 – 5-23; subtype C (CaP2) are shown in Table 5-4 and Figures 5-24 – 5-28.

None of the Subtype C constructs (apo, IDV, TPV, or CaP2) showed drastic changes in the determined conformational ensemble when different freezing methods were employed. The most variation, though still not dramatic, was found to occur when the protease was bound to IDV. This result is rational because the binding of IDV is known to be an entropy-driven process. Likely, the fastest freezing method is the most accurate, but it's suggested that none of the freezing conditions examined here capture a conformational ensemble that is an inaccurate description of the true conformational ensemble.

This work is significant because it addresses a concern previously expressed about the DEER method. The pulsed EPR data are collected at cryogenic temperatures because, at room temperature, the spin memory time is too short for detection of the spin echo. Often, the question of whether the conformations trapped at 40-80 K accurately represent the thermodynamic conformations sampled.

Single Point Drug-Induced Mutations Confer Alterations in the Conformational Ensemble of HIV-1 Protease

Figure 5-29 shows the background subtracted dipolar modulated echo curves (Fig 5-29A) and corresponding distance profiles (Figure 5-29B) for the four constructs. Data were analyzed via Tikhonov regularization with DeerAnalysis2008 (Jeschke, Timmel et al., 2006). The dotted lines in Figure 5-29B indicate the distances consistent

with the closed (33 Å) and semi-open (36 Å) populations (Blackburn, Veloro, 2009; Kear, Blackburn et al., 2009). It is clear that for these data, PRE, L63P and A71V have a most probable distance of about 33 Å, which corresponds with the closed conformation, whereas POST has a most probable distance of 37 Å, similar to Subtype B(si), which is consistent with a predominant semi-open conformational. Table 5-5 summarizes the results of Figure 5-29B. We have shown that a detailed analysis of the TKR distance profiles is possible via a Gaussian reconstruction analysis, leading to population ensembles for flap conformations (Blackburn, Veloro, 2009; Kear, Blackburn et al., 2009). Figure 5-29C shows the results of the population analysis of L63P HIV-1PR conformational sampling based upon a Gaussian reconstruction of the data in Figure 5-29B. These analyses provide details of the relative populations in the flap conformational ensemble comprised of states assigned to flap conformations of the tucked/curled, closed, semi-open, and wide-open states. Within the signal-to-noise of our experimental data, populations as low as 5% may be extracted with low error (Blackburn, Veloro, 2009; Kear, Blackburn et al., 2009). Based upon characterization of subtype B HIV-1PR (Galiano, Ding, 2009; Blackburn, Veloro, 2009; Kear, Blackburn et al., 2009; Torbeev, Raghuraman, 2009; Ding, and Simmerling, 2008; Hornak, Rizzo et al., 2006; Scott, and Schiffer, 2000; Heaslet, Rosenfeld et al. 2007; Foulkes-Murzycki, Scott et al., 2007), the average distances for the preceding populations are 24-28, 33, 36, and 42 Å, respectively.

Shown in Figure 5-30A are the relative populations of each distinct flap conformation that, in sum, make up the conformational ensembles of each of the HIV-1 PR apo-constructs, which correspond to the structures shown in Figure 5-30B. It is

noteworthy that although the distance between the MTSL spin-labels in the tucked/curled conformation are shorter than those in the closed-state, the experimentally obtained distances are consistent with structures seen in previous MD work, where the flaps are turned in towards the active site cavity and are described as a trigger for the wide-open states and where an opening to the active site pocket can be seen. Individual complete datasets from each of the four samples are shown in Figures 5-31 – 5-34. Figure 3-35 summarizes and compares the relative populations (%) for closed, semi-open, tucked/curled, and wide-open conformations of PRE apo (blue) and POST apo (red). Figures 3-36 – 3-39 show the mass spectra of spin-labeled constructs showing major peaks with charge states indicated. Figure 3-40 summarizes the analysis of the mass spectrometry data including theoretical and observed molecular weights.

The conformational ensemble of PRE, giving a predominantly closed conformation, is unexpected because, up to now, a ligand-free construct with a most probable distance corresponding with the closed flap conformation has not been reported. X-ray crystal structures of ligand-free HIV-1PR primarily exhibit the semi-open conformation (Ishima, Freedberg et al., 1999; Freedberg, Ishima et al., 2002; Todd, Semo et al., 2002; Todd, Freire et al., 1999; Todd, Luque et al., 2000; Hornak, Okur et al., 2006), and previous DEER studies have shown subtypes B, C, CRF01_A/E, F and multi-drug resistant constructs V6 and MDR769 exhibit predominantly semi-open flap state (Kear, Blackburn et al., 2009). Inhibitor-bound HIV-1PR has been shown to primarily adopt the closed conformation, caused at least in part by interaction between the Ile50/50' amide group and ligand via a conserved H₂O (Galiano, Bonora et al.,

2007; Galiano, Ding et al., 2009; Blackburn, Veloro, 2009; Kear, Blackburn et al., 2009; Ding, Layten et al., 2008; Rose, Craik, 1998). Intraflap hydrogen bonding, which may stabilize the closed flap conformation, has been observed in MD simulations while also demonstrating that the average flap distance in the closed state is similar for ligand-bound and unbound protease (Ding, Layten et al., 2008), indicating that despite having no flap-inhibitor interactions, ligand-free protease may adopt a closed conformation.

The conformational ensemble for Subtype B, reported in a previous work by Kear et al., consists of a most probable distance of 35.2 Å that corresponds with the semi-open conformation. The addition of the single mutation A71V, and to a slightly lesser extent L63P, shifts the conformational ensemble to favor the closed conformation with a most probable distance to 33 Å. This result is significant because it provides physical evidence that a single amino acid mutation is sufficient to cause a substantial shift in the flap conformations of HIV-1PR, and also because it suggests that it's not the lack of stabilizing mutations causing the staggering difference between the PRE and Subtype B_s (subscript "s" indicates the presence of three stabilizing mutations, namely Q7K, L33I, and L63I; subscript "i" indicates that the protease contains the inactivating D25N mutation) ensembles.

The PRE construct differs from Subtype B-LAI sequence by four polymorphisms, likely the result of natural genetic drift in the mother of the infected child from which the construct was derived. The POST construct differs from the PRE construct through an additional 8 drug-selected polymorphisms, including L10I, I15V, E34Q (negative to uncharged), M36I, T37N, I54A, R57E (positive to negative), and V82A. The combination of these mutations induces a substantial shift of the conformational

ensemble of the protease, from a primary conformation of closed to semi-open. M36I, located in the hydrophobic core, is thought to contribute to drug resistance and likely plays an important role in the conformational switch via the hydrophobic sliding mechanism (Foulkes-Murzycki, Scott et al., 2007).

The DEER results reported in this work show that drug-induced mutations, even outside of the active site pocket, can significantly affect the conformational ensemble of HIV-1PR, which likely play an important role in drug-resistance and efficacy. It is worthwhile to note here that the first generation of PIs used in treatment of HIV-1, including Saquinavir, Indinavir, and Ritonavir, were designed with respect to subtype B, but that subtype B virus constitutes only a small percentage of HIV worldwide. A second generation of PIs addressed the occurrence of resistant variants. Darunavir is the latest drug approved and is most effective against the highly drug resistant forms. The data reported here illustrates the importance of these later generation PIs, and also supports the need for ongoing drug-development focused on efficacy against common polymorphisms and prominent subtypes.

Electron Paramagnetic Resonance Studies of HIV-1 Protease Via Incorporation of the Unnatural Amino Acid Acetyl Phenylalanine

We have shown previously, via traditional SDSL-EPR approaches, that the FDA-approved HIV-1PR inhibitors induce changes in the conformational sampling of the flap regions of the protease. However, traditional approaches require removal of the native CYS residues. A non-native amino acid spin label, *p*-acetyl-L-phenylalanine, was received from Mark Fleissner and Wayne Hubbel at UCLA, who demonstrated its utility using T4 lysozyme (Jeschke, Polyhach et al., 2000). By incorporating the *p*-acetyl-L-phenylalanine, we can compare the distance profiles obtained for protease with our

original DEER results (with MTSL label). As mentioned in the introduction section, HIV-1PR was successfully expressed with p-acetyl-L-phenylalanine and subsequently spin labeled to yeast HIV-1PR K55K1. Preliminary DEER experiments (Figure 5-41) achieved low levels of SNR, and parameters are currently require optimization.

Table 5-1. Distance parameters for HIV-1PR C (apo) and various freezing conditions.

Freezing method	Tucked/curled	Closed	Semi-open	Wide-open
Isopentane/liquid N ₂ from 37 °C				
R(Å)	29.7	33.3	36.7	40.2
FWHM	3.1	2.8	4.0	3.3
% Conformation	6	12	55	27
Isopentane/liquid N ₂ from 25 °C				
R(Å)	29.7	33.3	36.7	40.2
FWHM	3.1	2.8	4.0	3.3
% Conformation	6	13	54	27
Liquid N ₂ from 25 °C				
R(Å)	29.7	33.3	36.7	40.2
FWHM	3.1	2.8	4.0	3.3
% Conformation	11	8	54	27
-20 °C from 25 °C				
R(Å)	27.2	33.8	37.1	40.0
FWHM	4.4	5.0	4.9	3.7
% Conformation	11	8	54	27

Table 5-2. Distance parameters for HIV-1PR C (IDV) and various freezing conditions.

Freezing method	Tucked/curled	Closed	Semi-open	Wide-open
Isopentane/liquid N ₂ from 37 °C				
R(Å)		32.5	37.9	40.0
FWHM		5.0	4.9	4.0
% Conformation		30	50	20
Isopentane/liquid N ₂ from 25 °C				
R(Å)	26.2	33.1	36.3	39.3
FWHM	4.8	4.8	3.7	3.7
% Conformation	6	27	50	17
Liquid N ₂ from 25 °C				
R(Å)		33.2	36.9	40.0
FWHM		4.6	4.8	4.0
% Conformation		30	50	20
-20 °C from 25 °C				
R(Å)	27.2	33.8	37.1	40.0
FWHM	4.4	5.0	4.9	3.7
% Conformation	6	27	50	17

Table 5-3. Distance parameters for HIV-1PR (TPV) and various freezing conditions.

Freezing method	Tucked/curled	Closed	Semi-open	Wide-open
Isopentane/liquid N ₂ from 37 °C				
R(Å)		33.0	37.5	41.0
FWHM		2.8	3.3	1.4
% Conformation		86	12	2
Isopentane/liquid N ₂ from 25 °C				
R(Å)		33.0	36.8	41.5
FWHM		2.7	2.6	1.7
% Conformation		84	12	4
Liquid N ₂ from 25 °C				
R(Å)		32.7	34.9	39.0
FWHM		3.1	3.3	3.0
% Conformation		83	12	5
-20 °C from 25 °C				
R(Å)		32.9	37.1	40.0
FWHM		2.8	4.9	3.7
% Conformation		82	12	6

Table 5-4. Distance parameters for HIV-1PR (CaP2) and various freezing conditions.

Freezing method	Tucked/curled	Closed	Semi-open	Wide-open
Isopentane/liquid N ₂ from 37 °C				
R(Å)	32.3	34.1		
FWHM	2.7	4.0		
% Conformation	10	90		
Isopentane/liquid N ₂ from 25 °C				
R(Å)	31.9	33.9		
FWHM	1.9	3.1		
% Conformation	12	88		
Liquid N ₂ from 25 °C				
R(Å)		33.2	36.9	40.0
FWHM		4.0	4.8	4.0
% Conformation		30	50	20
-20 °C from 25 °C				
R(Å)	27.2	33.9	37.1	40.0
FWHM	4.3	5.0	4.9	3.7
% Conformation	6	27	50	17

Table 5-5. Summary of distance parameters obtained from DEER EPR of HIV-1PR.

Construct	Range (span)	Most probable distance	Average distance
Subtype B*	24-45 Å (21)	35.2 Å	35.2 Å
PRE	25-45 Å (20)	33.2 Å	32.8 Å
POST	29-41 Å (12)	37.0 Å	36.7 Å
L63P	31-43 Å (12)	33.2 Å	34.4 Å
A71V	31-43 Å (12)	33.0 Å	33.2 Å

*Data from Kear, et al.

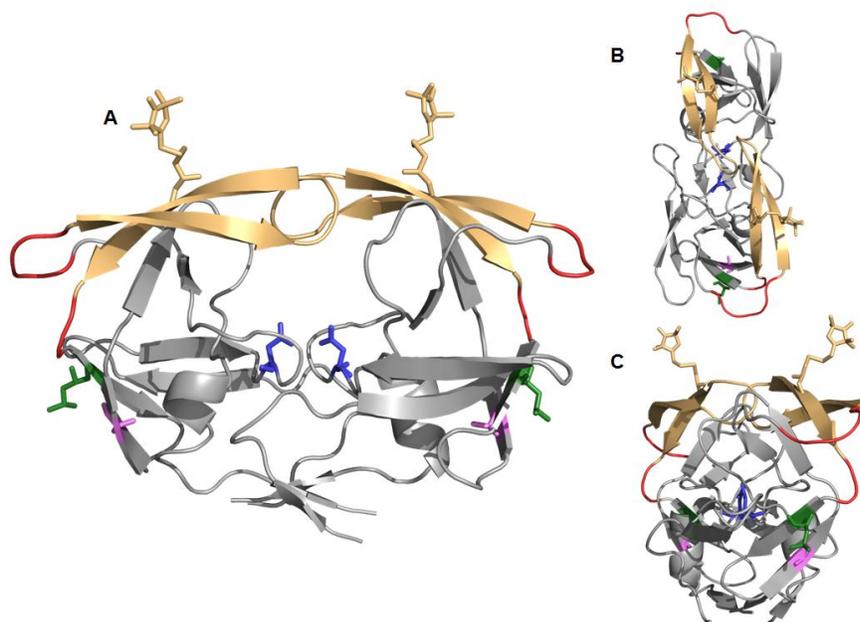


Figure 5-1. HIV crystal structure. A) Front view ribbon diagram of HIV-1PR (PDB ID 2PBX) with “flaps” and K55R1 reporter sites shown in tan, active site cavity and catalytic D25 residues in blue, the hinge regions in red, and L63 and A71 residues as green and pink, respectively. (B) Top View. (C) Side view.

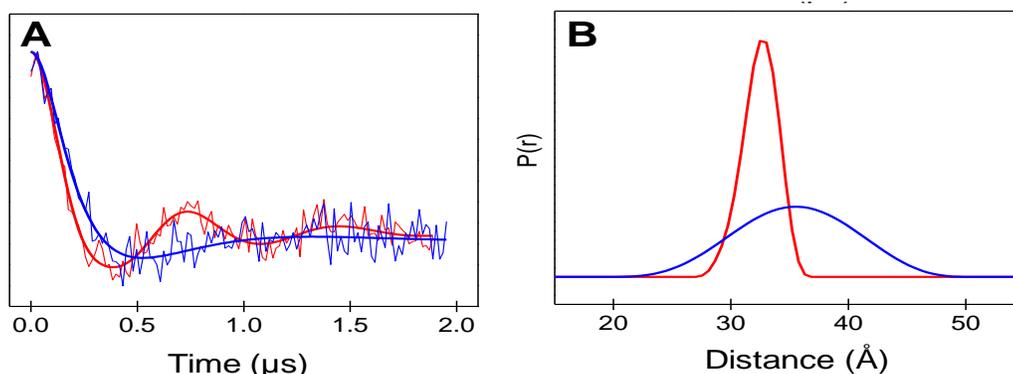


Figure 5-2. DEER results of subtype B HIV-1Pr with (blue) and without (red) the FDA-approved inhibitor Ritonavir. (A) Dipolar modulated echo data and (B) resulting distance distribution profile.

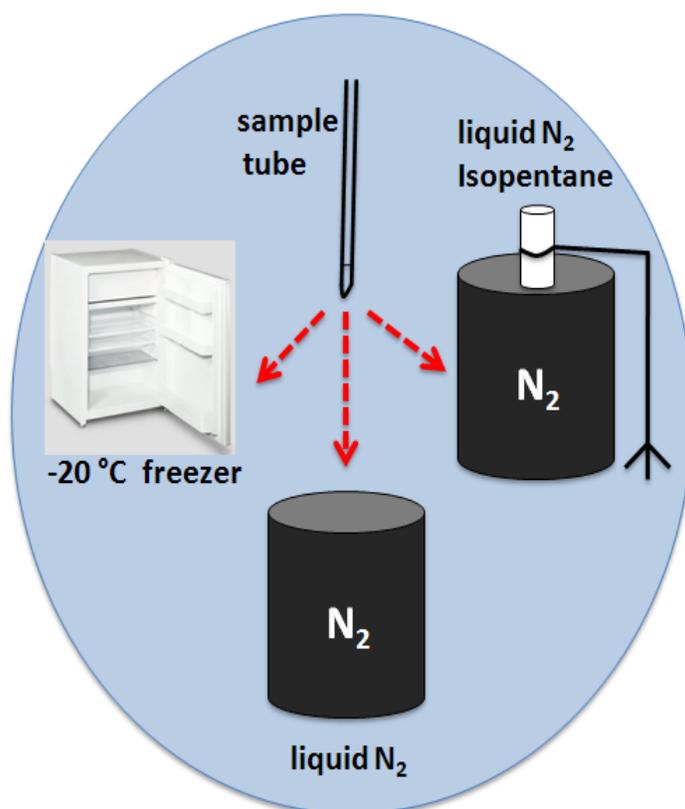


Figure 5-3. DEER experiments are typically run at cryogenic temperatures, ours at 65 K. Prior to inserting the sample into the temperature controlled cryostat, the sample must be frozen in a glass state. This work was designed to determine what, if any, effects are created by the method of freezing and speed at which sample freezing takes place. Four freezing techniques were considered: 1. placing the 25 °C sample tube at -20 °C freezer, 2. submerging the 25 °C sample in liquid N₂, 3. dipping the 25 °C sample tube in isopentane suspended in liquid N₂, and 4. dipping the 37 °C sample tube in isopentane suspended in liquid N₂.

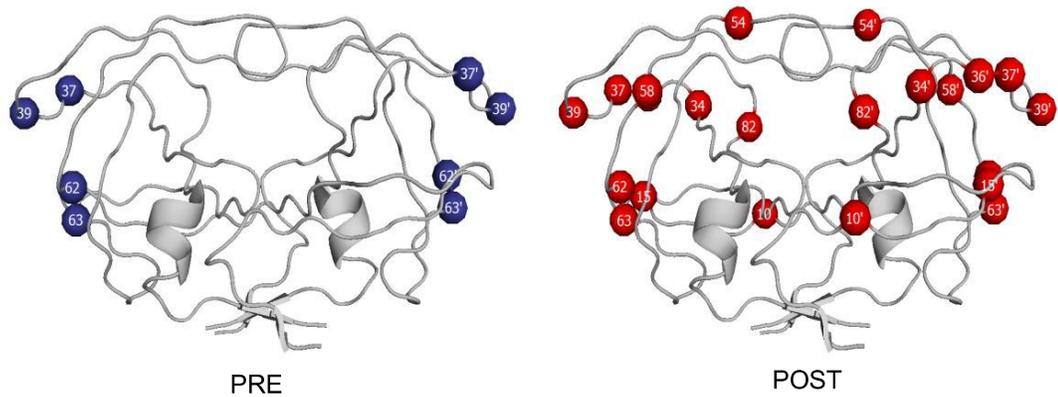


Figure 5-4. Locations of polymorphisms in PRE and POST, with respect to Subtype B (LAI).

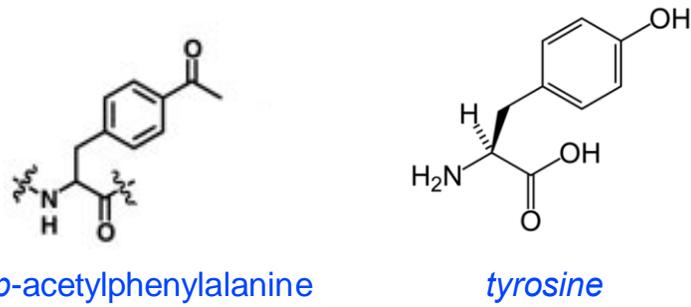


Figure 5-5. Unnatural amino acid *p*-acetylphenylalanine and structurally similar amino acid tyrosine.

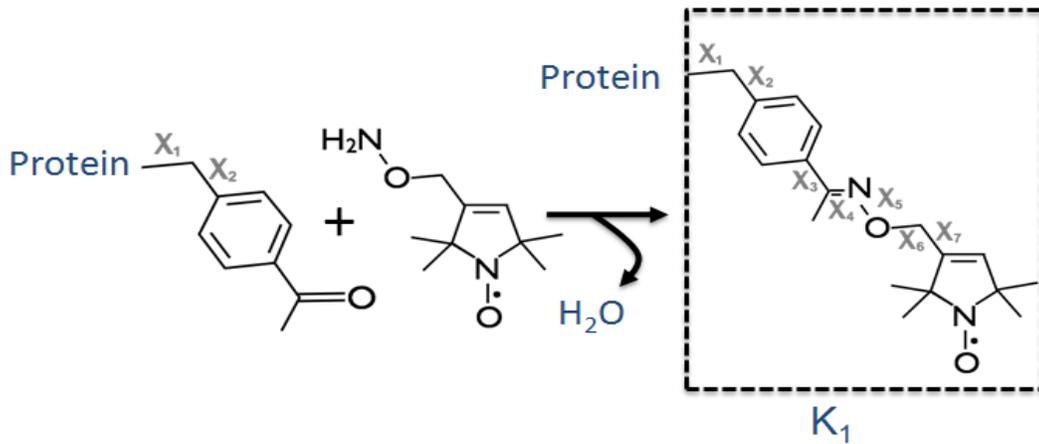


Figure 5-6. Chemical reaction for Incorporation of unnatural amino acid to form spin-labeled side-chain K1.

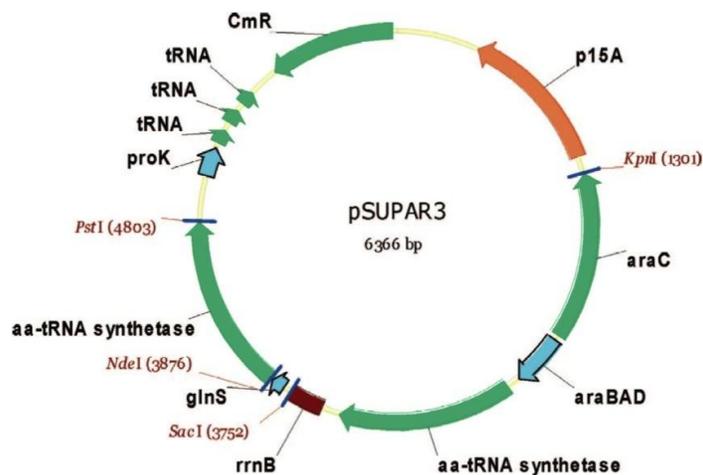
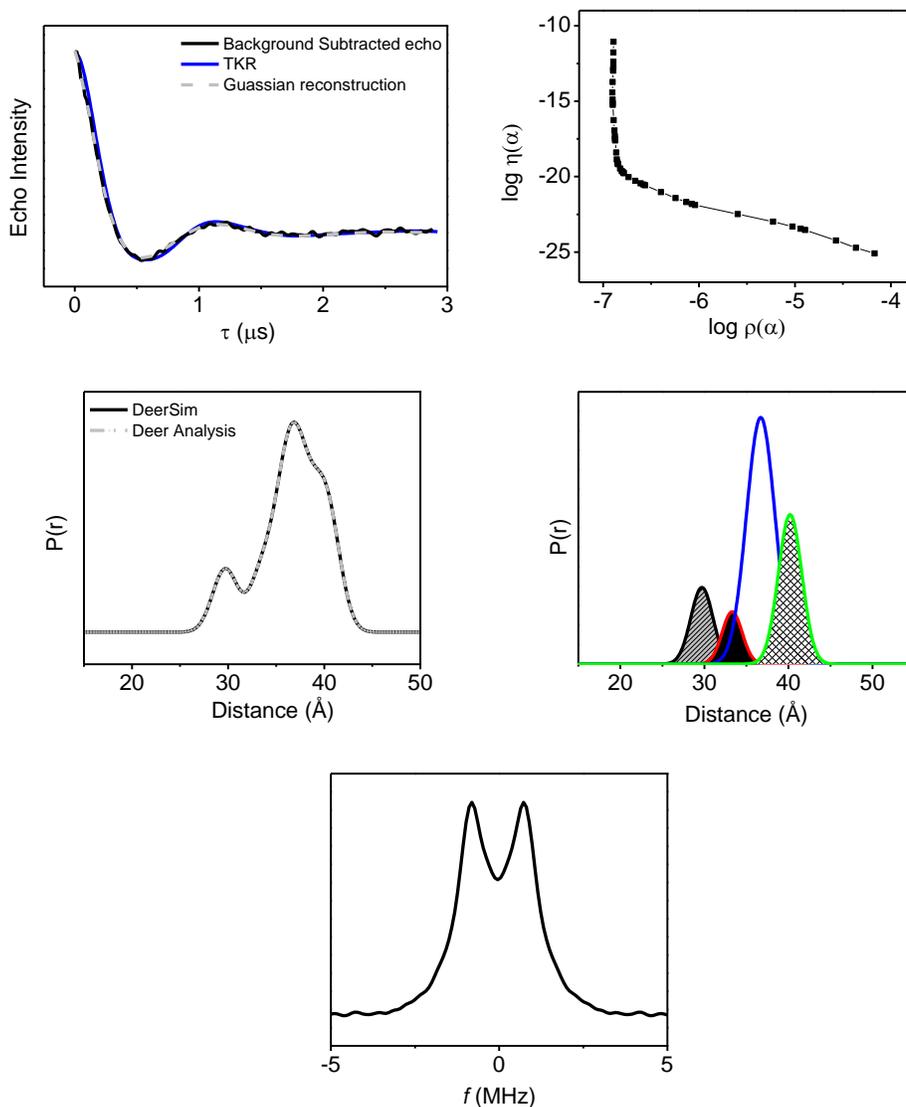


Figure 5-7. pSUPAR3 plasmid map, used for co-expression of *p*-acetylphenylalanine, which is subsequently reacted with a ketone-specific reagent (HO-4120) to produce ketoxime-linked spin label side chain.

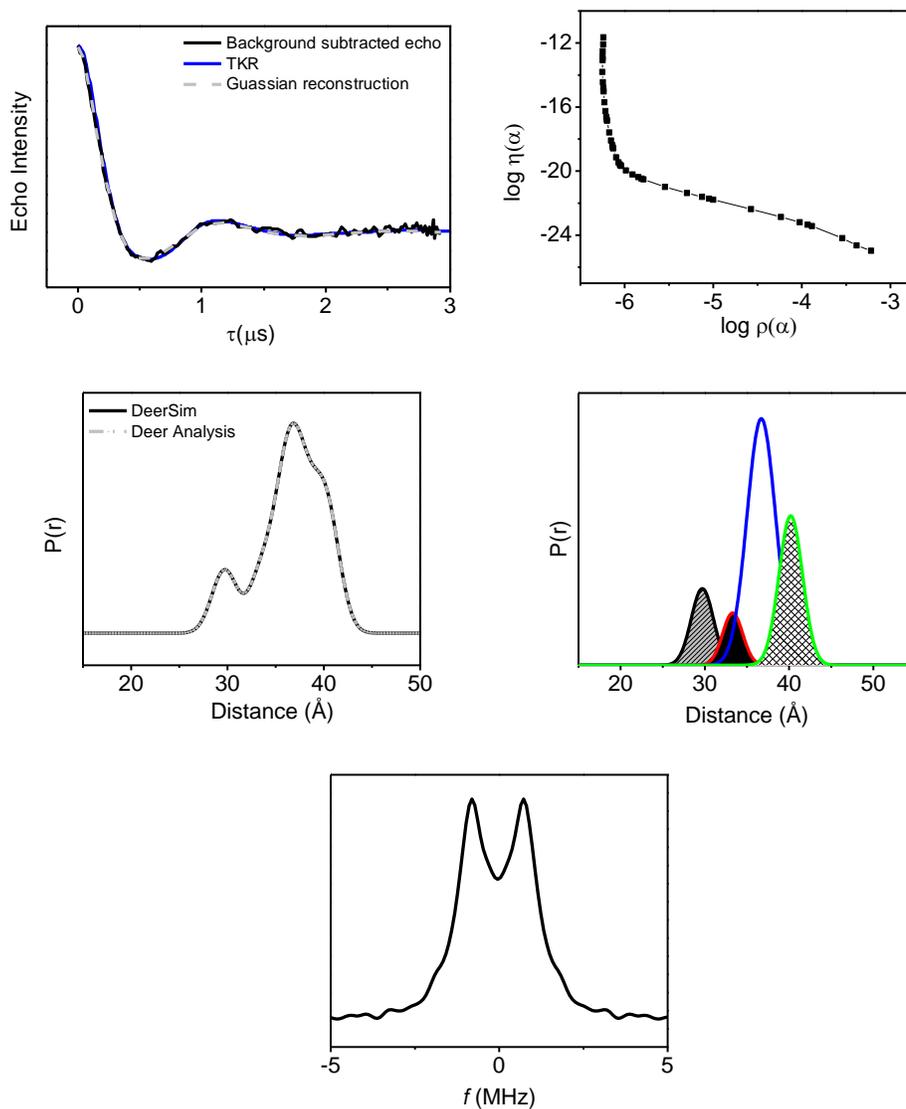
Construct	Sequence				
	Position 10	20	30	40	50
LAI	PQITLWQRPL	VTIKIGGQLK	EALLDTGADD	TVLEEMSLPG	RWKPKMIGGI
B	-----K-----	-----N-----	-----I-----	-----	-----
PRE	-----	-----N-----	-----T-T-----	-----	-----
L63P	-----K-----	-----N-----	-----I-----	-----	-----
A71V	-----K-----	-----N-----	-----I-----	-----	-----
POST	-----I-----	-----V-----	-----N-----	-----Q-IN-T-----	-----
Csi	-----K-----	-S-----I-----	-----N-----	-----I-----	-----IA-----
	Position 60	70	80	90	99
LAI	GGFIKVRQYD	QILIEICGHK	AIGTVLVGPT	PVNIIGRNLL	TQIGCTLNF
B	-----C-----	-----I-----A-----	-----	-----	-----A-----
PRE	-----C-----	-----VP-----A-----	-----	-----	-----A-----
L63P	-----C-----	-----P-----A-----	-----	-----	-----A-----
A71V	-----C-----	-----I-----A-----	V-----	-----	-----A-----
POST	-----AC-E-----	-----VP-----A-----	-----	-----A-----	-----A-----
Csi	-----C-----	-----I-----A-----K-----	-----	-----M-----	-----L-A-----

Figure 5-8. Sequence alignment for HIV-1 protease LAI with constructs B_{si}, C_{si}, PRE_i, I63P_{si}, A71V_{si}, and POST_i. Subscript “s” denotes the presence of three stabilizing mutations, which protect from autoproteolysis. Subscript “i” refers to inactive protease that contains the D25N mutation.



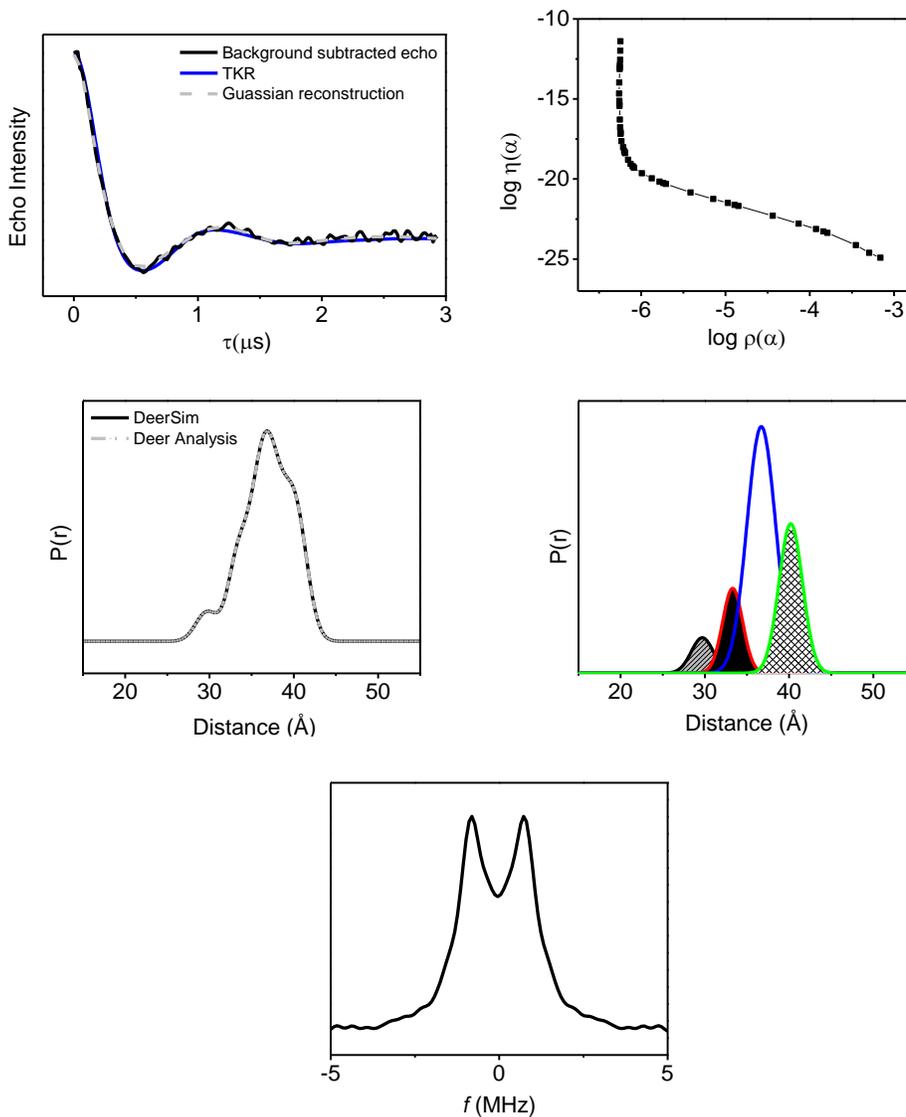
Csi apo -20°C	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	29.7	3.1	11
Closed	33.3	2.8	8
Semi-open	36.7	4.0	54
Wide-open	40.2	3.3	27

Figure 5-9. HIV-1PR Subtype C apo frozen at -20. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum; Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR.



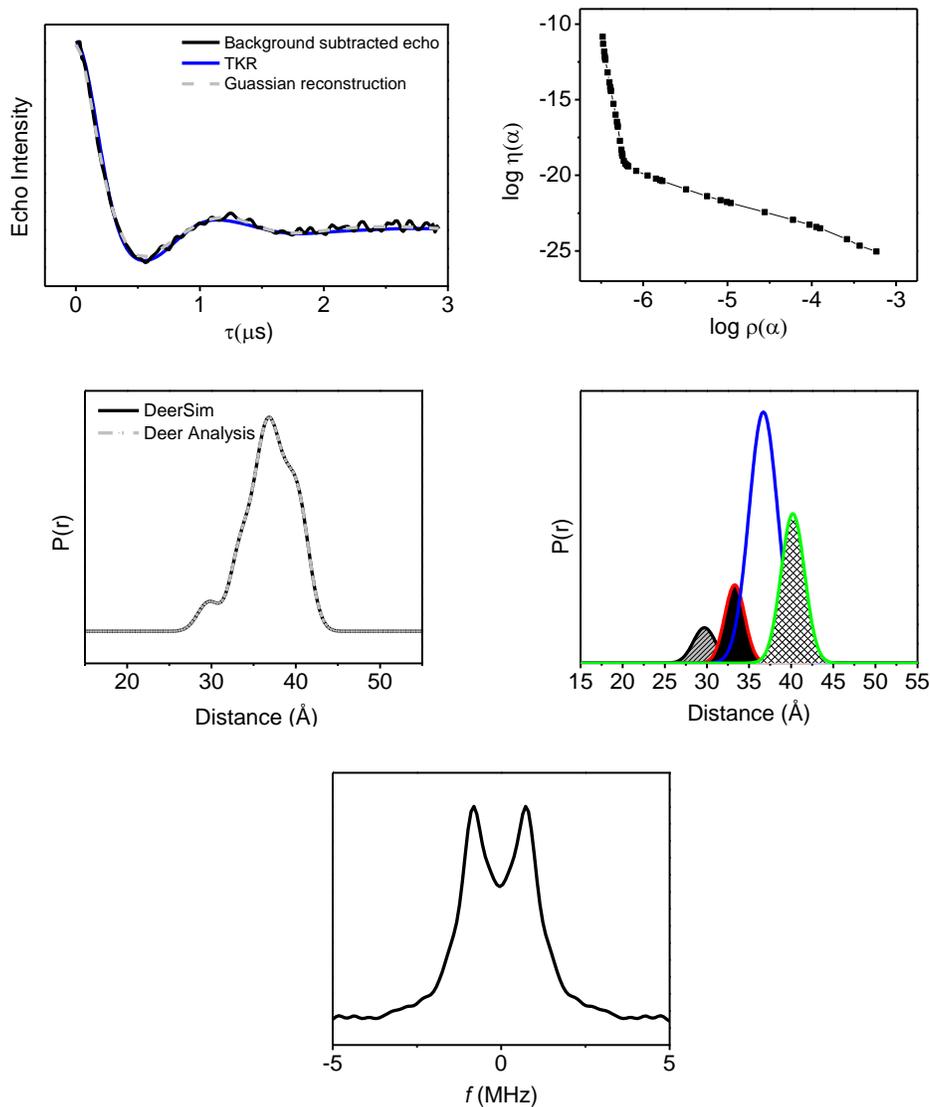
Csi apo N ₂	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	29.7	3.1	11
Closed	33.3	2.8	8
Semi-open	36.7	4.0	54
Wide-open	40.2	3.3	27

Figure 5-10. HIV-1PR Subtype C apo frozen in liquid nitrogen. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi apo IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	29.7	3.1	6
Closed	33.3	2.8	13
Semi-open	36.7	4.0	54
Wide-open	40.2	3.3	27

Figure 5-11. HIV-1PR Subtype C apo frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi apo 37°C IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	29.7	3.1	6
Closed	33.3	2.8	12
Semi-open	36.7	4.0	55
Wide-open	40.2	3.3	27

Figure 5-12. HIV-1PR Subtype C apo frozen in isopentane from 37C. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve (blue) and reconstructed echo curve (gray) corresponding to the sum of distance profiles; L curve; Distance profile from TKR analysis (black) overlain with summed Gaussian populations (gray dashed); Individual populations for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum; Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR.

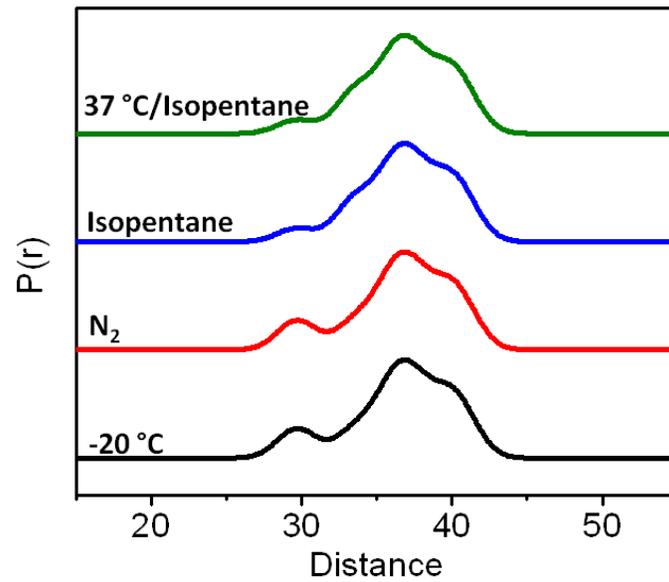
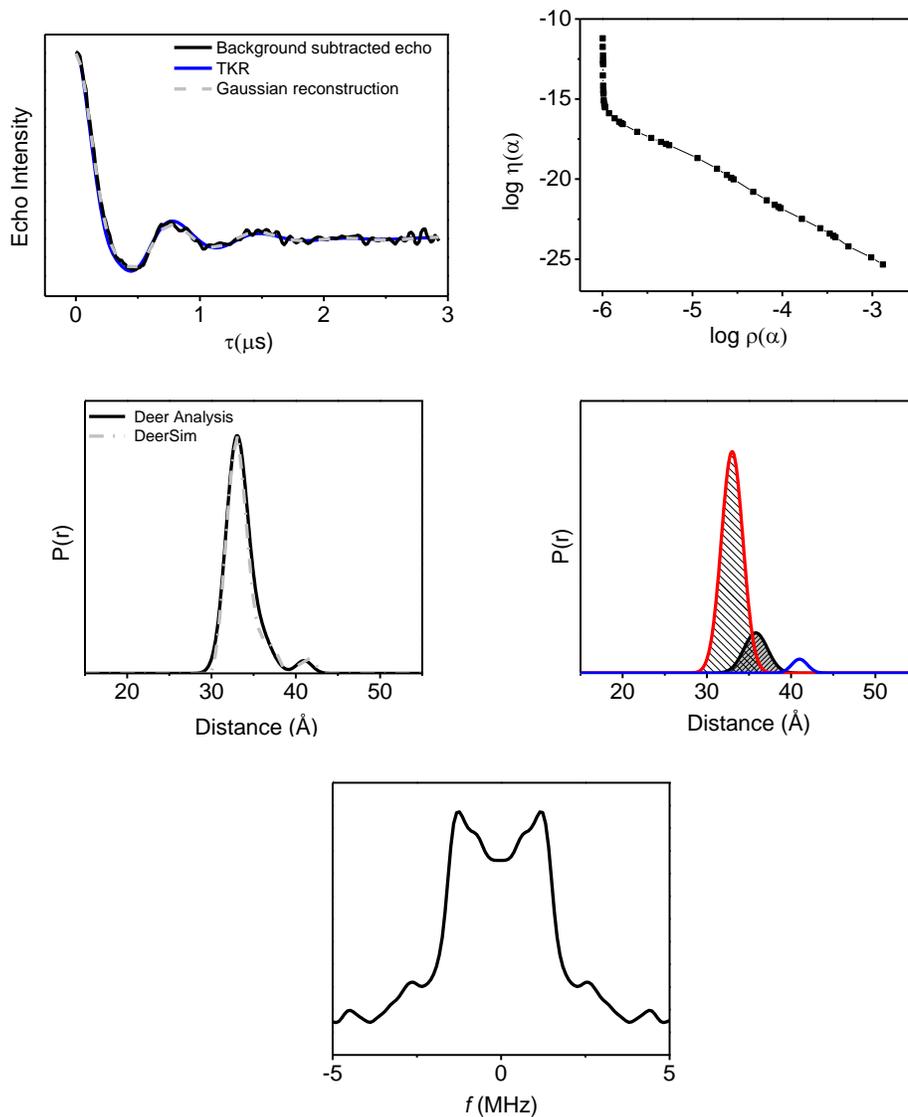
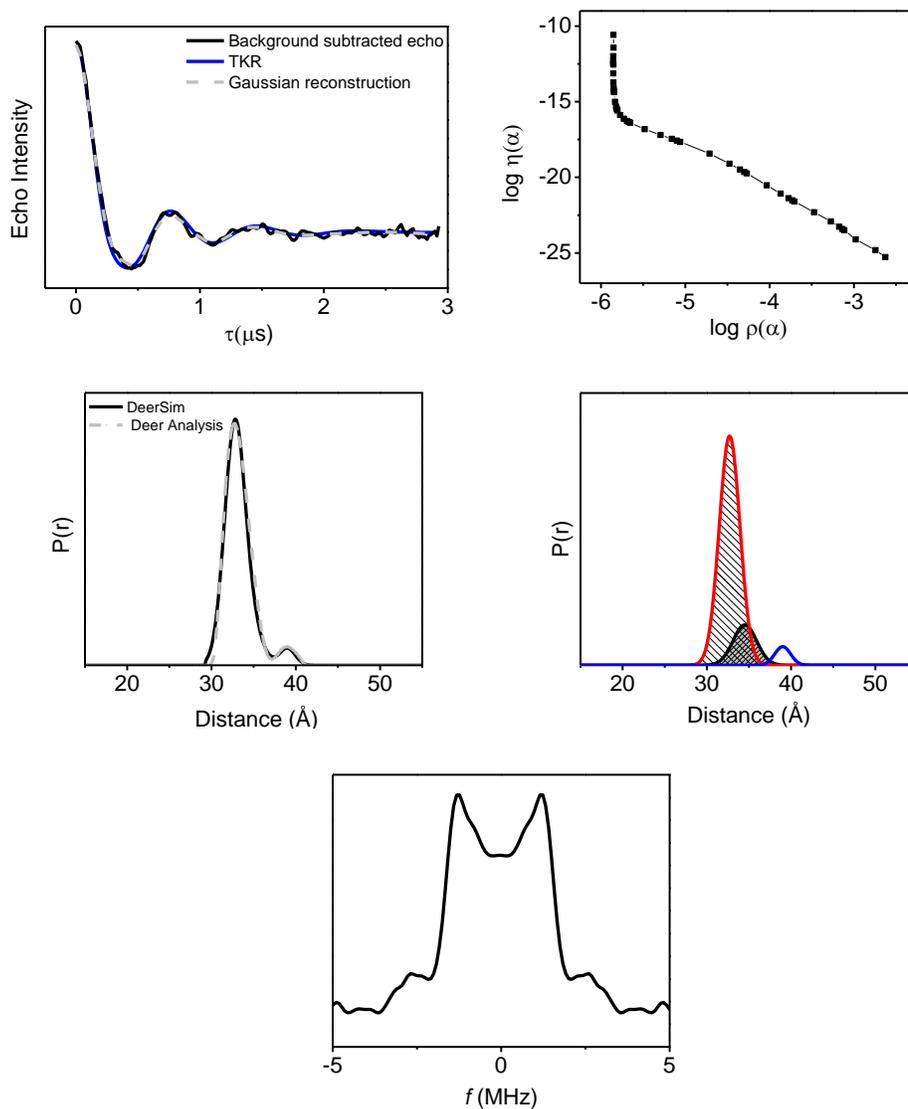


Figure 5-13. Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, apo after various freezing conditions, including -20 C, liquid nitrogen, isopentane/liquid nitrogen from room temperature, and isopentane/liquid nitrogen after heating to 37 °C.



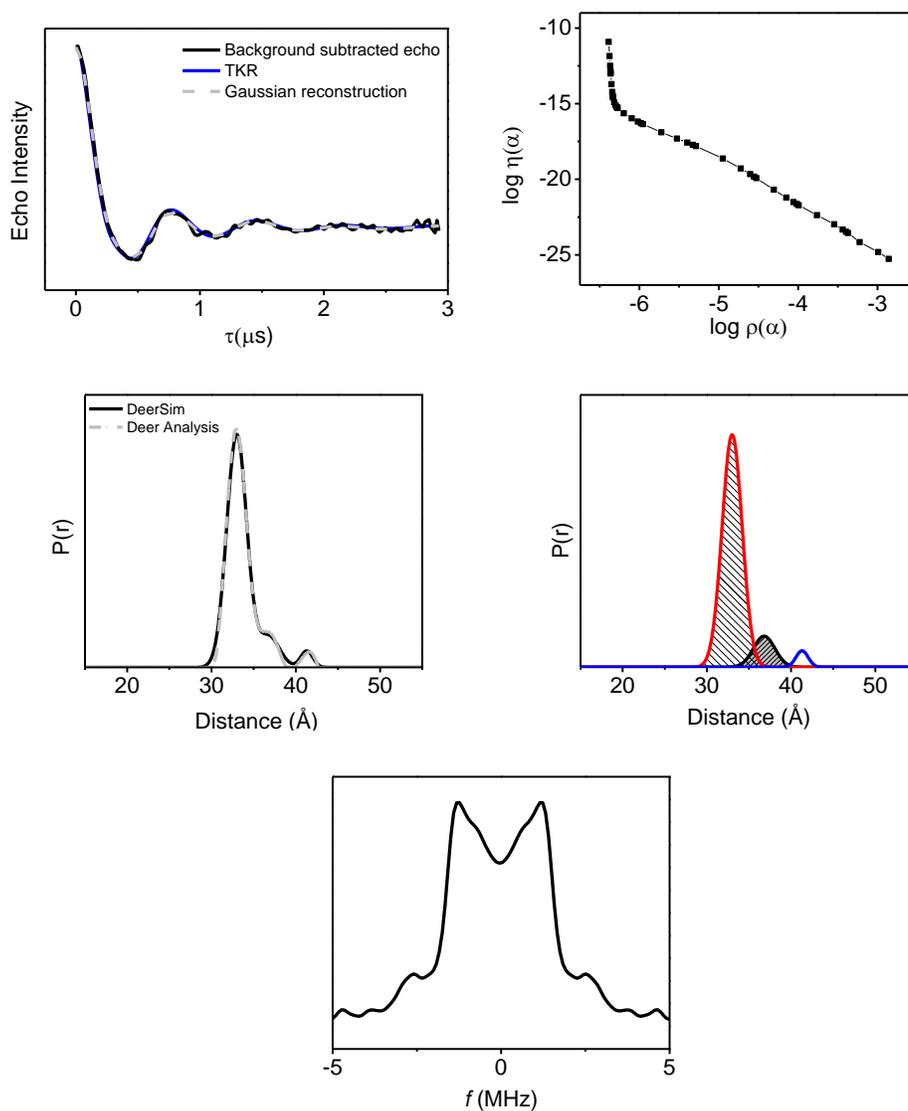
Csi TPV -20°C	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Closed	32.9	2.8	82
Semi-open	36.8	2.8	12
Wide-open	41.1	2.0	6

Figure 5-14. HIV-1PR Subtype C TPV-bound frozen at -20 °C. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



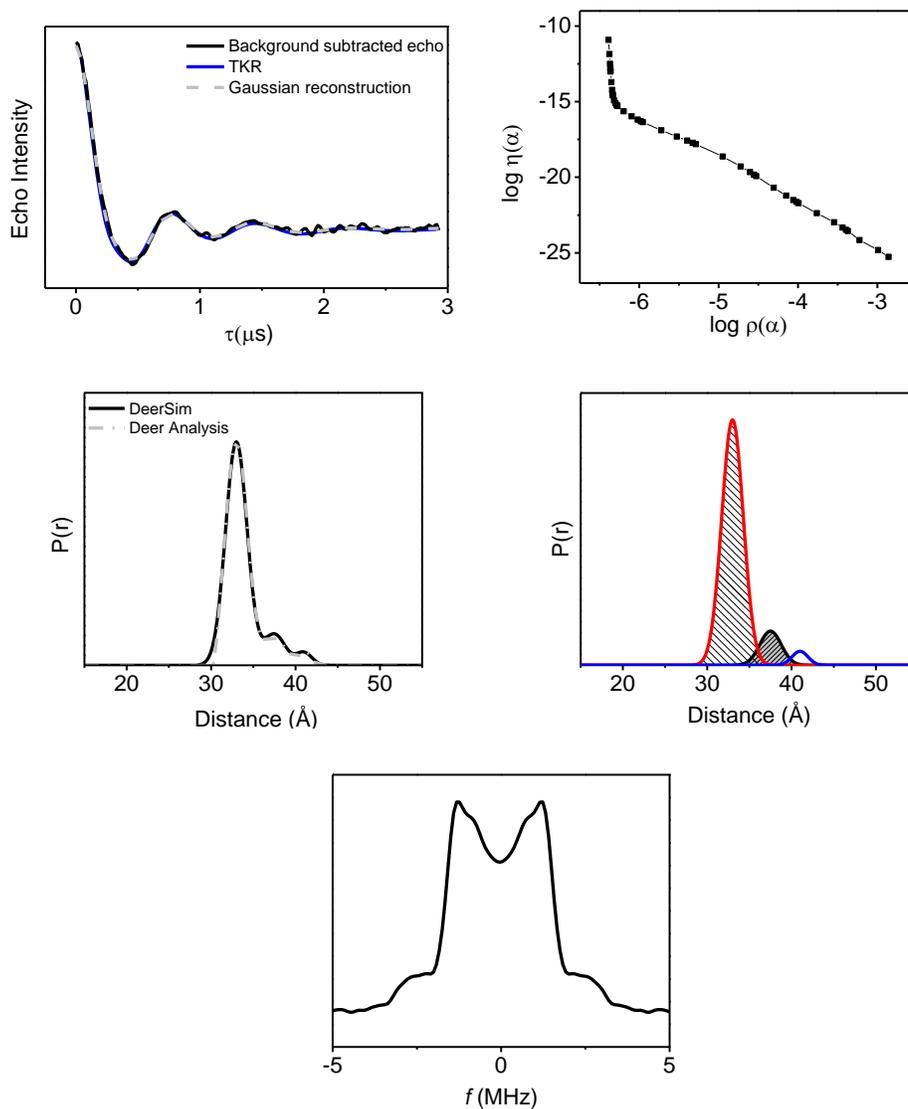
Csi TPV N ₂	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Closed	32.7	3.1	83
Semi-open	34.9	3.3	12
Wide-open	39.0	3.0	5

Figure 5-15. HIV-1PR Subtype C TPV-bound frozen in liquid nitrogen. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi TPV IP	R(Å) ±0.3 Å	FWHM (Å) ± 0.5 Å	% ensemble ± 5%
Closed	33.0	2.7	84
Semi-open	36.8	2.6	12
Wide-open	41.5	1.7	4

Figure 5-16. HIV-1PR Subtype C TPV-bound frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve (blue) and with the reconstructed echo curve (gray) corresponding to the sum of distance profiles; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum; Individual populations in Gaussian reconstruction of the TKR distance profile.



Csi TPV 37°C IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Closed	33.0	2.8	86
Semi-open	37.5	3.3	12
Wide-open	41.0	1.4	2

Figure 5-17. HIV-1PR Subtype C TPV-bound apo heated to 37 °C then frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction; Frequency domain spectrum

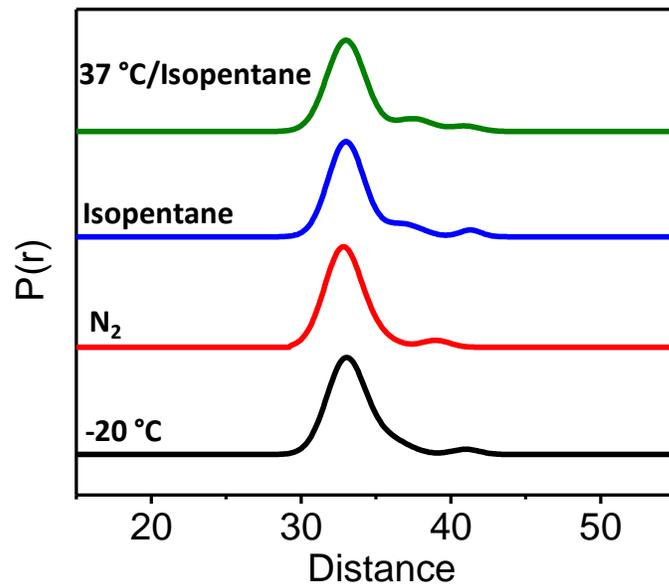
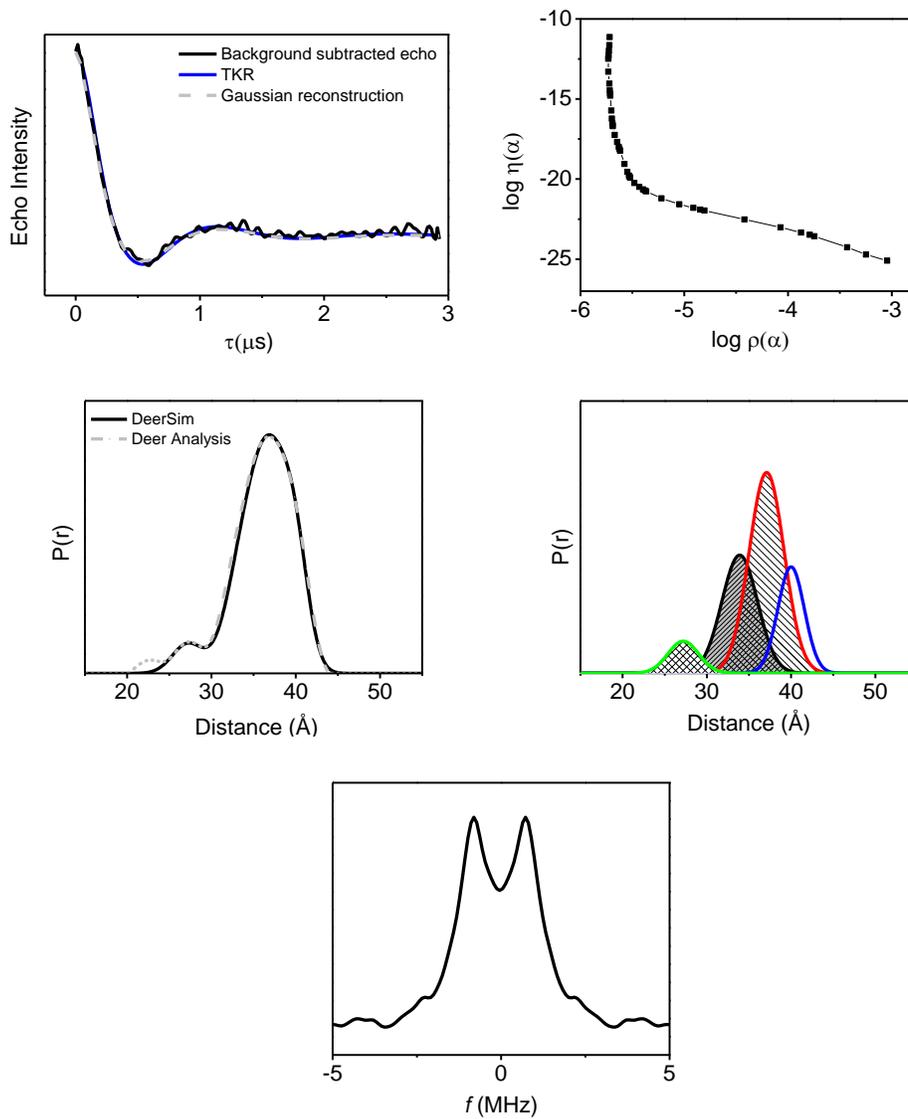
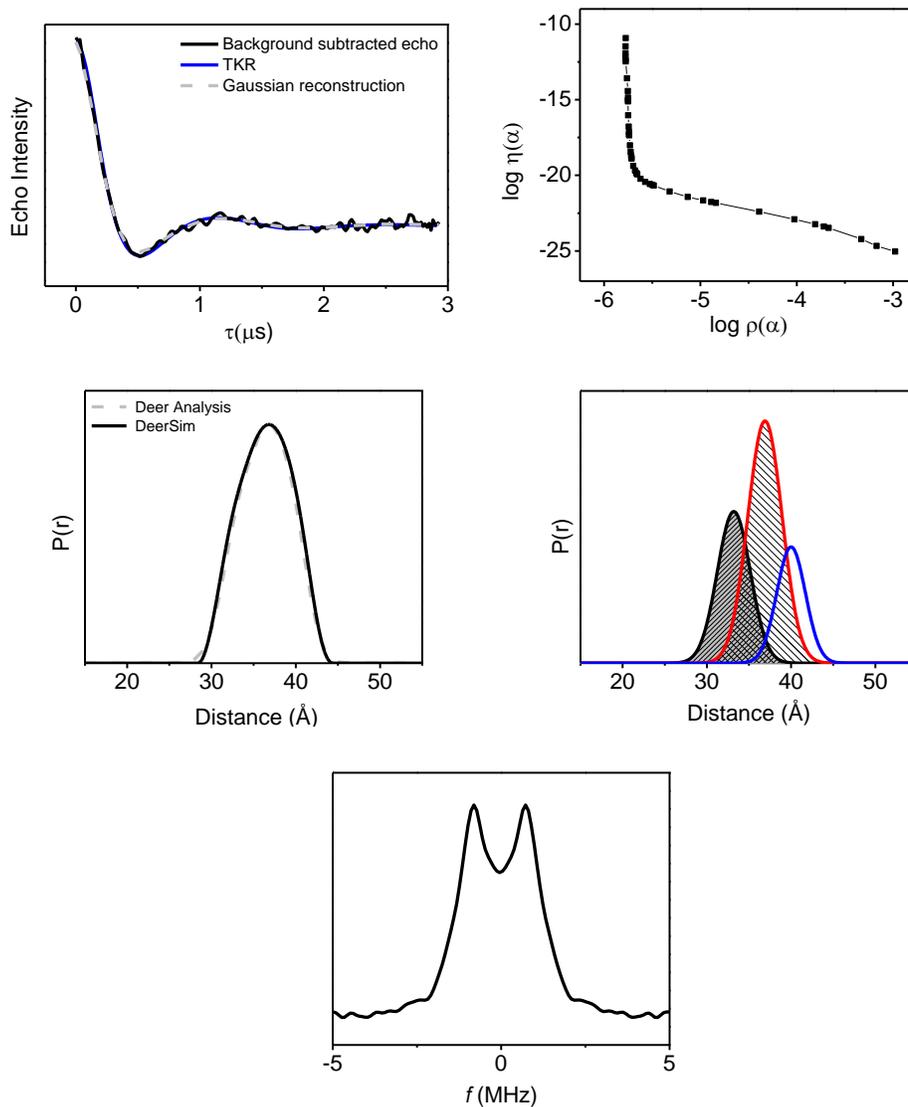


Figure 5-18. Stacked distance distribution profiles of HIV-1PR Subtype C K55R1 bound to TPV, after various freezing conditions, including -20 C, liquid nitrogen, isopentane/liquid nitrogen from room temperature, and isopentane/liquid nitrogen after heating to 37 °C.



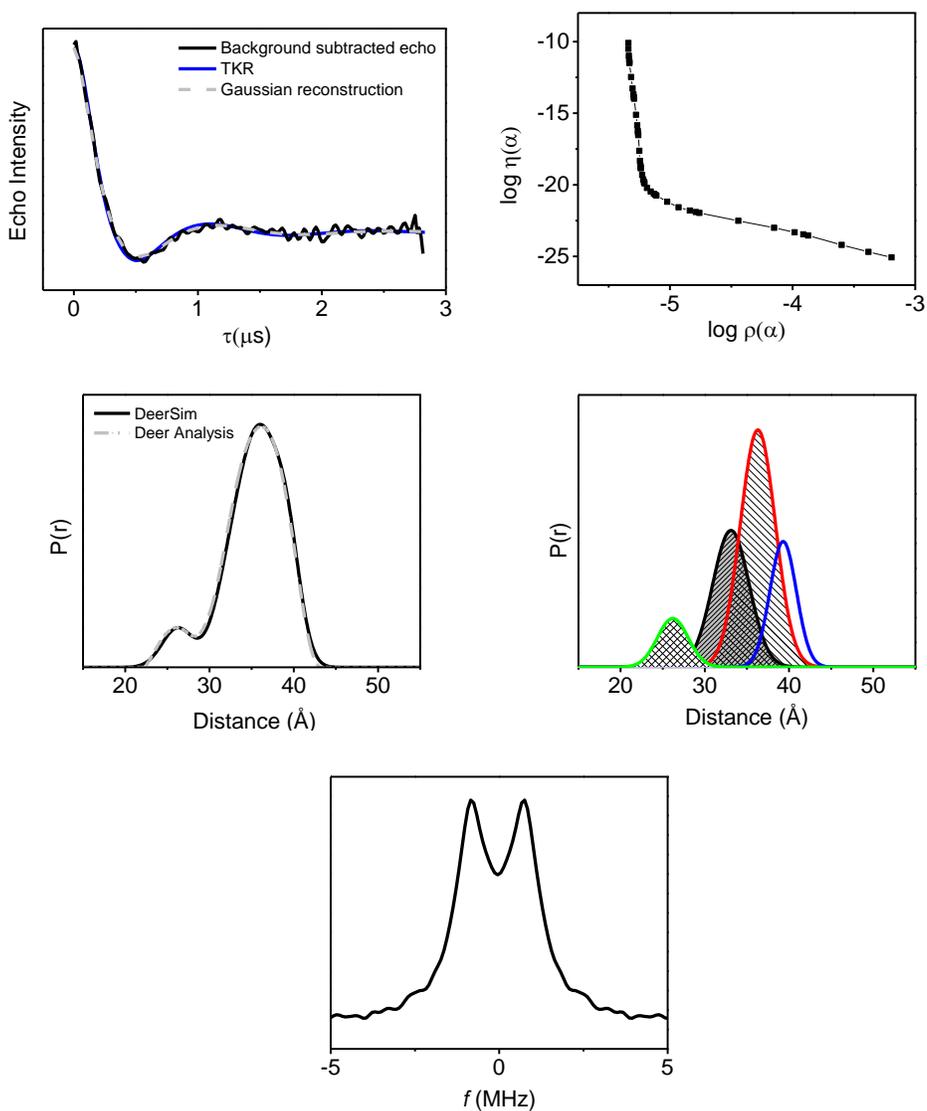
Csi IDV -20°C	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	27.2	4.3	6
Closed	33.9	5.0	27
Semi-open	37.1	4.9	50
Wide-open	40.0	3.7	17

Figure 5-19. HIV-1PR Subtype C IDV-bound frozen at -20. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile; Frequency domain spectrum.



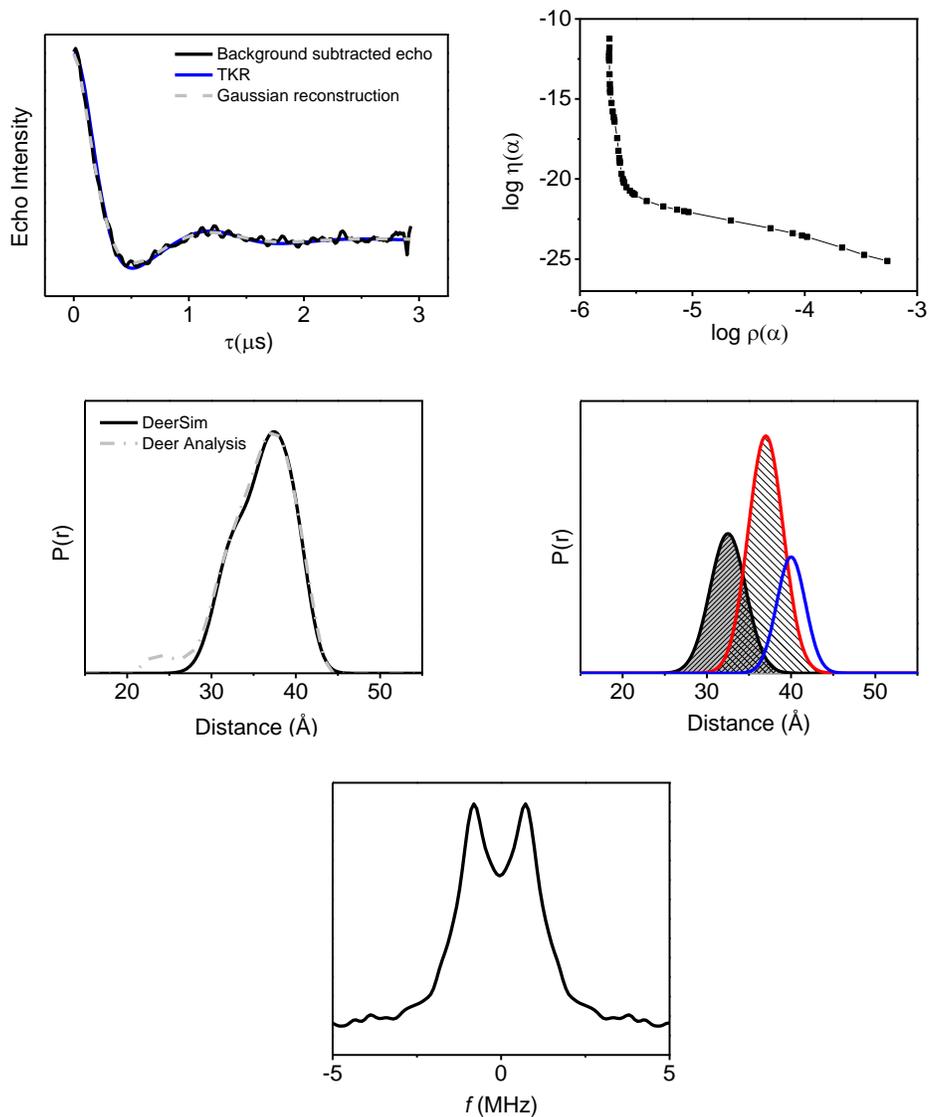
Csi IDV N ₂	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Closed	33.2	4.6	30
Semi-open	36.9	4.8	50
Wide-open	40.0	4.0	20

Figure 5-20. HIV-1PR Subtype C IDV-bound frozen in liquid nitrogen. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi IDV IP	R(Å) ±0.3 Å	FWHM (Å) ± 0.5 Å	% ensemble ± 5%
Tucked	26.2	4.8	6
Closed	33.1	4.8	27
Semi-open	36.3	3.7	50
Wide-open	3.93	3.7	17

Figure 5-21. HIV-1PR Subtype C IDV-bound frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi IDV 37°C IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Closed	32.5	5.0	30
Semi-open	37.9	4.9	50
Wide-open	40.0	4.0	20

Figure 5-22. HIV-1PR Subtype C IDV-bound apo heated to 37 °C then frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum

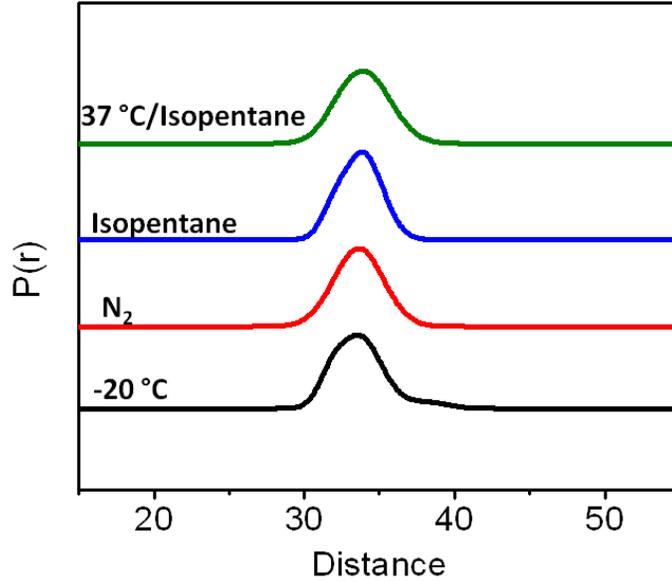
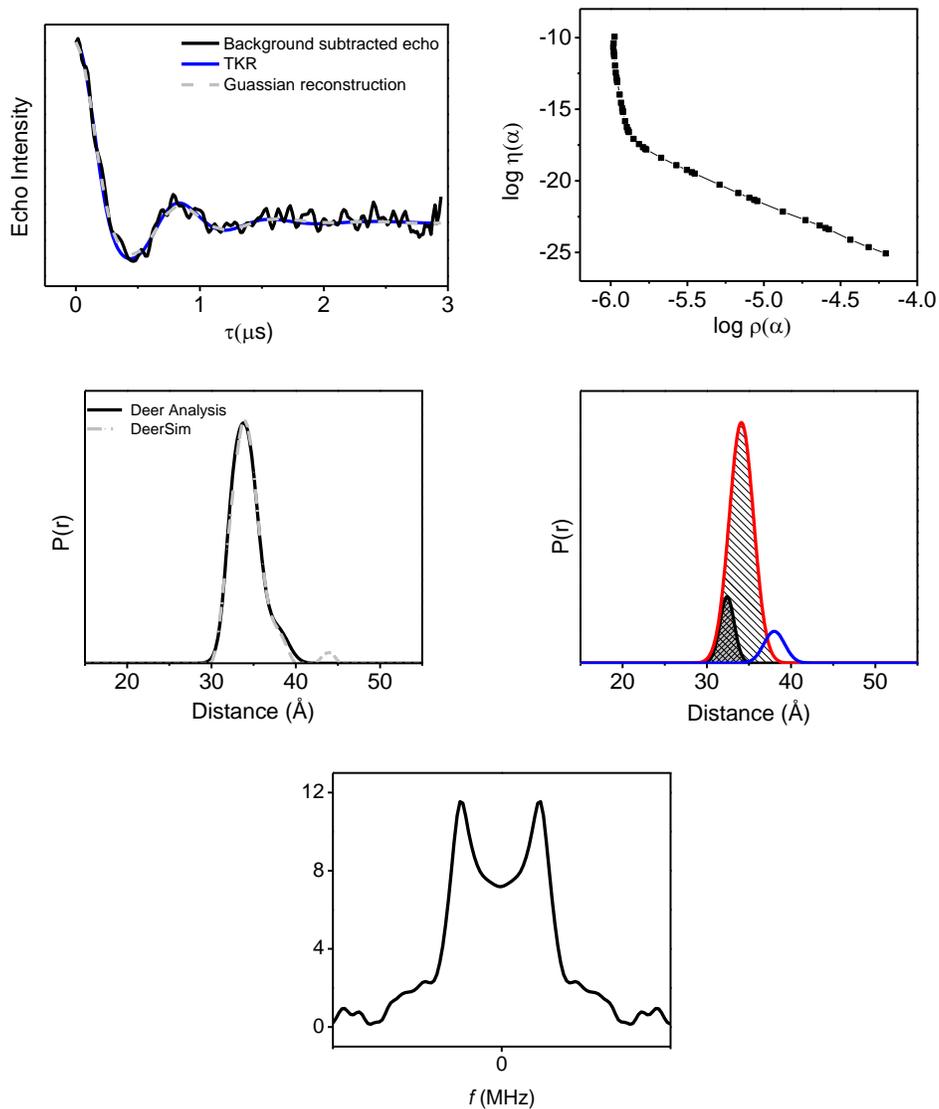
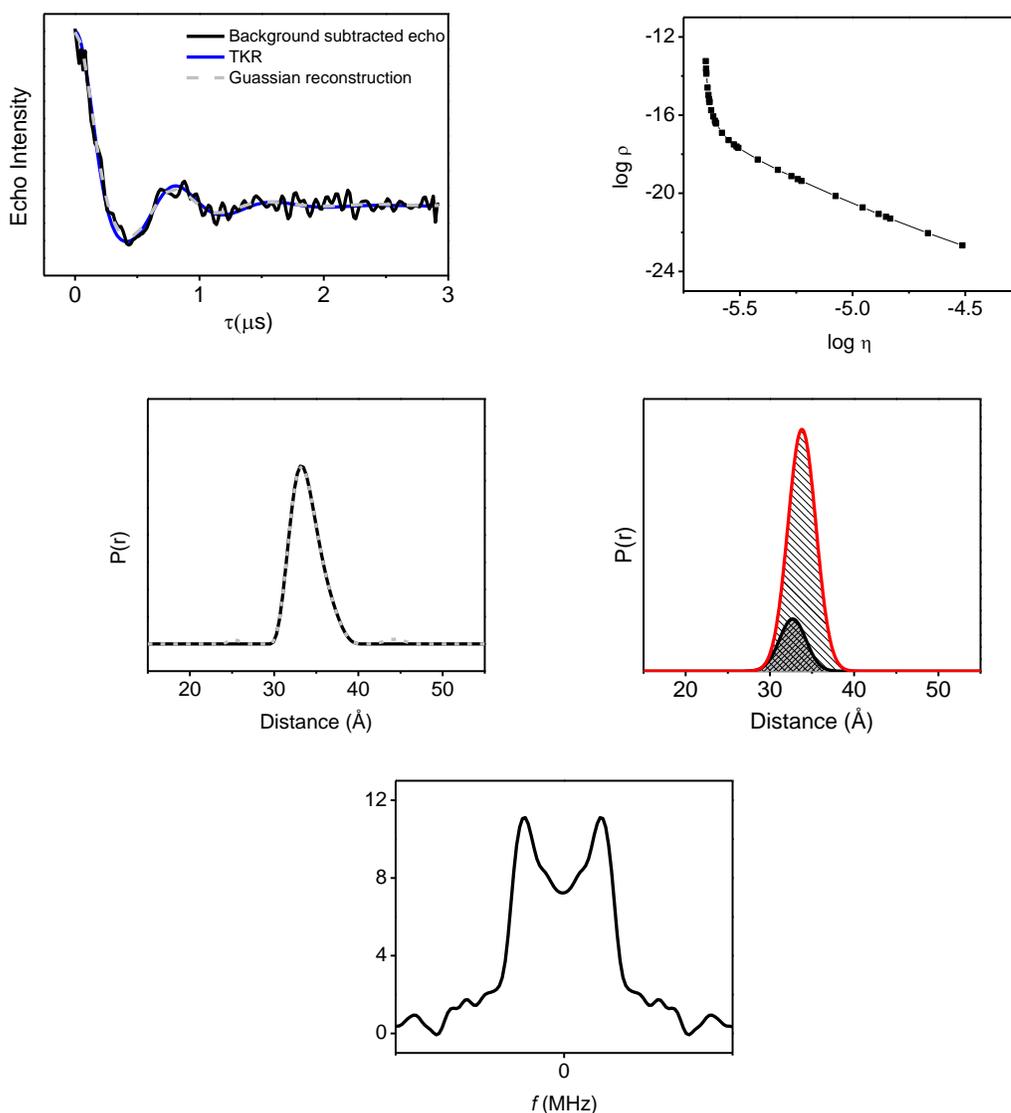


Figure 5-23. Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, bound with IDV after various freezing conditions, including -20 °C, liquid nitrogen, isopentane/liquid nitrogen from room temperature, and isopentane/liquid nitrogen after heating to 37 °C.



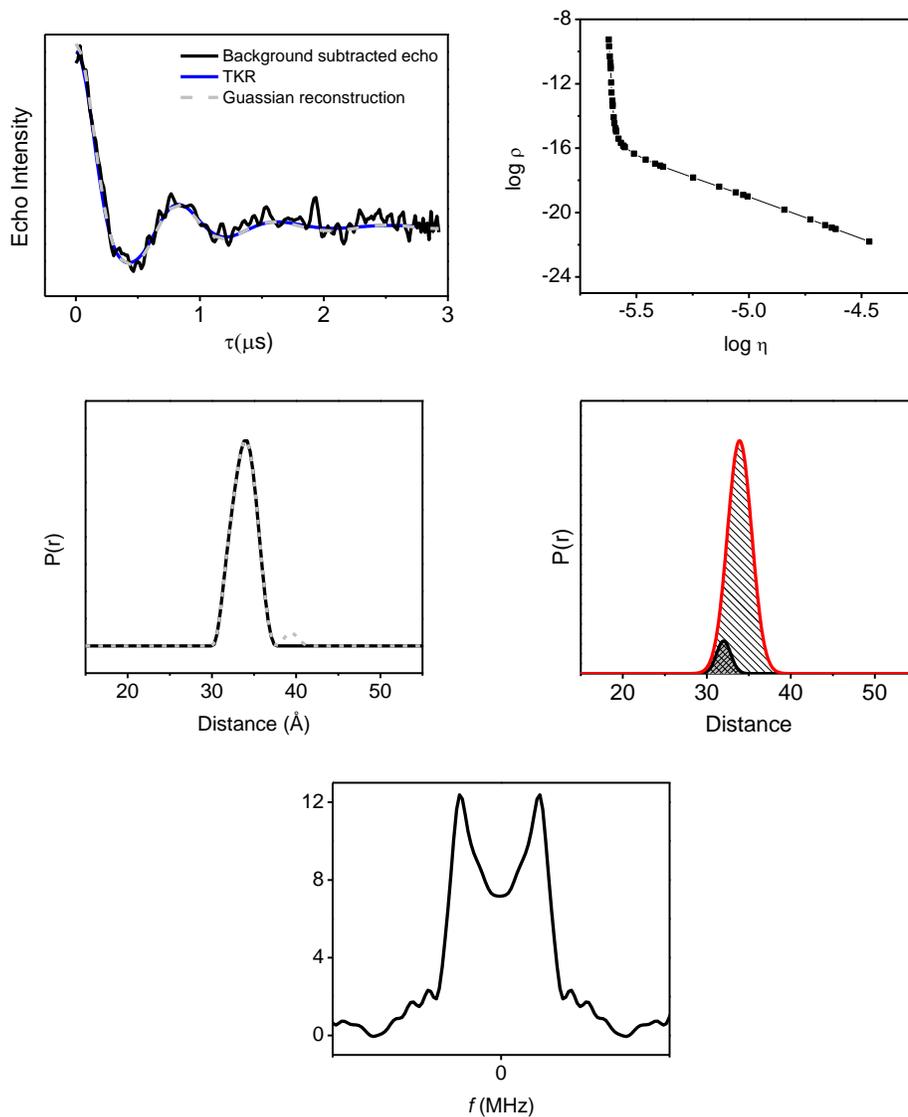
Csi CA-p2 -20°C	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	31.8	1.9	13
Closed	33.7	3.4	80
Semi-open	38.1	3.4	7

Figure 5-24. HIV-1PR Subtype C CA-p2-bound frozen at -20°C. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



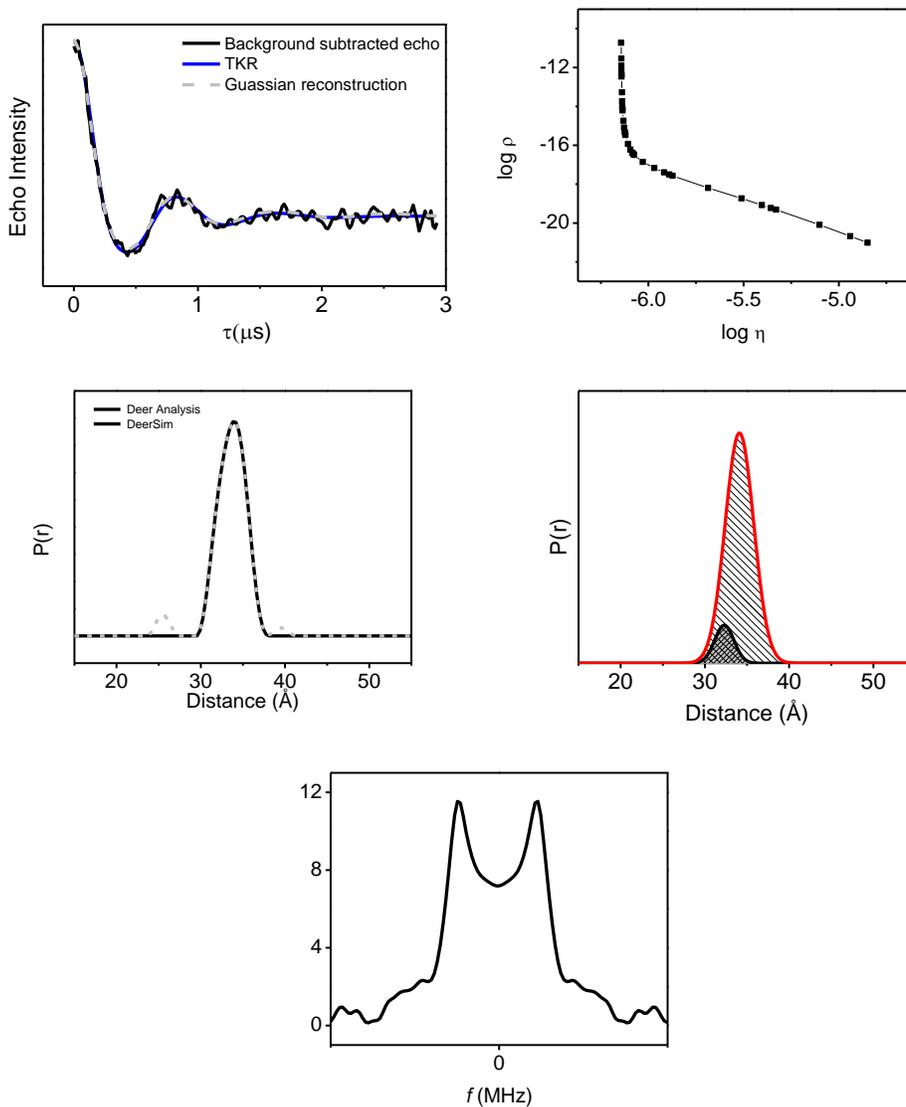
Csi CA-p2 N ₂	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	32.7	3.7	17
Closed	33.8	3.9	83

Figure 5-25. HIV-1PR Subtype C CA-p2-bound frozen in liquid nitrogen. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi CA-p2 IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	31.9	1.9	12
Closed	33.9	3.1	88

Figure 5-26. HIV-1PR Subtype C CA-p2-bound frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Csi CA-p2 37°C IP	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	32.3	2.7	10
Closed	34.1	4.0	90

Figure 5-27. HIV-1PR Subtype C CA-p2-bound apo heated to 37 °C then frozen in isopentane. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (blue) and with the reconstructed echo curve from DeerSim (gray) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (gray dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum

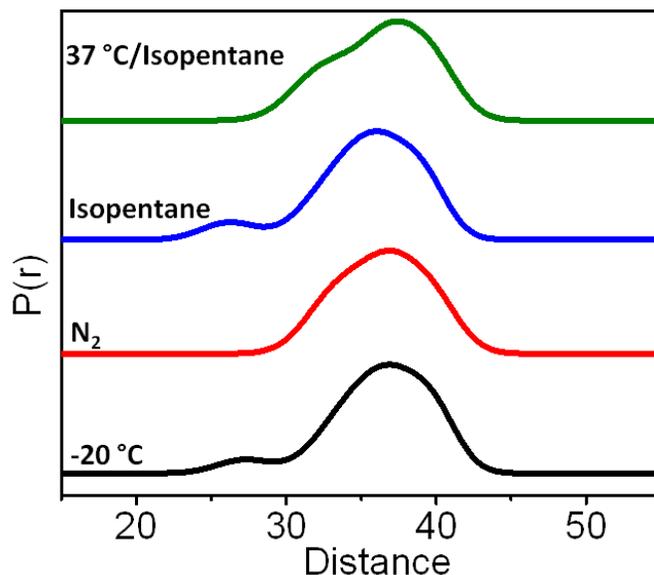


Figure 5-28. Stacked distance distribution profiles of HIV-1PR Subtype C, K55R1, bound with CA-p2 after various freezing conditions, including -20 °C, liquid nitrogen, isopentane/liquid nitrogen from room temperature, and isopentane/liquid nitrogen after heating to 37 °C.

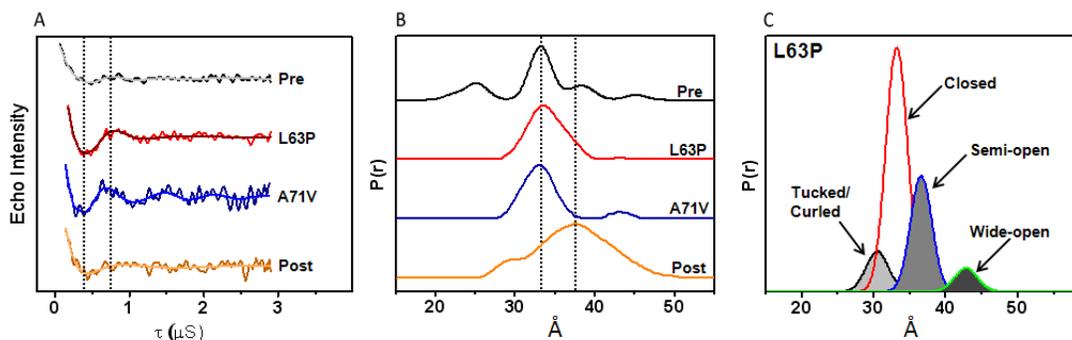


Figure 5-29. DEER results for drug pressure mutations. (A) Background subtracted DEER echo curves for PRE (grey), L63P (red), A71V (blue), and POST (orange) constructs, (B) distance profiles from TKR analysis, and (C) Gaussian populations used to regenerate the TKR distance profile of L63P, showing contributions from individual conformations described as tucked/curled (black), closed (red), semi-open (blue), and wide-open (green).

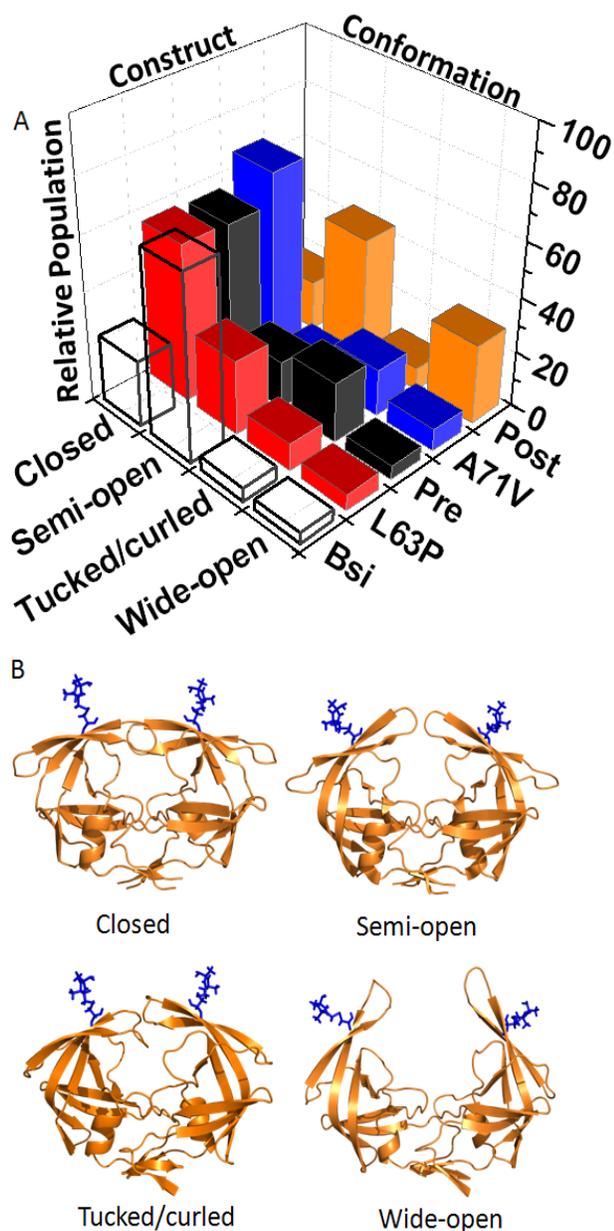
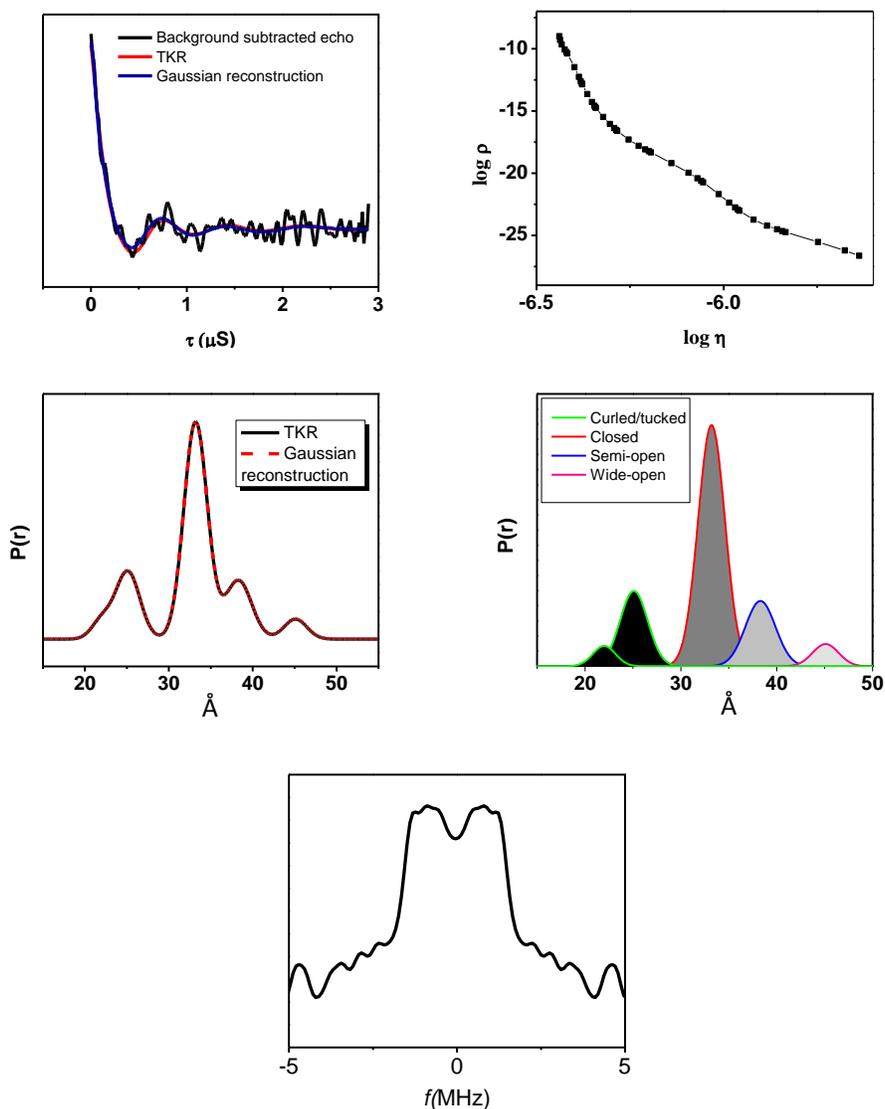
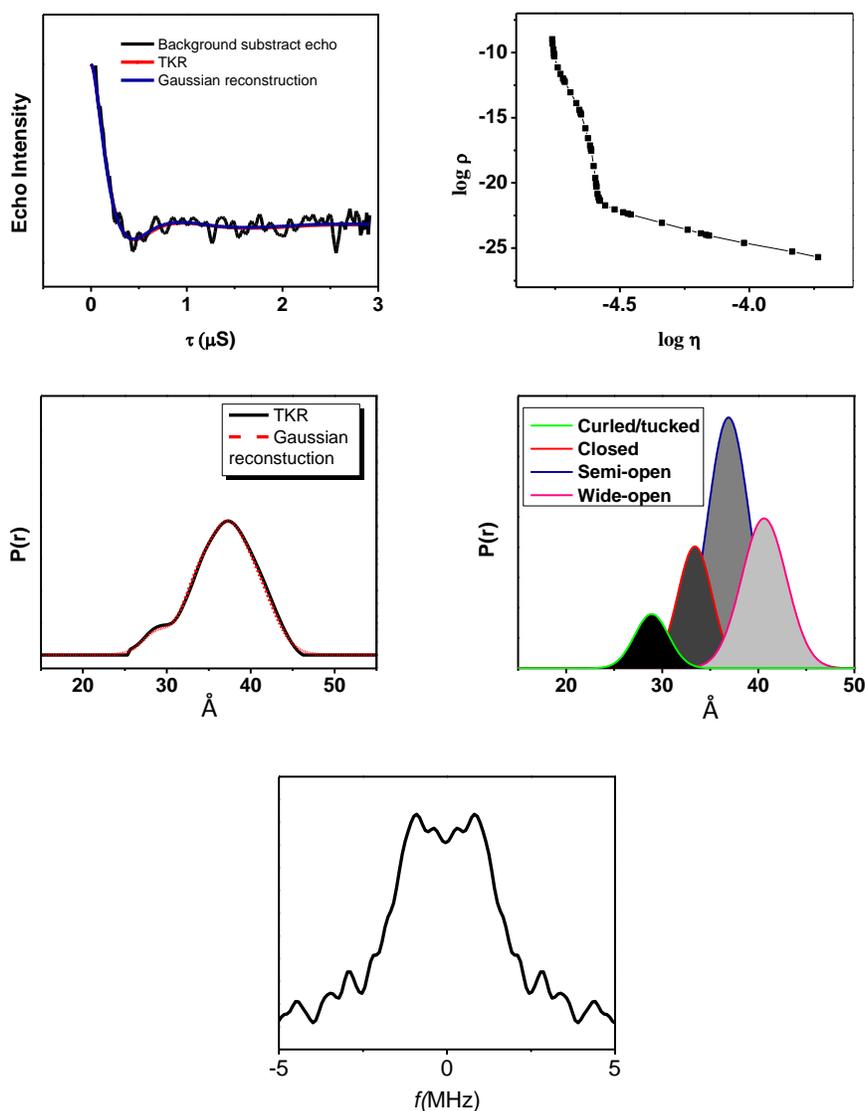


Figure 5-30. Population analysis and molecular dynamics simulations for HIV-1 PR. (A) Conformational ensembles consisting of tucked/curled, closed, semi-open and wide-open conformations determined via Gaussian regeneration of the DEER distance profiles for each construct. Error was approximated at (5%). Populations are summed to define the conformational ensemble of each construct. (B) Structures of HIV-1 PR in the closed, semi-open, tucked/curled and wide-open flap states. In each structure, the MTSL spin label at position 55 is shown in blue.



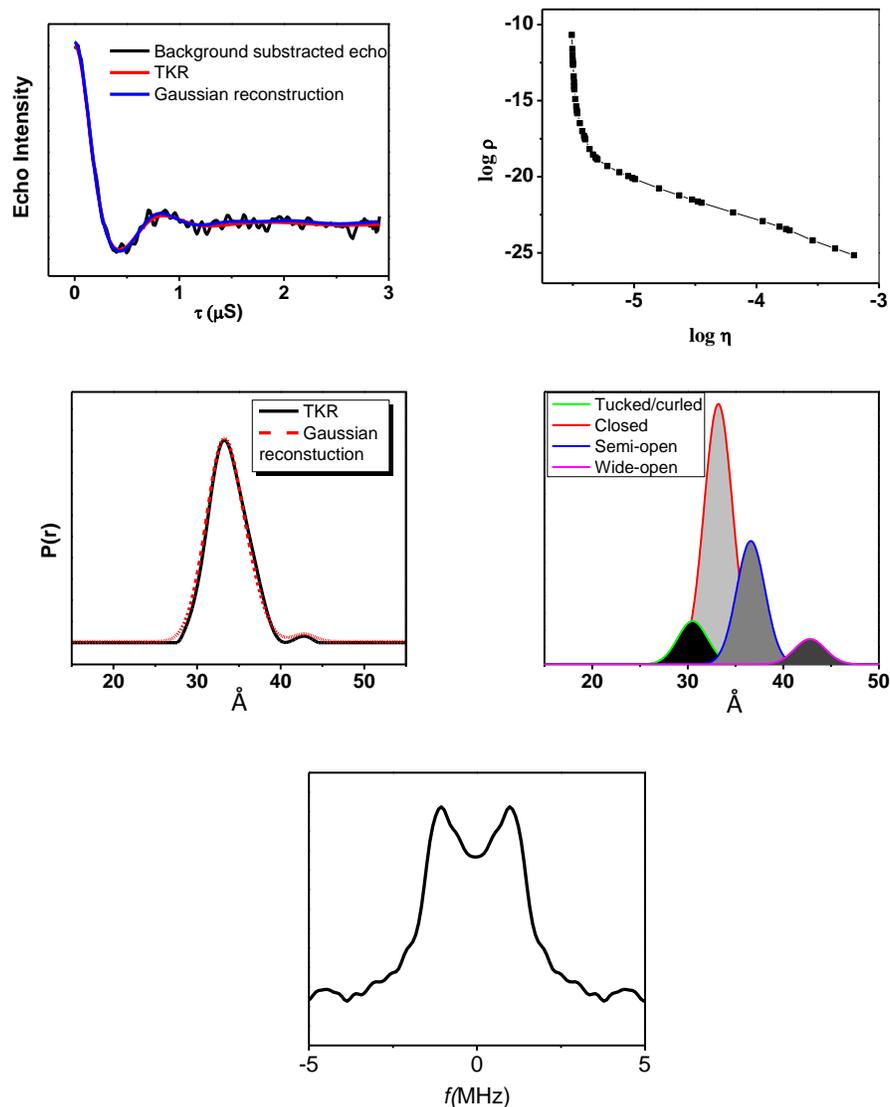
PRE apo	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	22.0	2.8	4
Tucked	25.1	3.4	18
Closed	33.2	3.3	56
Semi-open	38.3	3.7	17
Wide-open	45.1	3.2	5

Figure 5-31. HIV-1PR PRE apo. A Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red) and with the reconstructed echo curve from DeerSim (blue) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (red dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



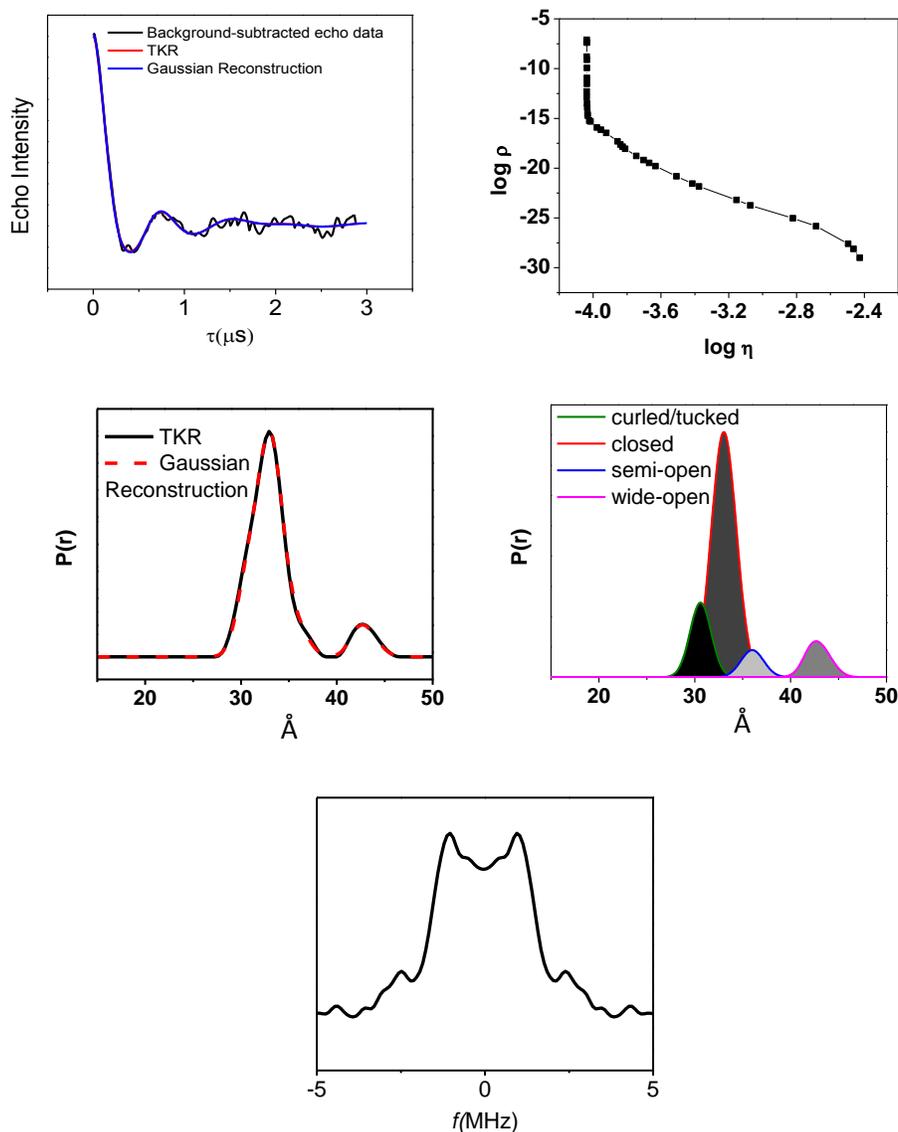
Post apo	R(Å) ± 0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	28.9	4.2	8
Closed	33.4	4.2	18
Semi-open	36.9	5.1	45
Wide-open	40.6	5.5	29

Figure 5-32. HIV-1PR POST apo. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red) and with the reconstructed echo curve from DeerSim (blue) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (red dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Subtype B L63P apo	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	30.5	3.7	10
Closed	33.2	3.5	57
Semi-open	36.6	3.5	27
Wide-open	42.8	3.8	6

Figure 5-33. HIV-1PR Subtype B L63P apo. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red) and with the reconstructed echo curve from DeerSim (blue) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (red dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum



Subtype B A71V apo	R(Å) ± 0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	30.6	2.6	17
Closed	33.0	3.0	67
Semi-open	36.0	2.8	7
Wide-open	42.8	3.0	9

Figure 5-34. HIV-1PR Subtype B A71V Apo. Background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red) and with the reconstructed echo curve from DeerSim (blue) corresponding to the sum of distance profiles comprising the Gaussian populations; L curve; Distance profile from TKR analysis (black) overlain with the summed Gaussian population profile (red dashed); Individual populations necessary for Gaussian reconstruction of the distance profile generated from TKR; Frequency domain spectrum

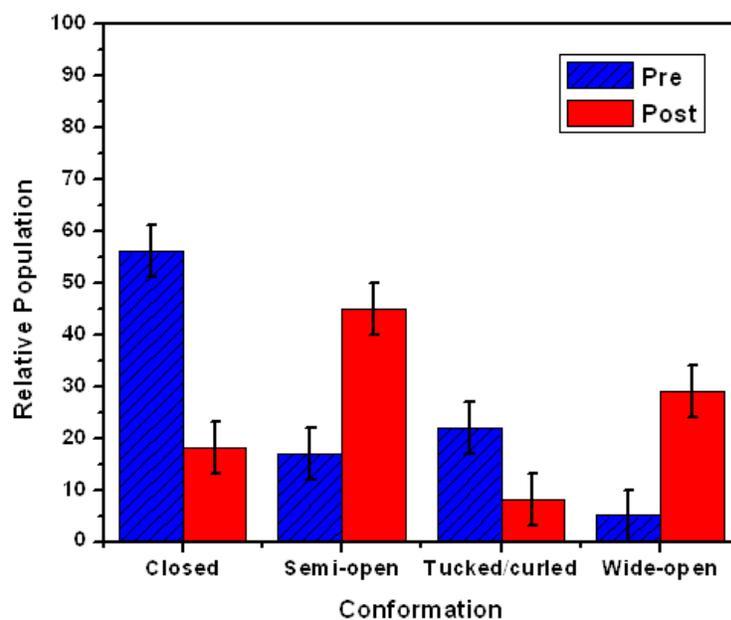


Figure 5-35. The relative populations (%) for closed, semi-open, tucked/curled, and wide-open conformations of PRE apo (blue) and POST apo (red).

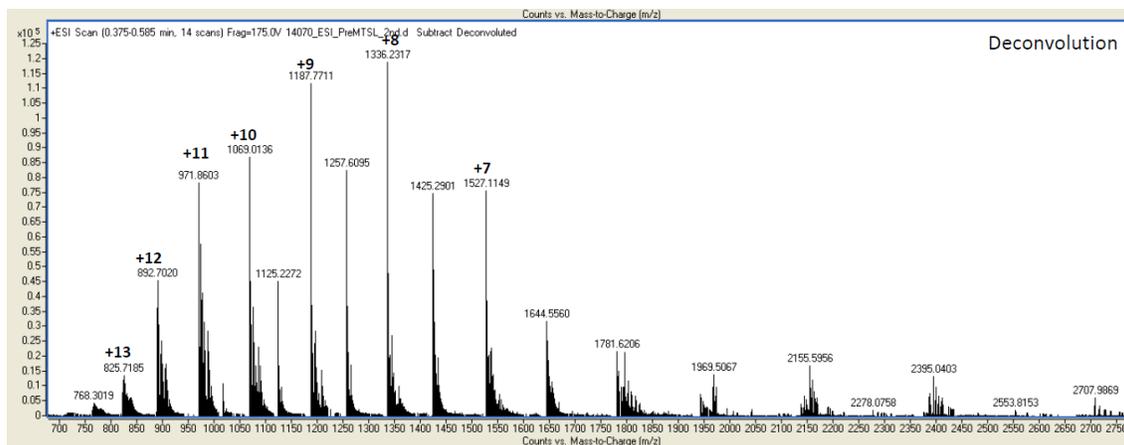


Figure 5-36. Mass spectra of spin-labeled PRE. The spectrum shows major peaks with charged states indicated.

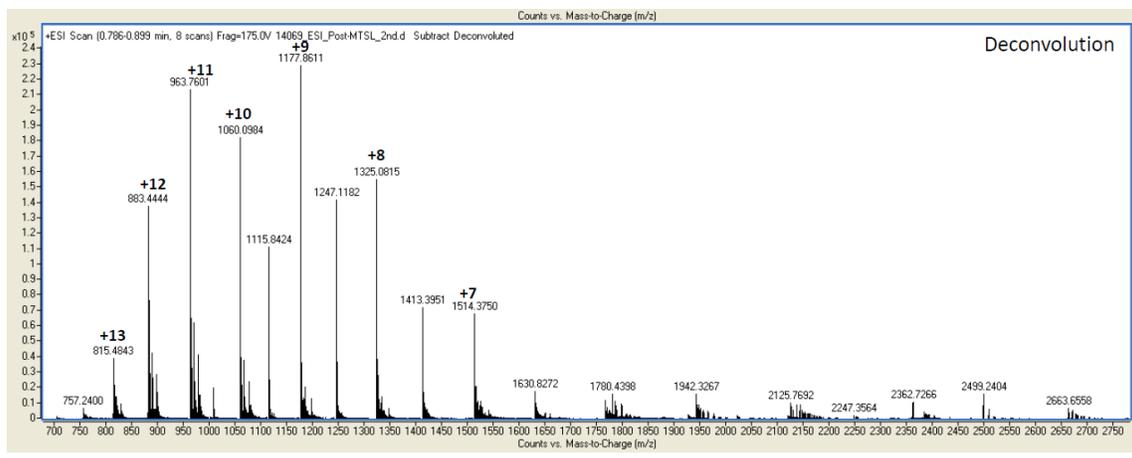


Figure 5-37. Mass spectra of spin-labeled POST. The spectrum shows major peaks with charged states indicated.

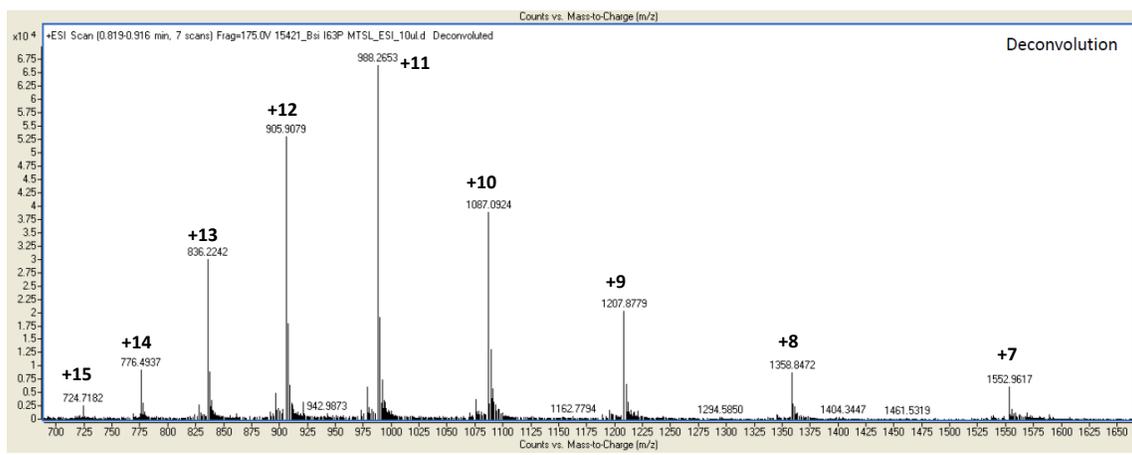


Figure 5-38. Mass spectra of spin-labeled L63P. The spectrum shows major peaks with charged states indicated.

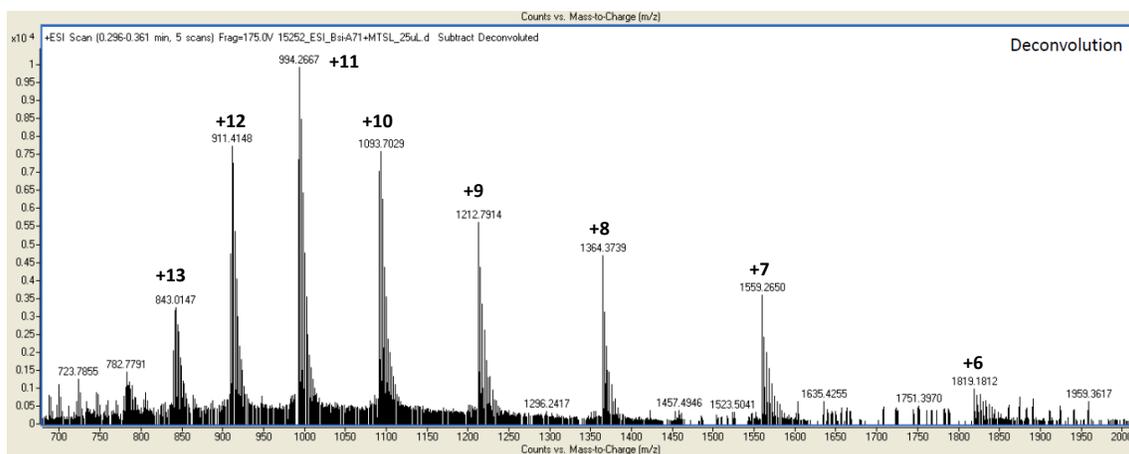
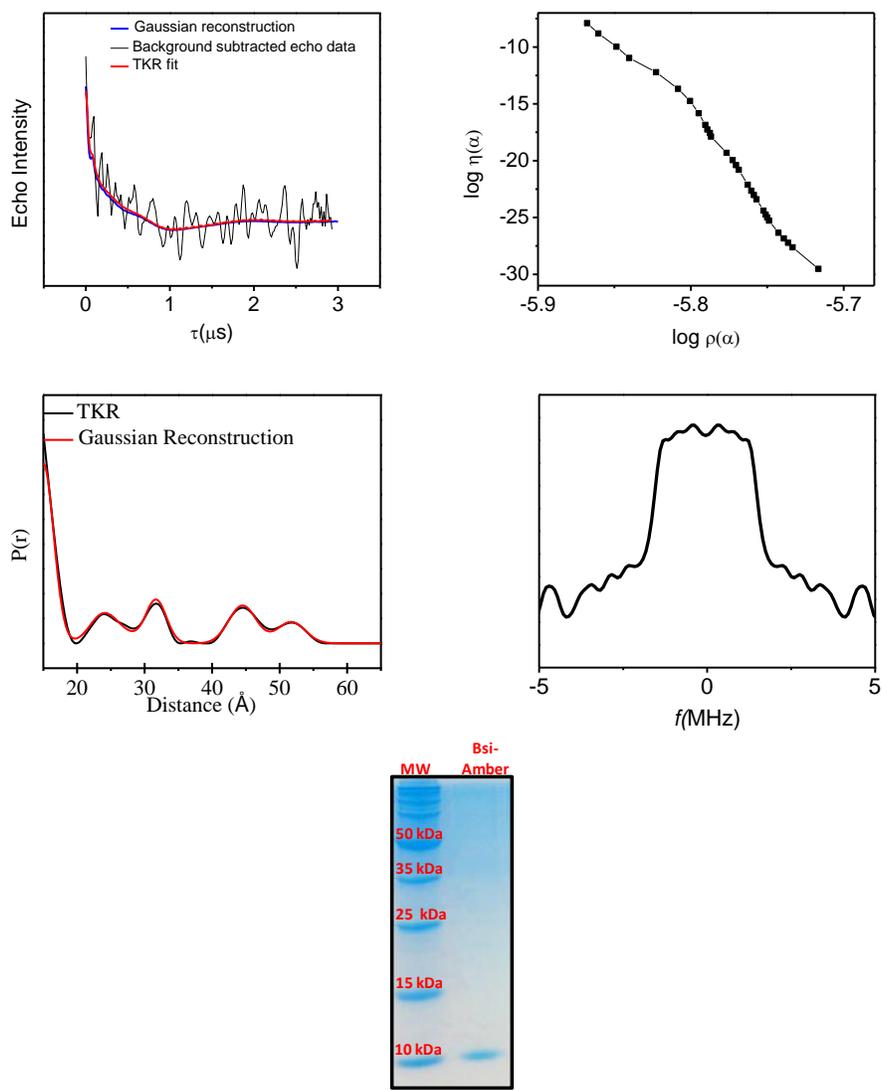


Figure 5-39. Mass spectra of spin-labeled A71V. The spectrum shows major peaks with charged states indicated.

HIV Protease Construct	Theoretical MW(Da) (without MTSL)	Theoretical MW(Da) (with MTSL)	Observed MW(Da) (with MTSL)
Pre	10690.6	10875.0	10874.9±0.2
Post	10601.4	10785.8	10785.6±0.2
Bsi L63P	10686.6	10871.0	10870.9±0.1
Bsi A71V	10730.8	10915.2	10915.0±0.2

Figure 5-40. Analysis of mass spectra for HIV-1 protease constructs.



Subtype B Amber apo	R(Å) ±0. 3Å	FWHM (Å) ± 0. 5Å	% ensemble ± 5%
Tucked	24.7	5.2	23
Closed	33.4	4.0	29
Semi-open	35.3	4.4	27
Wide-open	52.5	4.6	21

Figure 5-41. Bsi-Amber. Long pass filtered and background subtracted dipolar echo curve (black) overlaid with the TKR regenerated echo curve from DeerAnalysis (red); L curve utilized to choose the appropriate regularization parameter ($\lambda = 5$); Distance profile from TKR analysis; Frequency domain spectrum; Protein gel after Q-column isolation of Bsi-Amber.

CHAPTER 6 FUTURE DIRECTIONS

The results from SDSL EPR, DSC, and fluorescence experiments of GM2AP CYS constructs indicate that the flexible loops do not undergo conformational exchange as a result of ligand binding. But, loop flexibility is an important attribute in the ability of GM2AP to extract its multiple ligands. To further investigate how ligand binding alters the average conformation of GM2AP, uniformly ^{15}N labeled protein was generated for solution nuclear magnetic resonance (NMR) heteronuclear quantum coherence (HSQC) measurements were performed by our post-doctoral fellow Yong Ran. Figure 6-1 shows HSQC spectra of GM2AP and with GM2 micelles (red) at pH 4.8 (blue). The binding of GM2 also does not induce a shift in most of the resonances, only a few peaks are seen to shift or disappear. Indicating that GM2AP does not undergo a global conformational shift as a result of ligand binding. NMR assignments for GM2AP are lacking, and future NMR studies are aimed at mapping out the specific regions in GM2AP that are altered during interactions with ligands. In attempt to overcome the overlapping NMR resonance peaks from the fast conformational exchange sub μs timescale, crystallizations trials (Figure 6-2) have begun in attempted to gain amino acid specific assignments from solid state NMR.

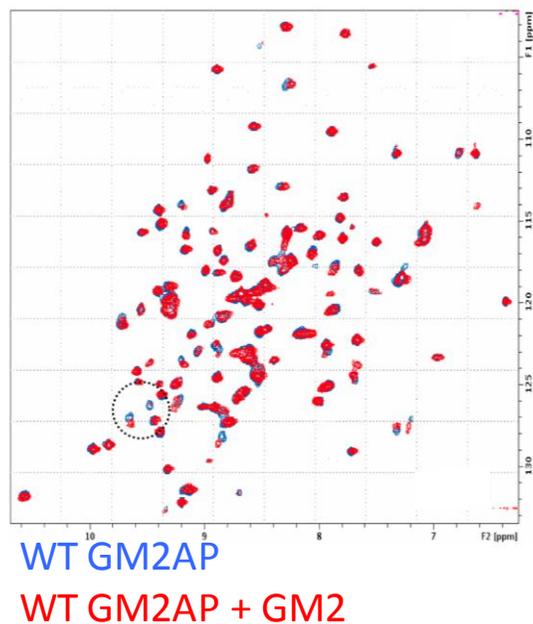


Figure 6-1. HSQC NMR spectra of uniformly labeled ^{15}N GM2AP upon addition of 4x molar excess GM2 micelles (red) at acidic pH 4.8 (blue). Sample buffers were exchanged via gel filtration chromatography and not as titrations. The final sample contains 0.15 mM GM2AP and 20mM NaOAc. For the GM2 binding experiments, the spectra were collected before and after incubate with 0.6mM GM2 micelles at room temperature. 2D ^1H , ^{15}N -HSQC spectra were collected on Bruker 600 Mhz spectrometer with 1mm cryoprobe at 20 °C. The spectra were processed with NMRPipe and Sparky.

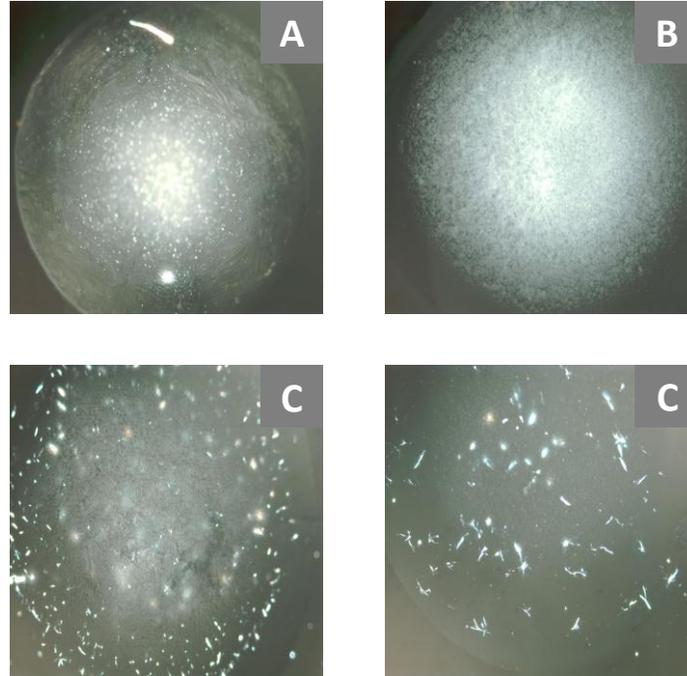


Figure 6-2. Crystallization trials of wild-type GM2AP. A) WT GM2AP hanging drop crystal trial in Hepes buffered solution (0.1 M, pH 7.5) containing 30% (w/v) polyethyleneglycol (MM 4000), 10% (v/v) propanol. B) WT GM2AP hanging drop crystal trial in Hepes buffered solution (0.1 M, pH 7.5) containing 25% (w/v) polyethyleneglycol (MM 4000), 10% (v/v) propanol. C) WT GM2AP hanging drop crystal trial in Hepes buffered solution (0.1 M, pH 7.5) containing 20% (w/v) polyethyleneglycol (MM 4000), 10% (v/v) propanol.

APPENDIX
CW AND PULSED EPR PARAMETERS

Table A-1. Typical CW EPR parameters used for work reported within

Parameter	Value
Number of points	1024
Center field	~3450 G
Number of scans	5-25
Sweep width	20 – 100 G
Acquisition time	40.63 sec
Frequency	~9.6 GHz
Power	20 dB – 20 mW
Receiver Gain	$1 \times 10^3 - 5 \times 10^5$
Modulation Amplitude	~1 G
Time Constant	0.082 – 0.164 sec
Receiver Phase	100

Table A-2. Typical pulsed EPR parameters used for DEER reported within.

Parameter	Value
Shot repetition time	4000
Sweep width	160-200 G
Number of scans	variable
Shots/point	100
Center Field	~3460 G
Low Field	~3432 G
Frequency	~9.6 GHz
Pulsed Attenuation	0.1
Video Bandwidth	25 MHz ⁵
Modulation Amplitude	~1 G
Time Constant	0.082 – 0.164 sec
Receiver Phase	100

Table A-3. 4-pulse table used for DEER experiments performed in deuterated matrices.

	+x Pulse		Acquisition Trigger	
Position	200	600	Position	
Pulse Length	16	32	Pulse Length	16
Pos. Display		8	Pos. Display	

LIST OF REFERENCES

- Ahn, V. E., K. F. Faull, et al. (2003). Crystal structure of saposin B reveals a dimeric shell for lipid binding. *PNAS* **100**: 38-43.
- Ahn, V. E., P. Leyko, et al. (2006). Crystal structures of saposins A and C. *Protein Sci* **15**: 1849-57.
- Altenbach, C., K. Cai, et al. (1999). Structural features and light-dependent changes in the sequence 306-322 extending from helix VII to the palmitoylation sites in rhodopsin: a site-directed spin-labeling study. *Biochemistry* **38**: 7931-7.
- Altenbach, C., S. L. Flitsch, et al. (1989). Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry* **28**: 7806-12.
- Altenbach, C., J. Klein-Seetharaman, et al. (1999). Structural features and light-dependent changes in the sequence 59-75 connecting helices I and II in rhodopsin: a site-directed spin-labeling study. *Biochemistry* **38**: 7945-9.
- Altenbach, C., Oh, K. J., Trabanino, R. J., Hideg, K., and Hubbell, W. L. (2001) Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry* **40**: 15471-15482.
- Altenbach, C., K. Yang, et al. (1996). Structural features and light-dependent changes in the cytoplasmic interhelical E-F loop region of rhodopsin: a site-directed spin-labeling study. *Biochemistry* **35**: 12470-8.
- Ashorn, P., T. J. McQuade, et al. (1990). An inhibitor of the protease blocks maturation of human and simian immunodeficiency viruses and spread of infection. *PNAS* **87**: 7472-6.
- Atkins, P. and J. d. Paula, Eds. (2005). *Elements of physical chemistry*. Physical Chemistry. New York, W.H. Freeman and Company.
- Baca, M. and S. B. Kent (1993). Catalytic contribution of flap-substrate hydrogen bonds in HIV-1 protease explored by chemical synthesis. *PNAS* **90**: 11638-11642.
- Badock, V, M. Raida, et al. (1998). Distinction between the three disulfide isomers of guanylin 99-115 by low-energy collision-induced dissociation. *Rapid Commun. Mass Spectrom* **12**: 1952–1956.
- Barnakov, A., C. Altenbach, et al. (2002). Site-directed spin labeling of a bacterial chemoreceptor reveals a dynamic, loosely packed transmembrane domain. *Protein Sci* **11**: 1472-81.

- Barnes, J. P., Z. Liang, et al. (1999). A multifrequency electron spin resonance study of T4 lysozyme dynamics. *Biophys J* **76**: 3298-306.
- Bean, M. F., S. Carr, et al. (1992). Characterization of disulfide bond position in proteins and sequence analysis of cystine-bridged peptides by tandem mass spectrometry. *Anal Biochem* **201**: 216–226.
- Beasusoleil, S. A., M. Jedrychowski, et al. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. *PNAS* **101**: 12130–12135.
- Beckey, H. D. (1969). Field ionization mass spectrometry. *Research/Development* **20**: 26.
- Berliner, L. J., J. Grunwald, et al. (1982). A novel reversible thiol-specific spin label: papain active site labeling and inhibition. *Anal Biochem* **119**: 450-5.
- Beu, S. C., D. Laude, et al. (1992). Elimination of axial ejection during excitation with a capacitively coupled open trapped-ion cell for Fourier-transform ion-cyclotron resonance mass-spectrometry. *Anal Chem* **64**: 177–180.
- Blackburn, M. E., A. M. Veloro, et al. (2009). Monitoring inhibitor induced conformational population shifts in HIV-1 protease by pulsed EPR spectroscopy. *Biochemistry* **48**: 8765-8767.
- Blakney, G. T., D. Robinson, et al. (2005). *53rd Amer. Soc. Mass Spectrom. Ann. Conf. on Mass Spectrom and Allied Topics*, San Antonio, TX.
- Bordignon, E., Klare, J. P., Doebber, M., Wegener, A. A., Martell, S., Engelhard, M., and Steinhoff, H. J. (2005) Structural analysis of a HAMP domain: the linker region of the phototransducer in complex with sensory rhodopsin II. *J Biol Chem* **280**: 38767-38775.
- Brustad, E. M., E. A. Lemke, et al. (2008). A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *J Am Chem Soc* **130**: 17664-5.
- Buchbinder, S. P., M. H. Katz, et al. (1994). Long-term HIV-1 infection without immunological progression. *AIDS* **8**: 1123-1128.
- Budil, D. E., Lee, S., Saxena, S., Freed, J. H. (1996) Nonlinear-Least-Squares Analysis of Slow-Motion EPR Spectra in One and Two Dimensions Using a Modified Levenberg-Marquardt Algorithm. *J Magn Reson* **120**: 155-189.
- Carr, J. K., B. T. Foley, et al. (1998). Reference sequences representing the principal genetic diversity of HIV-1 in the pandemic. *HIV Sequence Compendium*: 10-19.

- Chiang, Y.-W., P. P. Borbat, et al. (2005). Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *J Magn Reson* **177**: 184-196.
- Chiang, Y.-W. B., Peter P.; Freed, Jack H. (2005). Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *J Magn Reson* **177**: 184-196.
- Chin, J. W., T. A. Cropp, et al. (2003). Progress toward an expanded eukaryotic genetic code. *Chem Biol* **10**: 511-9.
- Ciaffoni, F., R. Salvioli, et al. (2001). Saposin D solubilizes anionic phospholipid-containing membranes. *J Biol Chem* **276**: 31583-9.
- Ciaffoni, F., M. Tatti, et al. (2006). Saposin B binds and transfers phospholipids. *J Lipid Res* **47**: 1045-53.
- Clemente, J., R. Hemrajani, et al., (2003). Secondary mutations M36I and A71V in the human immunodeficiency virus type 1 protease can provide an advantage for the emergence of the primary mutation D30N. *Biochemistry* **42**:15029-35.
- Clore, G. M. (2008) Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement. *Mol Biosyst* **4**: 1058-1069.
- Columbus, L. and W. L. Hubbell (2002). A new spin on protein dynamics. *Trends Biochem Sci* **27**: 288-295.
- Columbus, L. and W. L. Hubbell (2004). Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* **43**: 7273-87.
- Columbus, L., T. Kalai, et al. (2001). Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure. *Biochemistry* **40**: 3828-3846..
- Conzelmann, E., Burg, J., Stephan, G., and Sandhoff, K. (1982) Complexing of glycolipids and their transfer between membranes by the activator protein for degradation of lysosomal ganglioside GM2. *Eur J Biochem* **123**: 455-464.
- Conzelmann, E. and K. Sandhoff (1979). Purification and characterization of an activator protein for the degradation of glycolipids GM2 and GA2 by hexosaminidase A. *Hoppe Seylers Z Physiol Chem* **360**: 1837-49.
- Cooper, H. J., C. Hakansson, et al. (2005). The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom Rev* **24**: 201–222.

- Darmoise, A., P. Maschmeyer, et al. (2010). The immunological functions of saposins. *Adv Immunol* **105**: 25-62.
- Ding, F., M. Layten, et al. (2008). Solution structure of HIV-1 protease flaps probed by comparison of molecular dynamics simulation ensembles and EPR experiments. *J Am Chem Soc* **130**: 7184-5.
- Douek, D. C., M. Roederer, et al. (2009). Emerging concepts in the immunopathogenesis of AIDS. *Annu Rev Med* **60**: 471 - 484.
- Dougherty, R. C. (1981). Negative chemical ionization mass spectrometry: applications in environmental analytical chemistry. *Biomed Mass Spectrom* **8**: 283-92.
- Drahl, C. e. (2009). Protease Flaps Give HIV Drugs the Slip. *Chemical and Engineering News* **87**: 1.
- Eaton, G. R., S. S. Eaton, et al. (1998). *Foundations of Modern EPR*. River Edge, World Scientific Publishing Co. Pte. Ltd.
- Fanucci, G. E., Cadieux, N., Piedmont, C. A., Kadner, R. J., and Cafiso, D. S. (2002) Structure and dynamics of the beta-barrel of the membrane transporter BtuB by site-directed spin labeling. *Biochemistry* **41**: 11543-11551.
- Fanucci, G. E. and D. S. Cafiso (2006). Recent advances and applications of site-directed spin labeling. *Curr Opin Struct Biol* **16**: 644-653.
- Farahbakhsh, Z. T., C. Altenbach, et al. (1992). Spin labeled cysteines as sensors for protein-lipid interaction and conformation in rhodopsin. *Photochem Photobiol* **56**: 1019-33.
- Ferguson, M. A. and A. F. Williams (1988). Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu Rev Biochem* **57**: 285-320.
- Fleissner, M. R., E. M. Brustad, et al. (2009). Site-directed spin labeling of a genetically encoded unnatural amino acid. *PNAS* **106**: 21637-42.
- Foulkes-Murzycki, J., E. W. Scott, et al. (2007) Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease. *Structure* **15**: 225-33.
- Freedberg, D. I., R. Ishima, et al. (2002). Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations. *Protein Sci* **11**: 221-232.
- Furst, W., W. Machleidt, et al. (1988). The precursor of sulfatide activator protein is processed to three different proteins. *Biol Chem Hoppe Seyler* **369**: 317-28.

- Furst, W. and K. Sandhoff (1992). Activator proteins and topology of lysosomal sphingolipid catabolism. *Biochim Biophys Acta* **1126**: 1-16.
- Galiano, L. (2008). Site-directed spin labeling EPR studies of conformational dynamics in the flap region of HIV-1 protease Department of Chemistry. Gainesville, University of Florida. Ph.D.: **229**.
- Galiano, L., M. E. Blackburn, et al. (2009). Solute effects on spin labels at an aqueous-exposed site in the flap region of HIV-1 protease. *J Phys Chem B* **113**: 1673-80.
- Galiano, L., M. Bonora, et al. (2007). Inter-flap distances in HIV-1 Protease determined by pulsed EPR measurements. *J Am Chem Soc* **129**: 11004-5.
- Galiano, L., F. Ding, et al. (2009). Drug pressure selected mutations in HIV-1 protease alter flap conformations. *J Am Chem Soc* **131**: 430-1.
- Giehl, A., Lemm, T., Bartelsen, O., Sandhoff, K., and Blume, A. (1999) Interaction of the GM2-activator protein with phospholipid-ganglioside bilayer membranes and with monolayers at the air-water interface. *Eur J Biochem* **261**: 650-658.
- Gilbert, P. B. (2003). Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. *Statistics in Medicine* **22**: 573-593.
- Gorman, J. J., T. Wallis, et al. (2002). Protein disulfide bond determination by mass spectrometry. *J. Mass Spectrom. Rev* **21**: 183– 216.
- Gravel, R., Clarke, J., Kaback, M., Mahuran, D., Sandhoff, K., and Suzuki, K (1995) *The GM2 Gangliosidoses*, **7 ed.**, McGraw-Hill Inc., New York.
- Griffith, H. and H. M. McConnell (1965). A nitroxide-maleimide spin label. *PNAS* **55**: 8-11.
- Griffiths, G., B. Hoflack, et al. (1988). The mannose 6-phosphate receptor and the biogenesis of lysosomes. *Cell* **52**: 329-41.
- Griffiths, W.J., P. Jonsson, et al. (2001). Gas phase conformation can have an influence on peptide fragmentation. *Proteomics* **1**: 934–945.
- Goodenow, M.M., G. Bloom, et al. (2002). Naturally occurring amino acid polymorphisms in human immunodeficiency virus type 1 (HIV-1) Gag NC p7 and the C-cleavage site impact GagPol processing by HIV-1 protease. *Virology* **292**: 137-149.
- Gross, A., Columbus, L., Hideg, K., Altenbach, C., and Hubbell, W. L. (1999) Structure of the KcsA potassium channel from *Streptomyces lividans*: a site-directed spin labeling study of the second transmembrane segment. *Biochemistry* **38**: 10324-10335.

- Guo, Z., D. Cascio, et al. (2008). Structural determinants of nitroxide motion in spin-labeled proteins: solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci* **17**: 228-39.
- Hakansson, K., M. Chalmers, et al. (2003). Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem* **75**: 3256–3262.
- Hansen, P. C. (1998). *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, SIAM.
- Heaslet, H., R. Rosenfeld, et al. (2007). Conformational flexibility in the flap domains of ligand-free HIV protease. *Acta Crystallogr D Biol Crystallogr* **63**: 866-75.
- Hendrickson, C.L., J. Drader, et al. (1996). Fourier transform ion cyclotron resonance mass spectrometry in a 20 T resistive magnet. *S Rapid Commun. Mass Spectrom.* **10**: 1829– 1832.
- Ho, M. W. and J. S. O'Brien (1971). Gaucher's disease: deficiency of 'acid' -glucosidase and reconstitution of enzyme activity in vitro. *PNAS* **68**: 2810-3.
- Ho, S., R. Coman, et al. (2008). Drug-associated changes in amino acid residues in Gag p2, p7(NC), and p6(Gag)/p6(Pol) in human immunodeficiency virus type 1 (HIV-1) display a dominant effect on replicative fitness and drug response. *Virology* **378**: 272-281.
- Hoffman, E. d. and V. Stroobant (2007). *Mass spectrometry: principles and applications*. West Sussex, Jon Wiley and Sons Ltd.
- Hornak, V., A. Okur, et al. (2006). HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. *J Am Chem Soc* **128**: 2812-2813.
- Hornak, V., A. Okur, et al. (2006). HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *PNAS* **103**: 915-20.
- Hubbell, W. L., D. S. Cafiso, et al. (2000). Identifying conformational changes with site-directed spin labeling. *Nat Struct Biol* **7**: 735-739.
- Hubbell, W. L., A. Gross, et al. (1998). Recent advances in site-directed spin labeling of proteins. *Curr Opin Struct Biol* **8**: 649-656.
- Hubbell, W. L., H. S. McHaourab, et al. (1996). Watching proteins move using site-directed spin labeling. *Structure* **4**: 779-83.
- Hubbell, W. L. C., D. S.; Altenbach, C. (2000). Identifying conformational changes with site-directed spin labeling. *Nat Struct Biol* **7**: 735-739.

- Ishima, R., D. I. Freedberg, et al. (1999). Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. *Structure* **7**: 1047-1055.
- Ishima, R., R. Ghirlando, et al. (2001). Folded monomer of HIV-1 protease. *J Biol Chem* **276**: 49110-49116.
- Ishima, R., D. A. Torchia, et al. (2003). Solution structure of the mature HIV-1 protease monomer: insight into the tertiary fold and stability of a precursor. *J Biol Chem* **278**: 43311-43319.
- Jeschke, G. (2002). Distance Measurements in the Nanometer Range by Pulse EPR. *Chem Phys Chem* **3**: 927-932.
- Jeschke, G., V. Chechik, et al. (2006). DeerAnalysis2006 - A Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl Mag Reson* **30**: 473-498.
- Jeschke, G. and Y. Polyhach (2007). Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical Chemistry Chemical Physics* **9**(16): 1895-1910.
- Jeschke, G. K., A.; Jonas, U.; Godt, A. (2002). Direct Conversion of EPR Dipolar Time Evolution Data to Distance Distributions. *J Magn Reson* **155**: 72-82.
- Jeschke, G. P., H.; Zimmerman, H.; Godt, A. (2004). Sensitivity Enhancement in Pulse EPR Distance Measurements. *J Magn Reson* **169**: 1-12.
- Kantor, R., R. W. Shafer, et al. (2005). The HIV-1 Non-subtype B Workgroup: an international collaboration for the collection and analysis of HIV-1 non-subtype B data. *MedGenMed* **7**: 71.
- Karas, M., D. Bachman, et al. (1987). Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. *Int J Mass Spectrom Ion Proc* **78**: 53-68.
- Karas, M., D. Bachmann, et al. (1985). Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal Chem* **57**: 2935-2939.
- Kear, J. L., M. E. Blackburn, et al. (2009). Subtype polymorphisms among HIV-1 protease variants confer altered flap conformations and flexibility. *J Am Chem Soc* **131**: 14650-14651.
- Kolter, T., and Sandhoff, K. (2005) Principles of lysosomal membrane digestion: stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids. *Annu Rev Cell Dev Biol* **21**: 81-103.
- Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy*. New York, Springer Science and Business Media.

- Langen, R., K. J. Oh, et al. (2000). Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry* **39**: 8396-8405.
- Lawn, S. D. (2004). AIDS in Africa: the impact of coinfections on the pathogenesis of HIV-1 infection. *J Infect Dis* **48**: 1-12.
- Lemieux, M. J., B. L. Mark, et al. (2006). Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *J Mol Biol* **359**: 913-29.
- Li, S. C., H. Kihara, et al. (1985). Activator protein required for the enzymatic hydrolysis of cerebroside sulfate. Deficiency in urine of patients affected with cerebroside sulfatase activator deficiency and identity with activators for the enzymatic hydrolysis of GM1 ganglioside and globotriaosylceramide. *J Biol Chem* **260**: 1867-71.
- Liu, W., A. Brock, et al. (2007). Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* **4**: 239-44.
- Locatelli-Hoops, S., N. Remmel, et al. (2006). Saposin A mobilizes lipids from low cholesterol and high bis(monoacylglycerol)phosphate-containing membranes: patient variant Saposin A lacks lipid extraction capacity. *J Biol Chem* **281**: 32451-60.
- Louis, J. M., R. Ishima, et al. (2003). Revisiting monomeric HIV-1 protease. Characterization and redesign for improved properties. *J Biol Chem* **278**: 6085-6092.
- Low, M. G. and A. R. Saltiel (1988). Structural and functional roles of glycosylphosphatidylinositol in membranes. *Science* **239**: 268-75.
- Lundblad, R. L. (2004). *Chemical Reagents for Protein Modification*, 3rd ed.; CRC Press.
- Major, A. S., Joyce, S., and Van Kaer, L. (2006) Lipid metabolism, atherogenesis and CD1-restricted antigen presentation. *Trends Mol Med* **12**: 270-278.
- Mathias, J. D., Ran, Y., Carter, J. D., and Fanucci, G. E. (2009) Interactions of the GM2 Activator Protein with Phosphatidylcholine Bilayers: A Site Directed Spin Labeling Power Saturation Study. *Biophysical Journal* **97**: 1436-1444.
- Matsuda, J., M. Kido, et al. (2004). Mutation in saposin D domain of sphingolipid activator protein gene causes urinary system defects and cerebellar Purkinje cell degeneration with accumulation of hydroxy fatty acid-containing ceramide in mouse. *Hum Mol Genet* **13**: 2709-23.

- Matsuda, J., M. T. Vanier, et al. (2001). A mutation in the saposin A domain of the sphingolipid activator protein (prosaposin) gene results in a late-onset, chronic form of globoid cell leukodystrophy in the mouse. *Hum Mol Genet* **10**: 1191-9.
- Mahuran, D. J. (1998) The GM2 activator protein, its roles as a co-factor in GM2 hydrolysis and as a general glycolipid transport protein. *Biochim Biophys Acta* **1393**: 1-18.
- McHaourab, H. S., T. Kalai, et al. (1999). Motion of spin-labeled side chains in T4 lysozyme: effect of side chain structure. *Biochemistry* **38**: 2947-55.
- McHaourab, H. S., M. A. Lietzow, et al. (1996). Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**: 7692-7704.
- McHaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. (1996). Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**: 7692-7704.
- Mehl, E. and H. Jatzkewitz (1964). [A cerebroside-sulfatase from swine kidney]. *Hoppe Seylers Z Physiol Chem* **339**: 260-76.
- Miller, M., J. Schneider, et al. (1989). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**: 1149-1152.
- Mollaaghababa, R., H. J. Steinhoff, et al. (2000). Time-resolved site-directed spin-labeling studies of bacteriorhodopsin: loop-specific conformational changes in M. *Biochemistry* **39**: 1120-7.
- Morimoto, S., B. M. Martin, et al. (1989). Saposin A: second cerebroside activator protein. *PNAS* **86**: 3389-93.
- Nair, S.S., C. Nilsson, et al. (2006). De novo sequencing and disulfide mapping of a bromotryptophan-containing conotoxin by Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* **78**: 8082–8088.
- Noble, R. (2009). HIV types, subtypes, groups, and strains. 2010, from <http://www.avert.org/hiv-types.htm>.
- O'Brien, J. S. and Y. Kishimoto (1991). Saposin proteins: structure, function, and role in human lysosomal storage disorders. *FASEB J* **5**: 301-8.
- O'Brien, J. S., K. A. Kretz, et al. (1988). Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic locus. *Science* **241**: 1098-101.
- Osmanov, S., C. Pattou, et al. (2002). Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *Acquir Immune Defic Syndr* **29**: 184-190.

- Pannier, M., S. Veit, et al. (2000). Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn Reson* **142**: 331-40.
- Pasa-Tolic, Y.L., C. Masselon, et al. (2004). Proteomic analyses using an accurate mass and time tag strategy. *BioTechniques* **37**: 621–636.
- Perrin, L., L. Kaiser, et al. (2003). Travel and the spread of HIV-1 genetic variants. *Lancet Infect Dis* **3**: 22-7.
- Poole, C. P., Jr. (1983). *Electron Spin Resonance. A Comprehensive Treatise on Experimental Techniques, Second Edition*. Mineola, Dover Publications, Inc.
- Popolo, L., E. Ragni, et al. (2008). Disulfide bond structure and domain organization of yeast beta(1,3)-glucanosyltransferases involved in cell wall biogenesis. *J Biol Chem* **283**: 18553–18565.
- Qin, P. Z., J. Feigon, et al. (2005). Site-directed spin labeling studies reveal solution conformational changes in a GAAA tetraloop receptor upon Mg(2+)-dependent docking of a GAAA tetraloop. *J Mol Biol* **351**: 1-8.
- Rabenstein, M. D., and Shin, Y. K. (1995) Determination of the distance between two spin labels attached to a macromolecule. *PNAS* **92**: 8239-8243.
- Rambaut, A., D. Posada, et al. (2004). The causes and consequences of HIV evolution. *Nat Rev Genet* **5**: 52-61.
- Ran, Y., and Fanucci, G. E. (2008) A dansyl fluorescence-based assay for monitoring kinetics of lipid extraction and transfer. *Anal Biochem* **382**: 132-134.
- Ran, Y., Fanucci, G.E. (2009) Ligand Extraction Properties the GM2 Activator Protein and its Interactions with Lipid Vesicles. *Biophysical Journal* **97**: 257-266.
- Reeves, J. D. and R. W. Dorns (2002). Human immunodeficiency virus type 2. *J Gen Virol* **83**: 1253-1265.
- Rigat, B., Wang, W., Leung, A., and Mahuran, D. J. (1997) Two mechanisms for the recapture of extracellular GM2 activator protein: evidence for a major secretory form of the protein. *Biochemistry* **36**: 8325-8331.
- Rigat, B., Yeger, H., Shehnaz, D., and Mahuran, D. (2009) GM2 activator protein inhibits platelet activating factor signaling in rats. *Biochem Biophys Res Commun* **385**: 576-580.
- Robertson, D. L., B. H. Hahn, et al. (1995). Recombination in AIDS viruses. *J Mol Evol* **40**: 249-259.

- Rose, R.B., C. Craik, et al. (1998). Stroud, R.M. Domain flexibility in retroviral proteases: structural implications for drug resistant mutations. *Biochemistry* **37**: 2607-2621.
- Rossmann, M., R. Schultz-Heienbrok, et al. (2008). Crystal structures of human saposins C and D: implications for lipid recognition and membrane interactions. *Structure* **16**: 809-17.
- Sandhoff, K. and T. Kolter (1995). [Glycolipids of the cell surface--biochemistry of their decomposition]. *Naturwissenschaften* **82**: 403-13.
- Sandhoff, K. and T. Kolter (1996). Topology of glycosphingolipid degradation. *Trends Cell Biol* **6**: 98-103.
- Schaub, T. M., C. Hendrickson, et al. (2008). High-performance mass spectrometry: Fourier transform ion cyclotron resonance at 14.5 Tesla. *Anal Chem* **80**: 3985-3990.
- Schlote, W., K. Harzer, et al. (1991). Sphingolipid activator protein 1 deficiency in metachromatic leucodystrophy with normal arylsulphatase A activity. A clinical, morphological, biochemical, and immunological study. *Eur J Pediatr* **150**: 584-91.
- Schneider, M. F., S. J. Gange, et al. (2005). Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy. *AIDS* **19**: 2009-2018.
- Schwartz, J.C., M. Senko, (2002). A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* **13**: 659-669.
- Schweiger, A. and G. Jeschke (2001). *Principles of pulse electron paramagnetic resonance*. New York, Oxford University Press, Inc.
- Scott, W. and C. Schiffer, (2000). Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. *Structure* **8**: 1259-65.
- Senko, M.W., C. Hendrickson, et al. (1996). A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* **13**: 659-669.
- Senko, M.W., C. Hendrickson, et al. (1997). External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* **8**: 970-976.
- Smith, J. A. and R. Daniel (2006). Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses. *ACS Chem Biol* **1**: 217 - 226.
- Smock, R. G., and Gierasch, L. M. (2009) Sending signals dynamically. *Science* **324**: 198-203.

- Sparkman, O. D. (2000). *Mass spectrometry desk reference*. Pittsburgh, Global View Publishing.
- Spiegel, R., G. Bach, et al. (2005). A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: first report of saposin A deficiency in humans. *Mol Genet Metab* **84**: 160-6.
- Stone, T. L., T. Buckman, et al. (1965). Spin-labeled biomolecules. *PNAS* **54**: 1010-1017.
- Stults, C. L., C. C. Sweeley, et al. (1989). Glycosphingolipids: structure, biological source, and properties. *Methods Enzymol* **179**: 167-214.
- Sun, Y., D. P. Witte, et al. (2008). Neurological deficits and glycosphingolipid accumulation in saposin B deficient mice. *Hum Mol Genet* **17**: 2345-56.
- Syka, J. E., J. Coon, et al. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *PNAS* **101**: 9528–9533.
- Tanaka, K., H. Waki, et al. (1988). Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of flight Mass Spectrometry. *Rapid Commun Mass Spectrom* **2**: 151-153.
- Tang, X.J., P. Thibault, et al. (1993). Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation. *Anal Chem* **65**: 2824–2834.
- Thomson, J. J. (1913). Rays of positive electricity. *Proceedings of the Royal Society* **A89**: 1-20.
- Thomson, M. M., L. Perez-Alvarez, et al. (2002). Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect Dis* **2**: 461-471.
- Tikhonov, A. N. (1943). On the stability of inverse problems. *Dokl. Akad. Nauk USSR* **39**: 195-198.
- Tipton, J. D., J. D. Carter, et al. (2009). Sequential proteolysis and high-field FTICR MS to determine disulfide connectivity and 4-maleimide TEMPO spin-label location in L126C GM2 activator protein. *Anal Chem* **81**: 7611-7.
- Todd, M. J. and E. Freire (1999). The effect of inhibitor binding on the structural stability and cooperativity of the HIV-1 protease. *Proteins* **36**: 147-156.
- Todd, M. J., I. Luque, et al. (2000). Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. *Biochemistry* **39**: 11876-11883.

- Todd, M. J., N. Semo, et al. (1998). The structural stability of the HIV-1 protease. *J Mol Biol* **283**: 475-488.
- Torbeev, V. Y., K. Mandal, et al. (2008). Crystal structure of chemically synthesized HIV-1 protease and a ketomethylene isostere inhibitor based on the p2/NC cleavage site. *Bioorg Med Chem Lett* **18**: 4554-4557.
- Torbeev, V. Y., H. Raghuraman, et al. (2008). Dynamics of Flap Structures in Three HIV-1 Protease/Inhibitor Complexes Probes by Total Chemical Synthesis and Pulse-EPR Spectroscopy. *J Am Chem Soc* **131**: 884-885.
- Tsaprailli, G., H. Hair, et al. (1999). Influence of secondary structure on the fragmentation of protonated peptides. *J Am Chem Soc* **121**: 5142-5154.
- Tsybin, Y.L., C. Hendrockson, et al. (2006). Impact of Ion Magnetron Motion on Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Int J Mass Spectrom* **126**: 144-149.
- Uggerud, E. (2004). Electron capture dissociation of the disulfide bond-a quantum chemical model study. *Int J Mass Spectrom* **234**: 45-50.
- Velazquez-Campoy, A., M. J. Todd, et al. (2000). HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry* **39**: 2201-2207.
- Wang, L., Z. Zhang, et al. (2003). Addition of the keto functional group to the genetic code of Escherichia coli. *PNAS* **100**: 56-61.
- Wendeler, M., Werth, N., Maier, T., Schwarzmann, G., Kolter, T., Schoeniger, M., Hoffmann, D., Lemm, T., Saenger, W., and Sandhoff, K. (2006) The enzyme-binding region of human GM2-activator protein. *FEBS J* **273**: 982-991.
- Weil, J. A., J. R. Bolton, et al. (1972). *Electron Spin Resonance: Elementary Theory and Practical Applications*, John Wiley and Sons, Inc.
- Weiss, R. A. (1993). How does HIV cause AIDS? *Science* **260**: 1273 - 1279.
- Wenger, D. A., G. DeGala, et al. (1989). Clinical, pathological, and biochemical studies on an infantile case of sulfatide/GM1 activator protein deficiency. *Am J Med Genet* **33**: 255-65.
- Wilkening, G., T. Linke, et al. (2000). Degradation of membrane-bound ganglioside GM1. Stimulation by bis(monoacylglycero)phosphate and the activator proteins SAP-B and GM2-AP. *J Biol Chem* **275**: 35814-9.
- Wright, C. S., S. C. Li, et al. (2000). Crystal structure of human GM2-activator protein with a novel beta-cup topology. *J Mol Biol* **304**: 411-22.

- Wright, C. S., Mi, L. Z., and Rastinejad, F. (2004) Evidence for lipid packaging in the crystal structure of the GM2-activator complex with platelet activating factor. *J Mol Biol* **342**: 585-592.
- Wright, C. S., Mi, L. Z., Lee, S., and Rastinejad, F. (2005) Crystal structure analysis of phosphatidylcholine-GM2-activator product complexes: evidence for hydrolase activity. *Biochemistry* **44**: 13510-13521.
- Wright, C. S., Zhao, Q., and Rastinejad, F. (2003) Structural analysis of lipid complexes of GM2-activator protein. *J Mol Biol* **331**: 951-964.
- Xu, H., M. Freitas, et al. (2008). Identification and characterization of disulfide bonds in proteins and peptides from tandem MS data by use of the MassMatrix MS/MS search engine. *J Proteome Res* **7**: 138-144.
- Xu, Q., Kim, M., Ho, K. W., Lachowicz, P., Fanucci, G. E., and Cafiso, D. S. (2008) Membrane hydrocarbon thickness modulates the dynamics of a membrane transport protein. *Biophys J* **95**: 2849-2858.
- Yen, T.Y., H. Yan, et al. (2002). Characterizing closely spaced, complex disulfide bond patterns in peptides and proteins by liquid chromatography/electrospray ionization tandem mass spectrometry. *J Mass Spectrom* **37**: 15–30.
- Zhang, S., C. Van Pelt, et al. (2003). Automated chip-based nanoelectrospray-mass spectrometry for rapid identification of proteins separated by two-dimensional gel electrophoresis. *Electrophoresis* **24**: 3620–3632.
- Zhou, D., Cantu, C., 3rd, Sagiv, Y., Schrantz, N., Kulkarni, A. B., Qi, X., Mahuran, D. J., Morales, C. R., Grabowski, G. A., Benlagha, K., Savage, P., Bendelac, A., and Teyton, L. (2004) Editing of CD1d-bound lipid antigens by endosomal lipid transfer proteins. *Science* **303**: 523-527.
- Zubarev, R. A. (2004). Electron-capture dissociation tandem mass spectrometry. *Curr Opin Biotechnol* **15**: 12–16.
- Zubarev, R. A., N. Kruger, (1999). Electron capture dissociation of gases multiply-charged proteins is favored at disulfide bonds and other sites of hydrogen atom affinity. *J Am Chem Soc* **121**: 250-258.

BIOGRAPHICAL SKETCH

Jeffrey David Carter was born in 1983 and graduated from high school in 2002. Then pursued a biochemistry degree at Virginia Tech and graduated with a Bachelor of Science in 2006. He was admitted to the Department of Chemistry graduate program at the University of Florida in 2006, where he joined the research group of Dr. Gail E. Fanucci. He obtained his Ph.D. in chemistry in 2012.