

**STRUCTURAL AND DYNAMIC INVESTIGATION OF HIV-1 CAPSID  
ASSEMBLIES AND CAPSID MATURATION BY MAGIC ANGLE SPINNING  
NMR SPECTROSCOPY**

by

Mingzhang Wang

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Summer 2018

© 2018 Mingzhang Wang  
All Rights Reserved

**STRUCTURAL AND DYNAMIC INVESTIGATION OF HIV-1 CAPSID  
ASSEMBLIES AND CAPSID MATURATION BY MAGIC ANGLE SPINNING  
NMR SPECTROSCOPY**

by

Mingzhang Wang

Approved: \_\_\_\_\_

Brian J. Bahnson, Ph.D.  
Chair of the Department of Chemistry and Biochemistry

Approved: \_\_\_\_\_

George H. Watson, Ph.D.  
Dean of the College of Arts and Sciences

Approved: \_\_\_\_\_

Ann L. Ardis, Ph.D.  
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Tatyana Polenova, Ph.D.  
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Cecil Dybowski, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Murray V. Johnston, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

In-Ja Byeon, Ph.D.  
Member of dissertation committee

## ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest appreciation to my adviser, Professor Tatyana Polenova, who provided constructive suggestion, supervision, patient help and continuous support over the years of my graduate study. Her critical thinking in research guided me to be a professional scientist. I learnt much from her broad knowledge and inimitable passion for work, which will always illuminate my future life and career.

I would like to thank all my committee members for their time and constructive input during my doctoral studies. To Professor Cecil Dybowski and Professor Murray Johnston from the University of Delaware, I am grateful for their guiding and encouraging me to explore analytical methods in my research project. To Dr. In-Ja Byeon from the University of Pittsburgh who is also one of our collaborators, it is with my sincerest appreciation for her constructive insights and suggestions into this project.

I want to thank all of our collaborators on different projects, Professor Angela Gronenborn, Dr. In-Ja Byeon, Professor Jinwoo Ahn, Professor Peijun Zhang from the University of Pittsburgh, Dr. Eric Freed from National Cancer Institute, Professor Chris Aiken from the Vanderbilt University, Professor Juan Perilla from the University of Delaware, Professor Klaus Schulten from the University of Illinois, and Dr. Jochem Struppe from Bruker Biospin.

My thanks are also extended to all my past and current colleagues in the Polenova group for all their help, support and friendship. Particularly, I want to thank

Dr. Christopher Suiter, Dr. Si Yan and Dr. Manman Lu for teaching me a lot of biochemistry skills and NMR data processing. I also want to thank Dr. Guangjin Hou and Xingyu Lu for teaching me many aspects of NMR theory and experimental set-up. To Dr. Caitlin Quinn and Dr. Huilan Zhang, many thanks for their helpful idea and support in many collaborated projects. To Dr. Mingyue Li, Dr. Jenna Yehl, Dr. Changmiao Guo, Luke Onisk, Chunting Zhang, and Vera Vladimirova, I am thankful for their warmest company and help over the years of my life in the lab. Thanks to Matthew Fritz and Jodi Kraus for sharing the thoughts of life and the spirit of coffee with me.

I would like to thank Dr. Steve Bai for his assistance and support about solution NMR experiments and training. I am thankful for Shannon Modla at the DBI Bioimaging center for helping me to acquire TEM images of my research samples.

I am extremely grateful to my friends at the University of Delaware, Yu-ting Hung, Baxter Abraham, Sandy Hou, Yichen Duan, Mackenzie Williams, Jing Zhang for all their constant support, warmest company, and cheerful encouragement.

Finally, I dedicate this work to my family, my grandparents (who inspired me to become a chemist), parents, auntie Linda Chen, uncle Eddie Luo, cousin Emily Luo, and especially my husband Dr. Jiasheng Qian for their endless love, encouragement and support throughout my graduate studies.

This thesis is dedicated to all of them.

## TABLE OF CONTENTS

LIST OF TABLES .....	x
LIST OF FIGURES.....	xii
LIST OF ABBREVIATIONS .....	xxii
ABSTRACT .....	xxvi
Chapter	
1 INTRODUCTION TO HIV-1 STRUCTURAL BIOLOGY: MATURATION AND CAPSID ASSEMBLY .....	1
1.1 AIDS Epidemic and HIV .....	1
1.2 HIV-1 Viral Life Cycle.....	5
1.3 HIV-1 CA Assembly.....	7
1.4 HIV-1 Maturation and Maturation Inhibitors .....	9
1.4.1 Gag Polyprotein Cleavage and Maturation Intermediate CA- SP1 .....	9
1.4.2 Maturation Inhibitors .....	11
1.4.3 T8I Mutation in Spacer Peptide 1 (SP1) .....	12
REFERENCES.....	14
2 STRUCTURE AND DYNAMICS OF PROTEIN ASSEMBLIES BY MAGIC ANGLE SPINNING SOLID-STATE NMR.....	21
2.1 Introduction.....	21
2.2 Magic Angle Spinning NMR Spectroscopy.....	22
2.3 Resonance Assignments.....	23
2.3.1 Isotopic Labeling Schemes .....	24
2.3.2 Multidimensional NMR Experiments .....	25
2.3.3 Secondary Structure Analysis .....	28
2.4 Protein Dynamics by MAS NMR .....	29
2.4.1 NMR Timescale of Protein Motions .....	30

2.4.2	Experiments for Probing Motions on Milli- to Microsecond Timescales.....	32
2.4.3	Experiments for Probing Motions on Micro- to Nanosecond Timescales.....	33
2.5	Dynamic Nuclear Polarization Technique .....	34
REFERENCES.....		36
<b>3</b>	<b>STRUCTURE CHARACTERIZATION OF HIV-1 CA TUBULAR AND SPHERICAL ASSEMBLIES BY MAS NMR .....</b>	<b>46</b>
3.1	Introduction.....	46
3.2	Methods and Experiments.....	48
3.2.1	Materials.....	48
3.2.2	Expression and Purification of CA Cross-Linked Hexamer and Pentamer Proteins.....	49
3.2.3	Preparation of Cross-Linked CA Hexamer and Pentamer Assemblies .....	51
3.2.4	The Reversible Assembly of Cross-Linked CA Hexamers and the Effect of NaCl .....	53
3.2.5	Preparation of CA Cross-Linked Hexamer/Pentamer Co-Assemblies .....	55
3.2.6	Transmission Electron Microscopy (TEM) .....	57
3.2.7	MAS NMR Experiments.....	58
3.3	Results.....	59
3.3.1	Resonance Assignments of CA Cross-linked Tubular Assemblies .....	59
3.3.2	The Effect of Cross-Linking on Structure of CA Tubular Assemblies .....	69
3.3.3	The Effect of Morphology on MAS NMR Spectra: Pentamers vs. Hexamers.....	71
3.4	Conclusions and Future Outlook.....	74
REFERENCES.....		76
<b>4</b>	<b>INVESTIGATION OF HIV-1 MATURATION .....</b>	<b>79</b>
4.1	Introduction.....	79
4.2	Methods and Experiments.....	80

4.2.1	Materials.....	80
4.2.2	Expression and Purification of CA-SP1 Proteins.....	80
4.2.3	Transmission Electron Microscopy (TEM) .....	81
4.2.4	MAS NMR Experiments.....	83
4.2.5	Simulations of NMR Dipolar Lineshapes and Calculations of Dipolar Order Parameters .....	84
4.2.6	Molecular Dynamics (MD) Simulation Trajectories of CA-SP1 Assemblies .....	85
4.2.7	Calculations of Chemical Shifts from MD Trajectories .....	85
4.3	Results.....	85
4.3.1	Resonance Assignments of CA-SP1(T8I) Assemblies .....	85
4.3.2	Conformation of SP1 in the Wild Type and T8I Mutant .....	93
4.3.3	Allosteric Conformational and Dynamic Changes in CA Induced by T8I Mutant.....	97
4.3.4	Investigation of Maturation by an Integrated MD and MAS NMR Approach.....	104
4.3.4.1	Conformation of SP1 is Modulated by T8I Mutation in MD Simulations .....	104
4.3.4.2	Quantitative Analysis of Isotropic Chemical Shift: MD vs. MAS NMR.....	109
4.3.5	The Role of Dynamics in Capsid Maturation and Maturation Inhibition.....	114
4.4	Conclusions.....	115
	REFERENCES.....	117
5	DEVELOPMENTS AND APPLICATIONS OF <sup>19</sup> F MAS NMR.....	121
5.1	Introduction.....	121
5.2	<sup>19</sup> F MAS NMR of Fluorosubstituted Tryptophans .....	122
5.2.1	Methods and Experiments.....	123
5.2.1.1	Materials	123
5.2.1.2	<sup>19</sup> F MAS NMR Experiments.....	123
5.2.1.3	Determination of Chemical Shift Tensors.....	126
5.2.2	MAS and Magnetic Field Dependence of Sensitivity and Resolution .....	126

5.2.3	Chemical Shift Tensor Measurements .....	132
5.3	<sup>19</sup> F MAS NMR of HIV-1 CA Protein Assemblies.....	138
5.3.1	Methods and Experiments.....	138
5.3.1.1	Materials .....	138
5.3.1.2	Sample Preparation of Fluorinated CA Protein .....	138
5.3.1.3	Transmission Electron Microscopy (TEM) .....	140
5.3.1.4	MAS NMR Experiments.....	141
5.3.2	Introduction of <sup>19</sup> F into HIV-1 CA Protein.....	143
5.3.3	Resonance Assignments of Fluorine Sites .....	148
5.3.4	Chemical Shift Anisotropy and Relaxation Time Measurements .....	155
5.3.5	<sup>19</sup> F- <sup>19</sup> F Correlation Experiments .....	161
5.4	Conclusions and Future Outlook.....	165
	REFERENCES.....	167
6	CONCLUSIONS.....	172
Appendix		
A	COPYRIGHT LETTER.....	175

## LIST OF TABLES

Table 3.1	Resonance assignments for CA cross-linked hexamer assemblies, including backbone N, Co, C $\alpha$ , sidechain C $\beta$ , C $\gamma$ /C $\gamma$ 2, C $\delta$ /C $\delta$ 2, C $\epsilon$ , and C $\zeta$ . CSI 2.0 <sup>13</sup> secondary structure prediction: C-coils, H-helix, and B-sheets. ....	62
Table 4.1	Chemical shifts of the CTD-SP1 region (CA residues 221-231, SP1 residues 1-14) of CA-SP1(T8I) assemblies determined by MAS NMR experiments. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.....	92
Table 4.2	CSI 2.0 <sup>21</sup> and TALOS-N <sup>22</sup> secondary structure prediction based on experimental chemical shifts for CA-SP1(T8I) NL4-3, CA(A92E)-SP1 NL4-3, and CA-SP1 HXB2. Designations for C-coil and H-helix are as described in Materials and Methods. "--" indicates predictions that are less reliable due to missing chemical shifts. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.....	96
Table 4.3	Secondary structure predictions of the NTD determined from MAS NMR experiments in CA-SP1 assemblies. CSI 2.0 <sup>28</sup> and TALOS-N <sup>22</sup> programs were used to predict secondary structures by CSI and torsion angles, respectively. "--" The predictions are less reliable due to the missing chemical shifts. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. ....	102
Table 4.4	An analysis of chemical shift (CS) RMSD values, comparing experimental shifts of MAS NMR with calculated shifts of MD simulations, as well as X-ray structure (PDB: 5I4T). The CS associated with the residues of CA(V221)-SP1 in CA(A92E)-SP1 and CA-SP1(T8I) assemblies were used for the comparison. ....	113
Table 5.1	MAS frequency and magnetic field dependence of the sensitivity of <sup>19</sup> F MAS NMR experiments for the 5F-L/DL-tryptophan and 6F-L/DL-tryptophan. The figure was originally published in REFERENCES <sup>15</sup> . Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. ....	131

Table 5.2	MAS NMR experimental principal components of the chemical shift tensor, isotropic chemical shifts, reduced anisotropy, and asymmetry for the F-DL-tryptophans and F-L-tryptophans. The experimental reduced anisotropy of 4F-DL-tryptophan is -62.8 ppm; for asymmetries close to 1, the sign can be inverted, indicating a symmetric tensor. ....	135
Table 5.3	Summary of MAS NMR experimental principal components of the chemical shift tensors, isotropic chemical shifts, reduced anisotropy, and asymmetry for the F-DL-tryptophan acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES <sup>15</sup> . Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. ....	136
Table 5.4	Chemical shift perturbations detected in (H)NH HETCOR, (H)CH HETCOR, and <sup>13</sup> C- <sup>13</sup> C RFDR MAS NMR spectra of tubular assemblies of 5- <sup>19</sup> F-Trp, U- <sup>13</sup> C, <sup>15</sup> N CA vs. U- <sup>13</sup> C, <sup>15</sup> N CA. ....	146
Table 5.5	<sup>19</sup> F chemical shifts of HIV-1 CA protein in solution. ....	152
Table 5.6	MAS NMR experimental <sup>19</sup> F isotropic chemical shifts for the CA assemblies of WT and 5F-Trp-substituted mutant. All spectra were recorded at 19.96 T. ....	153
Table 5.7	Summary of MAS NMR experimental isotropic chemical shifts, reduced anisotropy, and asymmetry for the 5- <sup>19</sup> F-Trp CA wild type and mutants acquired at 11.74 T and 19.96 T. ....	158
Table 5.8	T <sub>1</sub> and T <sub>2</sub> relaxation time of 5- <sup>19</sup> F-Trp CA wild type (tubular assembly), and A14C/E45C/W184A/M185A mutant (cross-linked hexamer tubular assembly). ....	160
Table 5.9	<sup>19</sup> F- <sup>19</sup> F distances in HIV-1 CA WT and A14C/E45C/W184A/M185A protein. ....	163

## LIST OF FIGURES

Figure 1.1	Map of HIV/AIDS global situation and trends. Made by WHO. <a href="http://www.who.int/gho/hiv/en/">http://www.who.int/gho/hiv/en/</a> .....	2
Figure 1.2	Organization of the HIV-1 genome. <sup>3</sup> The total size of the HIV-1 genome is 9.7 kb. Each of the viral genes is depicted based on the relative position in the RNA genome. Dashed lines represent RNA splicing. The number in parenthesis is molecular weight of each protein. The figure was originally published in REFERENCES <sup>3</sup> , copyright by Nkeze et al., 2015. ....	4
Figure 1.3	HIV-1 immature and mature virions. Both cryo-electron tomography (Cryo-ET) images and schematic representations illustrate the architecture of immature and mature virions. The figure adapted from REFERENCES <sup>6,8</sup> . Permissions for reuse in this dissertation is granted by Springer Nature, copyright 2015, and American Society for Microbiology, copyright 2015. ....	5
Figure 1.4	Schematic of the HIV-1 life cycle. The figure was originally published in REFERENCES <sup>9</sup> Permission for reuse in this dissertation is granted by Springer Nature, copyright 2012.....	6
Figure 1.5	The all-atom model of HIV-1 CA conical lattice with the structures of pentamers and hexamers (a) and the structure of CA monomer (b). The structures of pentamer and hexamer are not native units of the cone, and they are cross-linked pentamer and hexamer of CA mutants. PDB ID: 3J3Y for CA cone model <sup>18</sup> , 4XFX for CA monomer, NTD in dark gray and CTD in light gray <sup>29</sup> , 3P05 for CA pentamer <sup>17</sup> , and 3H47 for CA hexamer <sup>30</sup> .....	8
Figure 1.6	A model of CA tube (NTD in gray and CTD in pink) PDB ID 3J4F <sup>18</sup> , and transmission electron microscope (TEM) images of CA WT tubular assemblies. <sup>24</sup> .....	9
Figure 1.7	Schematic diagram of the HIV-1 Gag sequential cleavage and virus maturation process. <sup>40</sup> The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. ....	10

Figure 1.8	Chemical structures of Bevirimat molecule (a) and PF-46396 molecule (b). .....	12
Figure 2.1	Schematic illustration of MAS NMR setup. The sample is packed into a rotor and placed into an NMR probe at the magic angle of 54.74° with respect to the magnetic field. The effect of MAS on NMR lineshapes for an I=1/2 nucleus has been depicted in a-c. The broad static <sup>13</sup> C powder pattern due to dipolar interaction (a) is split into a series of spinning sidebands (b). When the MAS frequency exceeds the magnitude of the anisotropic interaction, the interaction is averaged out into an isotropic peak (c). Figure is adapted from REFERENCES <sup>24</sup> . Permission for reuse in this dissertation is granted by American Chemical Society, copyright 2015. ....	23
Figure 2.2	Magnetization transfer pathways of homonuclear and heteronuclear NMR experiments, including DARR/CORD, NCA, NCO, NCACX, NCOCX, CANCO(CX) and CONCA. Atoms in residue (i) are labeled in orange, and the ones in the neighboring residue (i-1) are labeled in purple. Strong transfers are marked with a black arrow, while weaker transfers or transfers that require longer mixing time are marked with gray arrows and lightly colored atoms. ....	27
Figure 2.3	Timescales of protein dynamic processes and MAS NMR measurements of dynamics. The figure is adapted from REFERENCES <sup>65,68</sup> . Permissions for reuse in this dissertation is granted by Springer Nature, copyright 2007, and Elsevier B.V., copyright 2005. ....	32
Figure 2.4	Schematic of DNP transfer pathways in a system containing <sup>1</sup> H and <sup>13</sup> C. Green arrows distinguish between indirect and direct DNP transfer. <sup>85</sup> Permissions for reuse in this dissertation is granted by Elsevier B.V., copyright 2017.....	35
Figure 3.1	Structures of CA hexamer (A14C/E45C/W184A/M185A) and CA pentamer (N21C/A22C/W184A/M185A), NL4-3 strain. Mutation sites, cysteine (pink) and alanine (green), are labeled as sticks in 3D structures, showing cross-linked subunits. PDB: 3H4E and 3P05. <sup>4,5</sup> .....	48
Figure 3.2	Purification of monomer (25.6 kDa) CA A14C/E45C/W184A/M185A and N21C/A22C/W184A/M185A proteins assessed by SDS-PAGE gel.....	50

Figure 3.3	Oligomerization of CA cross-linked hexamer and pentamer assemblies were characterized by non-reducing SDS-PAGE gel with molecular weight (MW) marker. Lane 1 is CA cross-linked pentamer protein solution (no DTT); Lane 2 is CA cross-linked pentamer assembly (pelleted with 2.4 M NaCl); Lane 3 is CA cross-linked hexamer protein solution (no DTT); Lane 4 is CA cross-linked hexamer assembly (pelleted with 2.4 M NaCl). The mismatch of MW marker and protein is due to the fact the current MW marker was not calibrated for this large oligomer assay, and CA hexamer and pentamer run slower than the MW maker.....	52
Figure 3.4	(a) Schematic diagram of dialysis process to remove NaCl from CA cross-linked assemblies. (b) TEM images of CA hexamer illustrate the process of disassembly and re-assembly.....	54
Figure 3.5	Oligomerization of CA cross-linked hexamer assemblies was characterized by non-reducing SDS-PAGE gel with molecular weight (MW) marker. Lane 1 is CA hexamer protein solution (no DTT); Lane 2 is CA hexamer assembly pellet (2.4 M NaCl); Lane 3 is CA hexamer assembly pellet (1 M NaCl); Lane 4 is CA cross-linked hexamer protein solution after removing NaCl; Lane 5 is CA hexamer re-assembly pellet with 2.4 M NaCl. ....	55
Figure 3.6	(a) TEM images of CA hexamer/pentamer co-assemblies, molar ratio of 10:1 and 20:1. (b) Co-sedimentation assay of CA hexamer/pentamer (10:1) co-assemblies. T: total, P: pellet, S: supernatant. ....	56
Figure 3.7	TEM images of CA cross-linked hexamer tubes (a-b) and pentamer spheres (c-d) with 1 M and 2.4 M NaCl. ....	57
Figure 3.8	The amino acid sequence of CA A14C/E45C/W184A/M185A mutant. The residues that contain ambiguous backbone resonance assignments are labeled in gray. ....	60
Figure 3.9	2D CORD, NCACX and NCOCX MAS NMR spectra for CA cross-linked hexamer tubular assemblies, showing assignments for S109-T110. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz; the data were processed with 60° shifted sine bell apodization process in both dimensions. ....	61

Figure 3.10	Superposition of 2D NCA and CORD (aliphatic region) MAS NMR spectra for CA cross-linked hexamer (orange) and CA WT tubular assemblies (gray), illustrating the chemical shift perturbations and peak intensity differences. The perturbed resonances of CA WT (blue) and CA hexamer (black) are labeled. The residues that exhibit large perturbations are highlighted in orange on a 3D structure of CA monomer. Four mutations (C14, C45, A184, A185) are shown in purple sticks. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz, and processed with 30° shifted sine bell apodization in both dimensions.....	70
Figure 3.11	Chemical shift changes between CA and CA cross-linked hexamer: sum of $^{13}\text{C}\alpha$ , $^{15}\text{N}$ chemical shift perturbations (CSP). Note that a CSP > 1.2 ppm is considered a significant perturbation, while a CSP < 0.6 ppm (1*STD) is not considered major given the mutations as well as experimental and systematic error. ....	71
Figure 3.12	Structure of a HIV-1 conical capsid (a) and a tubular capsid (b), and their building blocks. TEM images and $^{13}\text{C}$ - $^{13}\text{C}$ correlation spectra (aliphatic region) of CA hexamer and pentamer co-assemblies (c, f), CA cross-linked tubular assemblies (d, g), and CA WT tubular assemblies (e, h). All the spectra were acquired at 20.0 T and the MAS frequency of 14 kHz; (f) were processed with Lorentz-to-Gauss (GM) 80 (gray) and Exponential-Multiply (EM) 100 (blue) in both dimensions, and (g, h) were process with 45° shifted sine bell apodization in both dimensions.....	73
Figure 3.13	Superposition of 2D CORD (aliphatic region) MAS NMR spectra for CA hexamer (gray) in tubular assemblies and CA cross-linked pentamer (cyan) in co-assemblies with CA cross-linked hexamer. The expansions around $\text{C}\alpha$ - $\text{C}\beta$ regions of Thr and Ala are shown to illustrate chemical shift perturbations. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz, processed with Lorentz-to-Gauss (GM) 80 in both dimensions. The spectra of CA cross-linked pentamer in co-assemblies with the CA cross-linked hexamer were acquired with non-uniformly sampling (NUS). ....	74
Figure 4.1	TEM images <sup>9</sup> of tubular assemblies of CA <sup>7</sup> and CA-SP1 <sup>5</sup> variants. ....	82

- Figure 4.2 Superposition of dipolar-based spectra of HIV-1 CA and CA-SP1(T8I) tubular assemblies.<sup>9</sup> 2D CORD (top) and NCACX (bottom) MAS NMR spectra for CA NL4-3 (gray) and CA-SP1(T8I) NL4-3 (orange) illustrate the multiple chemical shift perturbations and peak intensity differences. The residues that performed peak splitting are labeled in blue. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz, with 60° shifted sine bell apodization process in both dimensions. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License..... 87
- Figure 4.3 A stretch of sequential walk for the CTD-SP1 region (residues CA: G225-L231 to SP1: A1-M14) of CA-SP1(T8I) tubular assemblies.<sup>9</sup> These residues were assigned by a combination of various types of MAS NMR spectra, including 2D CORD (orange labels), 2D/3D(NUS) NCACX (brown/blue labels, respectively), 2D/3D NCOX (brown/blue labels, respectively), 3D CONCA (purple labels), and 2D direct-INADEQUATE (green labels). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License..... 91
- Figure 4.4 Superposition of 2D <sup>13</sup>C-<sup>13</sup>C correlation spectra for CA and CA-SP1 variants, highlighting SP1 resonances and residues exhibiting chemical shift or peak intensity changes. Left expansions: CA NL4-3 (orange) and CA-SP1(T8I) NL4-3 (gray). Right expansions: CA(A92E)-SP1 NL4-3 (red), CA-SP1 HXB2 (blue) and CA-SP1(T8I) NL4-3 (gray). Chemical shift perturbations are present in several residues of the CTD (V221)-SP1 tail, CypA loop, MHR, and NTD β-hairpin. All spectra were acquired at 20.0 T and an MAS frequency of 14 kHz. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License..... 94
- Figure 4.5 Superposition of 2D direct-INADEQUATE spectra for HIV-1 CA-SP1 variants: CA(A92E)-SP1 NL4-3 (green, black labels) and CA-SP1(T8I) NL4-3 (gray, orange labels), displaying the multiple chemical shift perturbations in residues of CypA loop and CTD (V221-) SP1 tail. The spectra of CA(A92E)-SP1 assemblies were acquired at 20.0 T, with 60° shifted sine bell apodization process in both dimensions. The spectra of CA-SP1(T8I) assemblies were recorded at 14.4 T, processed with 90° shifted sine bell apodization. Both of them were performed at MAS frequency of 14 kHz. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License..... 95

- Figure 4.6 Chemical shift changes between CA and CA-SP1 plotted along the linear amino acid sequence I, II, III: sum of  $^{13}\text{C}\alpha$ ,  $^{15}\text{N}$  chemical shift perturbations (CSP): CA NL4-3 and CA-SP1(T8I) NL4-3 (orange), CA(A92E) NL4-3 and CA(A92E)-SP1 NL4-3 (brown), CA HXB2 and CA-SP1 HXB2 (blue). Note that a CSP > 0.5 ppm is considered a significant perturbation, while a CSP < 0.3 ppm (dashed gray) is negligible within experimental and systematic error. IV: Plot (grey) of the  $^{13}\text{C}\alpha$  -  $^{15}\text{N}$  backbone peak intensity ratio (CA-SP1(T8I) NL4-3/CA NL4-3) vs. residue number (for non-overlapping peaks with resonance assignments in the 2D NCA MAS NMR spectra). Residues with peak intensity ratio > 1 indicate attenuated motions on micro- to millisecond timescales. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 98
- Figure 4.7 Comparison of nano-to-microsecond timescale dynamics for CypA loop residues observed in tubular assemblies of HIV-1 CA and CA-SP1 variants. (a) Experimental (solid black) and simulated (dashed blue)  $^1\text{H}$ - $^{15}\text{N}$  dipolar lineshapes for CypA loop residues in (top to bottom): CA NL4-3, CA NL4-3 A92E, CA-SP1 NL4-3 A92E and CA-SP1 NL4-3 T8I. (b)  $^1\text{H}$ - $^{15}\text{N}$  Dipolar order parameters are plotted vs. residue number for CypA loop residues in CA and CA-SP1 tubular assemblies. The CA-SP1(T8I) mutant exhibits the same order of magnitude attenuation of loop dynamics as previously observed in the CA(A92E) escape mutant<sup>7</sup>. .... 100
- Figure 4.8 Two typical MD simulation frames showing hexameric subunits of CTD-SP1 (a) and the CTD-SP1(T8I) (b). Stride plots of secondary structure for each chain along with the MD trajectories for CA CTD-SP1 WT (c), and CA CTD-SP1(T8I) mutant (d). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 106
- Figure 4.9 Helix/coil/sheet probability plots of the CTD (V221)-SP1 subdomain averaged over the MD trajectories: CA CTD-SP1 WT (left), and CA CTD-SP1(T8I) (right). The expanded scale (0-0.05; right hand side) is shown for the “Sheet” content. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 107

- Figure 4.10 Clustering analysis of the MD trajectory of CA(CTD)-SP1 WT (top) and CA(CTD)-SP1(T8I) mutant (bottom) identified six major sub-populations. Structure clusters were extracted from the MD trajectory. One chain is highlighted in blue to illustrate the differences in secondary structure present in the SP1 region. CTD(V221)-SP1 segment is colored in green. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 108
- Figure 4.11 Correlation of MAS NMR chemical shifts with SHIFTX2 predicted shifts from MD trajectory and X-ray structures (PDB: 5I4T): CA CTD-SP1 WT. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 111
- Figure 4.12 Correlation of MAS NMR chemical shifts with SHIFTX2 predicted shifts from MD trajectory and X-ray structures (PDB: 5I4T): CA CTD-SP1(T8I). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License. .... 112
- Figure 4.13 Schematic illustration of hypothesized mechanism: HIV-1 capsid maturation (top), and its inhibition by T8I mutation in SP1 or MI small-molecule. Involved elements are depicted, viral protease in orange sphere and MI in red star. One CA chain is highlighted in blue and CTD(V221)-SP1 segment is colored in green..... 115
- Figure 5.1 Chemical structures of tryptophan and various fluorotryptophan molecules (a) DL-Trp, (b) 4F-DL- Trp, (c) 5F-DL-Trp, (d) 6F-DL-Trp, (e) 7F-DL- Trp, (f) L- Trp, (g) 4F-L- Trp, (h) 5F-L- Trp, (i) 6F-L- Trp, and (j) 7F-L- Trp. .... 123
- Figure 5.2 (a-d) MAS frequency dependence of the linewidth and signal intensity of fluorosubstituted tryptophan crystalline solids: (a) 6F-L-tryptophan, (b) 6F-DL-tryptophan, (c) 5F-L-tryptophan, (d) 5F-DL-tryptophan. 1D <sup>19</sup>F spectra were acquired at 14.1 T (564.278 MHz <sup>19</sup>F Larmor frequency) and MAS frequencies of 20 kHz, 30 kHz, 35 kHz, 40 kHz, 50 kHz, and 60 kHz. For (e) 5F-DL-tryptophan, <sup>19</sup>F 1D spectra were also acquired at 19.96 T (800.095 MHz <sup>19</sup>F Larmor frequency) and a MAS frequency of 40 kHz, with (right spectrum) and without <sup>1</sup>H decoupling (left spectrum). The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. .... 128

Figure 5.3	Summary of $^{19}\text{F}$ MAS NMR spectra of fluorosubstituted tryptophans (4F-DL-tryptophan, 5F-DL-tryptophan, 5F-L-tryptophan) acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES <sup>15</sup> . Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. ....	129
Figure 5.4	Summary of $^{19}\text{F}$ MAS NMR spectra of fluorosubstituted tryptophans (6F-DL-tryptophan, 6F-L-tryptophan, 7F-DL-tryptophan) acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES <sup>15</sup> . Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. ....	130
Figure 5.5	$^{19}\text{F}$ MAS NMR spectra of fluorosubstituted tryptophans acquired at 9.4 T (376.476 MHz $^{19}\text{F}$ Larmor frequency) and a MAS frequency of 7 kHz: (a) 4F-DL-tryptophan, (b) 7F-DL-tryptophan, (c) 5F-L-tryptophan (d) 5F-DL-tryptophan, (e) 6F-L-tryptophan (f) 6F-DL-tryptophan. $^{19}\text{F}$ MAS NMR spectra of fluorosubstituted tryptophans acquired at 14.1 T ( $^{19}\text{F}$ Larmor frequency of 564.278 MHz) and a MAS frequency of 8 kHz: (g) 5F-L-tryptophan, (h) 5F-DL-tryptophan, (i) 6F-L-tryptophan, (j) 6F-DL-tryptophan. The peaks colored in light grey in (b) arise from the $^{19}\text{F}$ signal of the sample spacer. The figure was originally published in REFERENCES <sup>15</sup> . Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018. ....	134
Figure 5.5	Structures of the HIV-1 CA monomer (W23, W80, W117, W133 and W184 are shown in purple stick representation with fluorine atoms as yellow spheres), a section of the CA tube and the A14C/E45C/W184A/M185A CA cross-linked hexamer (NTD; grey, CTD; teal)....	144

- Figure 5.6 (a) 3D structure of an HIV-1 CA monomer showing the positions where 5-F-Trp was incorporated. Five 5-<sup>19</sup>F-Trp residues (W23, W80, W117, W133 and W184) are shown in purple sticks, and the fluorine atoms in yellow spheres. (b-d) 2D (H)NH HETCOR (b), (H)CH HETCOR (c) spectra, and <sup>13</sup>C-<sup>13</sup>C RFDR spectra in aliphatic region (d) of tubular assemblies of U-<sup>13</sup>C, <sup>15</sup>N CA (grey), and 5-<sup>19</sup>F-Trp, U-<sup>13</sup>C, <sup>15</sup>N CA (teal), acquired at 19.96 T; the MAS frequency of 60 kHz. Note the disappearance of resonances corresponding to the indole ring of Trp residues in the fluorinated sample. The residues for which large chemical shift perturbations were detected in the MAS NMR spectra (b-d) are shown as sticks. Nitrogen and hydrogen atoms are shown in blue and red, respectively..... 145
- Figure 5.7 <sup>19</sup>F solution NMR (a) and MAS NMR (b) spectra of 5-<sup>19</sup>F-Trp CA WT, fluorosubstituted CA mutants and CA cross-linked hexamer (A14C/E45C /W184A/M185A). The solution spectra were recorded at 14.1 T. The MAS spectra were recorded at 19.96 T and a MAS frequency of 40 kHz. Assignments for each fluorine site are shown in purple labels. .... 151
- Figure 5.8 TEM images of 5-<sup>19</sup>F-Trp CA WT and mutant assemblies, including 5-<sup>19</sup>F-Trp, U-<sup>15</sup>N CA tubes (a), 5-<sup>19</sup>F-Trp, U-<sup>15</sup>N W80Y tubes (c), 5-<sup>19</sup>F-Trp, U-<sup>15</sup>N W23I non-tube assemblies (b), 5-<sup>19</sup>F-Trp, U-<sup>13</sup>C, <sup>15</sup>N A14C/E45C/W184A/M185A tubes and soluble cross-linked hexamers (d)..... 152
- Figure 5.9 Comparison of <sup>19</sup>F MAS NMR spectra of 5-<sup>19</sup>F-Trp, U-<sup>15</sup>N CA tubular assemblies that were recorded at different magnetic fields and MAS frequencies. .... 155
- Figure 5.10 <sup>19</sup>F MAS NMR spectra of 5-<sup>19</sup>F-Trp,U-<sup>15</sup>N CA tubes (11.74 T; MAS frequency 4 kHz (a), and 19.96 T; MAS frequency 15 kHz (b))..... 157
- Figure 5.11 <sup>19</sup>F MAS NMR spectra of 5-<sup>19</sup>F-Trp CA mutant assemblies acquired at 19.96 T; the MAS frequency was 15 kHz without <sup>1</sup>H decoupling.... 157
- Figure 5.12 T<sub>1</sub> buildup curves and T<sub>2</sub> decay curves of 5-<sup>19</sup>F-Trp CA wild type (a) and A14C/E45C/W184A/M185A mutant assemblies (b). Experiments were acquired at 19.96 T and MAS frequency of 40 kHz. .... 159

Figure 5.13 2D  $^{19}\text{F}$ - $^{19}\text{F}$  correlation spectra of 5- $^{19}\text{F}$ -Trp CA (a), CA A14C/E45C/W184A/M185A (b), 5F-Trp-substituted CA mutant assemblies (c). (d) Experimental  $^{19}\text{F}$ - $^{19}\text{F}$  RFDR buildup curves of the W80-W133 and W133-W80 cross-peak intensities in the 5- $^{19}\text{F}$ -Trp A14C/E45C/W184A/ M185A CA cross-linked hexamer tubes (MAS frequency of 40 kHz). (e) Simulated  $^{19}\text{F}$ - $^{19}\text{F}$  RFDR buildup curves for  $^{19}\text{F}$ - $^{19}\text{F}$  distances of 5-20 Å, prepared by Matthew Fritz in our lab. .... 164

## LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AIDS	Acquired Immunodeficiency Syndrome
BVM	Bevirimat
CA	Capsid
CCR5	C-C Chemokine Receptor Type 5
CD4+	Cluster of Differentiation 4
CODEX	Center-Band Only Detection of Exchange
CORD	Combined $R_{2n}^v$ Driven
COSY	Correlation Spectroscopy
CP	Cross Polarization
CSA	Chemical Shift Anisotropy
CST	Chemical Shift Tensor
CSI	Chemical Shift Index
CSP	Chemical Shift Perturbation
CTD	C-Terminal Domain
CypA	Cyclophilin A
CXCR4	C-X-C Chemokine Receptor Type 4
DARR	Dipolar Assisted Rotational Resonance
DBI	Delaware Biotechnology Institute
DIPSHIFT	Dipolar Chemical Shift Correlation

DNA	Deoxyribonucleic Acid
DNP	Dynamic Nuclear Polarization
EM	Electron Microscopy
Env	Envelope
ET	Electron Tomography
FFT	Fast Fourier Transform
FID	Free Induction Decay
FRET	Förster Resonance Energy Transfer
HBA	Herzfeld-Berger Analysis
HIV	Human Immunodeficiency Virus
HETCOR	Heteronuclear Correlation
IN	Integrase
INADEQUATE	Incredible Natural Abundance Double Quantum Transfer Experiment
LP	Linear Prediction
MA	Matrix
MAS	Magic Angle Spinning
MD	Molecular Dynamics
MHR	Major Homology Region
MI	Maturation Inhibitors
MS	Mass Spectrometry
NC	Nucleocapsid
NOESY	Nuclear Overhauser Effect Spectroscopy
NMR	Nuclear Magnetic Resonance
NTD	N-Terminal Domain

NUS	Nonuniform Sampling
PARS	Phase-Alternating R-Symmetry
PDB	Protein Databank
PDSB	Proton Driven Spin Diffusion
PEG	Polyethylene Glycol
PIC	Pre-Integration Complex
PLUQ	PACSYlite Unified Query
PR	Enzyme Protease
QM	Quantum Mechanical
RF	Radio Frequency
RAP	Reduced Adjoining Protonation
RFDR	Radio Frequency Driven Recoupling
RMSD	Root Mean Squared Deviation
RNA	Ribonucleic Acid
RT	Reverse Transcriptase
RTC	Reverse Transcription Complex
SAIL	Stereo-Array Isotope Labeling
SAXS	Small-Angle X-Ray Scattering
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SP1	Spacer Peptide 1
SP2	Spacer Peptide 2
SPINAL	Small Phase Incremental Alternation
SU	Surface Membrane Protein
TALOS	Torsion Angle Likelihood Obtained From Shift and Sequence Similarity

TEM	Transmission Electron Microscopy
TM	Trans-Membrane Protein
TOBSY	Total Through-Bond Correlation Spectroscopy
TPPI	Time-Proportional Phase Incrementation
UNAIDS	The Joint United Nations Programme on HIV/AIDS
WHO	World Healthy Organization
WT	Wild Type

## ABSTRACT

Human immunodeficiency virus (HIV) is the causative agent for the acquired immune deficiency syndrome (AIDS), a global pandemic. The curative treatments or vaccines for HIV are still lacking to date. The HIV-1 capsid (CA) protein plays essential roles in the HIV-1 life cycle, involving in two critical assembly events, formation of the immature viral particle as a component of Gag polyprotein and re-assembly into a mature capsid core after Gag cleavage. During CA maturation, the Gag polyprotein cleaves into its constituent domains. The final step in the Gag processing cascade is the cleavage of spacer peptide 1 (SP1) from the C-terminal domain of CA. Following maturation, approximately 1000-1500 copies of CA protein arrange into a conical core to protect the viral genome in a mature virion. Initial research has identified that maturation inhibitors (MI) prevent the cleavage of CA from SP1, and formation of infectious virions. Thus, the HIV-1 CA maturation is an attractive target for therapeutic intervention. However, atomic-level understanding of CA maturation and maturation inhibition is lacking.

In this dissertation, I employed magic angle spinning (MAS) solid-state NMR spectroscopy to investigate the structural rearrangements and dynamic changes accompanying capsid maturation. By examining the final-step maturation intermediate, CA-SP1, I demonstrate the presence of as well as quantified dynamic helix-coil equilibrium in the CA-SP1 assemblies, and discovered that it is inhibited by the T8I mutation in the SP1 domain that phenocopies the MI-bound state. I have combined MAS NMR and molecular dynamics (MD) simulations to obtain

unprecedented atomic-level quantitative insights into the assemblies of maturation intermediates, unavailable from other techniques. Overall, the results indicate that modulation of protein dynamics appears to be a determining factor in capsid maturation and maturation inhibition. This study is described in Chapter 4.

To further understand the role of dynamics in the capsid assembly, I have also applied MAS NMR to probe atomic-resolution structures of the CA capsid protein assemblies of different morphologies and of model pentameric and hexameric CA building blocks. This study is described in Chapter 3.

In Chapter 5, I have explored  $^{19}\text{F}$  as an NMR probe for applications to biological assemblies. A robust protocol for  $^{19}\text{F}$  MAS NMR spectroscopy has been developed based characterization of structure and dynamics in fluorinated solids by examining fluorosubstituted tryptophans. I have applied this approach to the HIV-1 capsid protein assemblies. The results exhibit the fast and ultrafast MAS frequencies are beneficial to resolution enhancement and recording interfluorine distances.

I have incorporated 5-F-Trp into CA protein, and resonances were assigned by mutagenesis. The  $^{19}\text{F}$  chemical shifts for the five tryptophans are distinct, reflecting differences in local environment. Spin diffusion and radio frequency driven recoupling (RFDR) experiments were performed at fast MAS frequencies of 35-60 kHz and permitted establishing the  $^{19}\text{F}$ - $^{19}\text{F}$  correlations, yielding interatomic distances as long as 23 Å. Fast MAS frequencies of 35-60 kHz are essential and yield narrow lines. I demonstrate the potential of  $^{19}\text{F}$  NMR for structural analysis in large protein assemblies by MAS NMR.

## Chapter 1

### INTRODUCTION TO HIV-1 STRUCTURAL BIOLOGY: MATURATION AND CAPSID ASSEMBLY

#### 1.1 AIDS Epidemic and HIV

Health problems are one of the major factors that affect people's living condition and quality in modern society. Extensive research has been devoted to fighting off devastating illnesses and epidemics, one of which is the acquired immune deficiency syndrome (AIDS), caused by human immunodeficiency virus (HIV). Since the AIDS epidemic was first recognized in the 1980s, HIV has infected more than 70 million people and taken 35 million people's lives worldwide. Lacking protective vaccine, it remains a serious human pathogen for the world. According to the latest data that were reported by the Joint United Nations Programme on HIV/AIDS (UNAIDS), in 2016, HIV newly infected 1.8 million people and approximately 36.7 million people are living with HIV globally.<sup>1</sup> The global prevalence of HIV among adults is shown on the map (Fig 1.1) prepared by World Health Organization (WHO). Active therapeutic treatments have been developed to control the HIV infection; meanwhile, HIV vaccines have been tested in clinical trials. However, there is no effective treatment to cure or prevent HIV so far.

HIV is a retrovirus that has been identified in two different types, HIV-1 and HIV-2. HIV-1 is the major causative agent of AIDS for human beings, which infects by attacking the functional cells in the immune system, for instance CD4<sup>+</sup> T-cells.<sup>2</sup>

Over time, HIV depletes a great number of the host cells to weaken human immune system, so that the body loses the resistance to other infections and diseases.

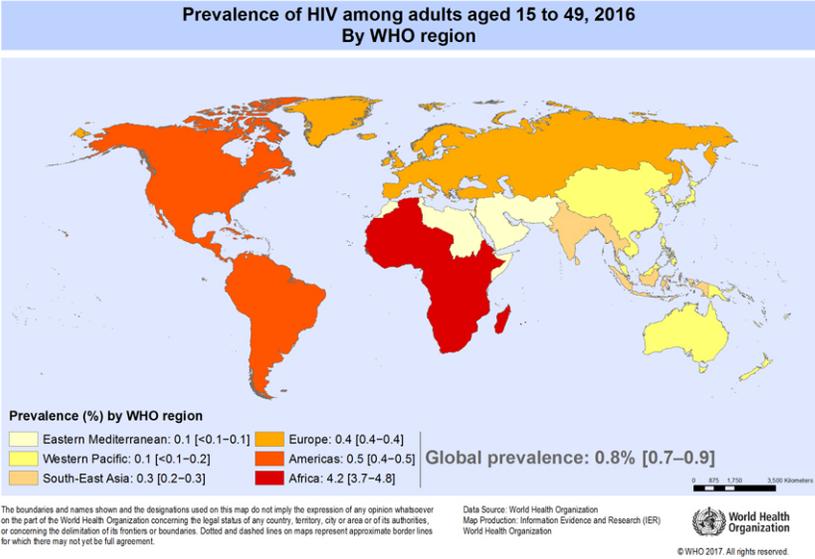


Figure 1.1 Map of HIV/AIDS global situation and trends. Made by WHO. <http://www.who.int/gho/hiv/en/>

The HIV virus is an icosahedral particle with a diameter of 120 nm. During its life cycle, immature and mature virions remain before and after viral maturation. The total size of HIV-1 genome is approximately 9.7 kb, and is comprised of three major genes (gag, pol, env), as well as a number of regulatory and accessory genes (vif, vpr, vpu, nef, tat and rev), shown in Fig 1.2.<sup>3</sup> The gag gene encodes a polyprotein precursor, named as Gag (Pr55<sup>Gag</sup>), the major structural protein as well as driver in virus assembly and release. The pol gene encodes the viral enzyme protease (PR), reverse transcriptase (RT), and integrase (IN). The GagPol polyprotein is a rare

frameshifting event during Gag cleavage process. The env-encoded envelope (Env) glycoproteins, including surface membrane protein (SU) and trans-membrane protein (TM), trigger the membrane fusion reaction during the virus entry.<sup>4</sup> The maturation process is a key step involving alteration of immature to mature virions (Fig 1.3). In an immature virion, Gag polyproteins form a sphere arranged from hexameric building blocks.<sup>5</sup> To form an infectious particle, the viral protease (PR) triggers virus maturation by cleaving various sites in both Gag and GagPol polyproteins. The Gag polyprotein is cleaved into its components, matrix (MA or p17), capsid (CA or p24), nucleocapsid (NC or p7), p6, and two spacer peptides, SP1 and SP2. Following the cleavage, mature CA subunits reassemble into a cone-shape lattice to protect the viral genome, when a mature virion forms.<sup>6,7</sup> The work presented in this dissertation focuses on the mature CA lattice formation by characterizing the structure and dynamics of assemblies of CA protein as well as of the maturation intermediate CA-SP1 protein.

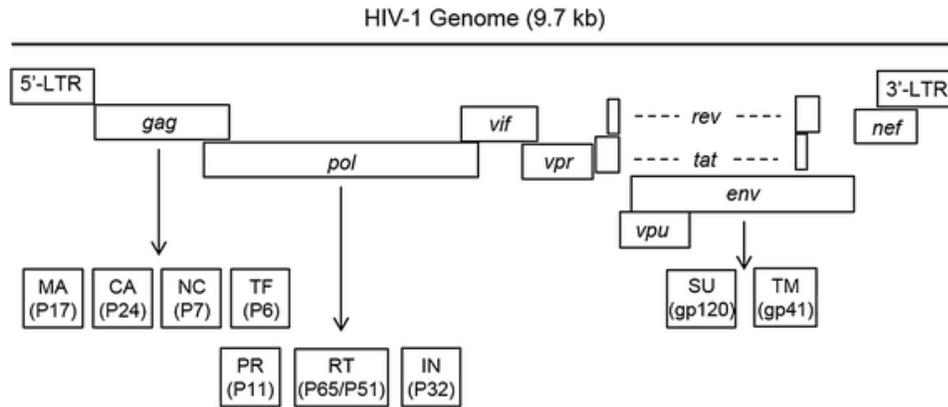


Figure 1.2 Organization of the HIV-1 genome.<sup>3</sup> The total size of the HIV-1 genome is 9.7 kb. Each of the viral genes is depicted based on the relative position in the RNA genome. Dashed lines represent RNA splicing. The number in parenthesis is molecular weight of each protein. The figure was originally published in REFERENCES<sup>3</sup>, copyright by Nkeze et al., 2015.

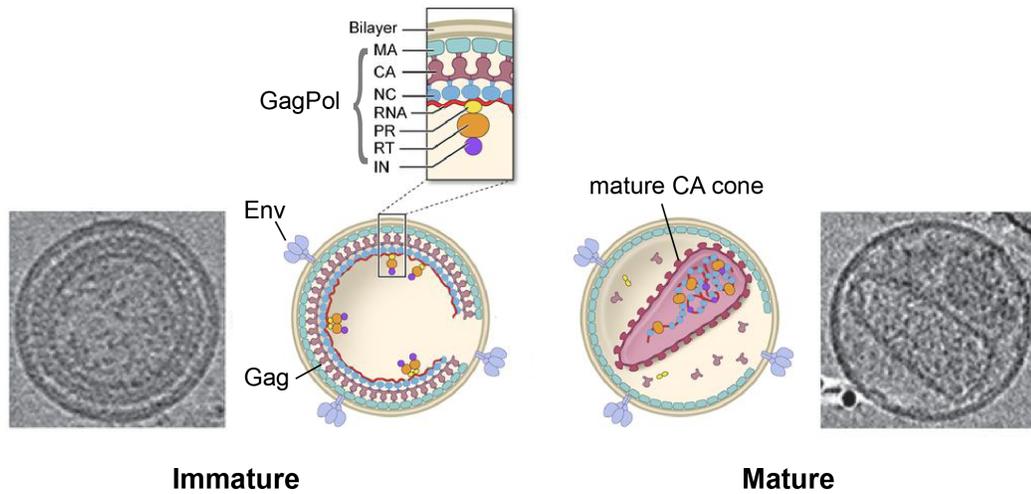
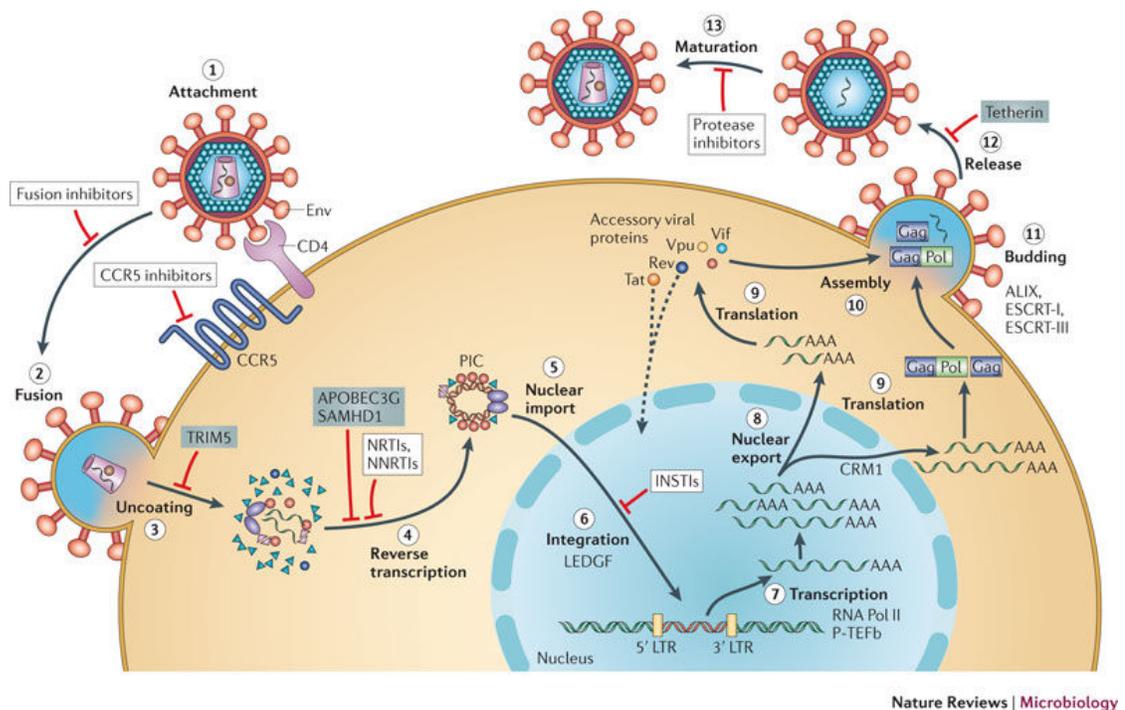


Figure 1.3 HIV-1 immature and mature virions. Both cryo-electron tomography (Cryo-ET) images and schematic representations illustrate the architecture of immature and mature virions. The figure adapted from REFERENCES<sup>6,8</sup>. Permissions for reuse in this dissertation is granted by Springer Nature, copyright 2015, and American Society for Microbiology, copyright 2015.

## 1.2 HIV-1 Viral Life Cycle

The HIV-1 viral life cycle is complex and contains various phases (Fig 1.4<sup>9</sup>), which are classified into early and late replication phases. The early phases start when the envelope (Env) glycoprotein of an infectious viral particle binds to the primary receptor, CD4 and co-receptor, CCR5 or CXCR4 on the surface of a CD4<sup>+</sup> T-cell.<sup>10</sup> Following the virion's attachment, fusion of the viral and cellular membranes occurs to release the viral core into the cytoplasm of the target cell. The mature capsid undergoes uncoating process where the capsid core rearranges to become the reverse transcription complex (RTC) by capsid immediate or complete disassembly to activate reverse transcription of the viral RNA to double-stranded DNA.<sup>11</sup> The capsid uncoating mechanism remains poorly understood. Through various transportins and

nucleoporins, the pre-integration complex (PIC) is imported into the host cell nucleus and followed by the proviral integration into the host genome.<sup>12</sup> Subsequently, the late phases begin with the proviral transcription that yields different sizes of viral mRNAs. The large mRNAs are exported from the nucleus, composing templates for viral protein production. Protein components are synthesized and assemble into a viral particle. Afterwards, an immature virion buds and releases from the infected cell. Finally, an infectious viral particle, containing a conical mature capsid lattice, is formed by protease-mediated maturation step. In the entire HIV-1 life cycle, several critical steps can be considered as potential targets for antiviral intervention<sup>13-15</sup> by using a variety of clinical inhibitors and cellular restriction factors.



Nature Reviews | Microbiology

Figure 1.4 Schematic of the HIV-1 life cycle. The figure was originally published in REFERENCES<sup>9</sup> Permission for reuse in this dissertation is granted by Springer Nature, copyright 2012.

### 1.3 HIV-1 CA Assembly

The capsid protein (CA) performs irreplaceable functions during HIV-1 morphogenesis. It is involved in two critical assembly events; one is the formation of the immature viral particle, where CA is a component of Gag polyprotein<sup>5,16</sup>, and the other is re-assembly of 1000-1500 CA subunits into a cone-shape lattice after Gag cleavage, resulting in an infectious viral particle.<sup>17-19</sup>

The mature CA lattice is a cone-shape structure and thought to be comprised of hexamers and pentamers. In one all-atom model, there are 216 hexamers and 12 pentamers<sup>18,20</sup> that allow adoption of different curvatures for the ovoid closure (Fig 1.5a). These hexamers and pentamers are made up of CA monomers. The CA monomer is 25.6 kDa, with 231 amino acids, and folds into N-terminal domain (NTD) and C-terminal domain (CTD) that are connected by a flexible linker region<sup>6,21,22</sup>, Fig 1.5b. CA NTD begins with a beta-hairpin structure, followed by 7 helical domains and a Pro-rich dynamic loop that binds to Cyclophilin A (a host cell factor) to regulate viral infectivity, named CypA loop.<sup>23-25</sup> CA CTD folds into 4 helices and a 3<sub>10</sub> helix, where several essential interfaces (e.g. dimer interface and trimer interface), play critical roles in CA multimerization.<sup>17,18,26</sup> Specifically, the major homology region (MHR), spanning residues 153-172, is a highly conserved motif in retroviruses and serves an important function in virus assembly.<sup>6,27</sup>

*In vitro*, CA proteins can assemble into different morphologies, for instance sheets, tubes, spheres and cones.<sup>20,28</sup> In our lab, we established a protocol to prepare tubular assemblies by using buffers containing 1-2.4 M NaCl, which yields high-quality samples for solid-state NMR studies. Both the wild type (WT) and mutant CA proteins assemble into stable tubes, Fig 1.6, representing the predominant hexameric

building blocks in the mature CA lattice. In this dissertation, my research mainly concerns tubular assemblies of CA.

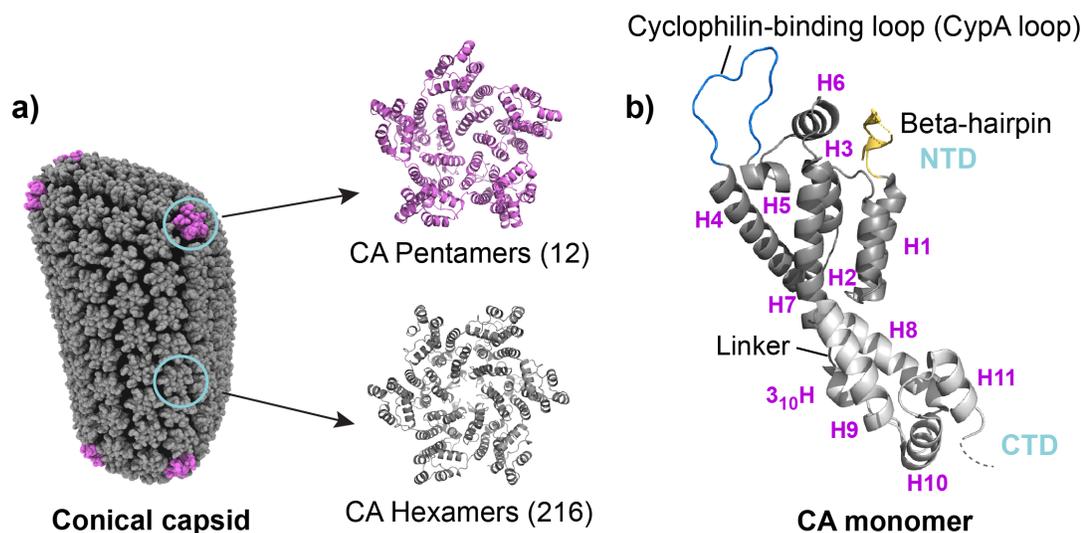


Figure 1.5 The all-atom model of HIV-1 CA conical lattice with the structures of pentamers and hexamers (a) and the structure of CA monomer (b). The structures of pentamer and hexamer are not native units of the cone, and they are cross-linked pentamer and hexamer of CA mutants. PDB ID: 3J3Y for CA cone model<sup>18</sup>, 4XFX for CA monomer, NTD in dark gray and CTD in light gray<sup>29</sup>, 3P05 for CA pentamer<sup>17</sup>, and 3H47 for CA hexamer<sup>30</sup>.

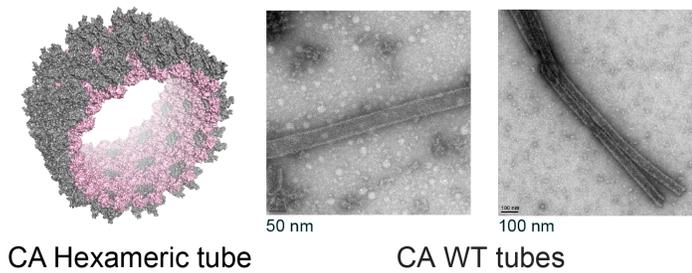


Figure 1.6 A model of CA tube (NTD in gray and CTD in pink) PDB ID 3J4F<sup>18</sup>, and transmission electron microscope (TEM) images of CA WT tubular assemblies.<sup>24</sup>

#### 1.4 HIV-1 Maturation and Maturation Inhibitors

HIV-1 maturation is a critical step to trigger formation of infectious particles that occurs during the late phase of HIV replication cycle. In the past decades, small-molecule maturation inhibitors have been discovered to prevent the Gag polyprotein cleavage and further block the formation of mature virions.<sup>31,32</sup> The Gag polyprotein cleavage during HIV-1 CA maturation, as well as its inhibition by maturation inhibitors will be discussed in this section.

##### 1.4.1 Gag Polyprotein Cleavage and Maturation Intermediate CA-SP1

During the late phase of HIV-1 life cycle, Gag polyprotein is synthesized. Gag serves as the structural protein of the viral particles. In parallel, due to a ribosomal frameshifting event, GagPol polyprotein that contains viral enzymes (PR, TR, and IN) is produced at approximately 5% amount compared to Gag population.<sup>6,7</sup> The GagPol polyprotein is a rare frameshifting event during Gag cleavage process. Gag polyprotein is composed of, from N- to C-terminus, MA, CA, SP1, NC, SP2 and p6 proteins, arranged as a spherical lattice in the immature virion. During maturation, the Gag polyprotein undergoes a sequential cleavage into its constituent proteins, Fig 1.7.

Along with the separation of Gag and GagPol precursor polyproteins, the Gag cleavage begins at the SP1-NC site, followed by the cleavage at the sites of MA-CA and SP2-p6, resulting in the dissection of CA from the membrane-bound MA layer. The final cleavage occurs at the sites of CA-SP1 and NC-SP2, respectively, to form the mature virion. However, this complicated process is still not understood in detail, and three possible pathways have been proposed on the basis of the available evidence, to explain the maturation mechanism (Fig 1.7): i) slow rearrangement of the immature-like lattice to form the mature capsid cone (displacive)<sup>33</sup>, ii) disassembly of the immature lattice induced by SP1 cleavage from CA<sup>34-36</sup>, *de novo* the mature CA cone reassembly from a pool of CA monomers<sup>37,38</sup>, and iii) a sequential combination of displacive and *de novo* processes<sup>39</sup>.

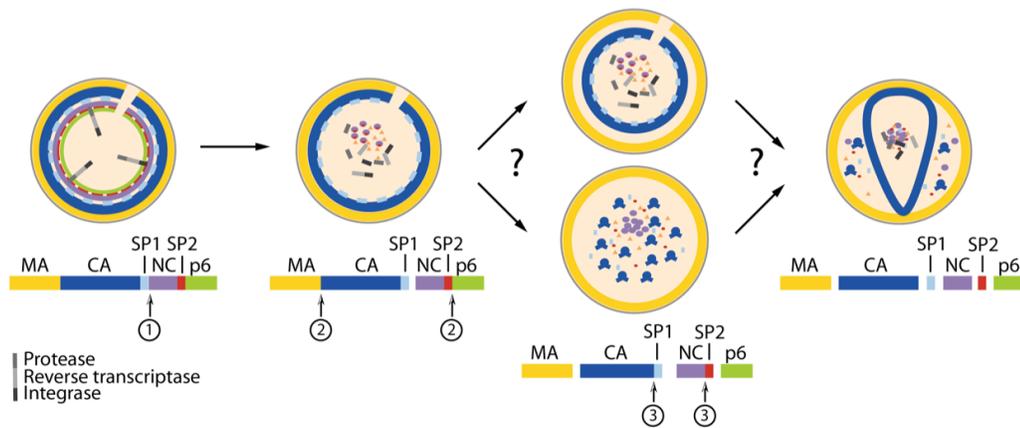


Figure 1.7 Schematic diagram of the HIV-1 Gag sequential cleavage and virus maturation process.<sup>40</sup> The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

Among the Gag's components, the 14-residues SP1 was proposed to play a key role as a molecular switch during the final step of the maturation process. As the last-step maturation intermediate, CA-SP1 has attracted much attention. In the context of the immature and mature lattice, SP1 forms different conformations. The high-resolution structure of the assembled immature HIV-1 particles has recently been solved by cryo-EM.<sup>41,42</sup> Cryo-EM results have indicated the presence of a six-helix bundle of SP1 in the immature Gag or modified Gag constructs.<sup>38,42,43</sup> Recent findings on the crystal structure of the CA CTD-SP1 Gag fragment suggest the CA-SP1 junction can form a six-helix bundle under certain crystallization conditions<sup>44</sup> (PDB: 5I4T). In tubular assemblies of the CA-SP1 maturation intermediate, the SP1 region adopts a dynamic random coil structure, as shown by magic angle spinning (MAS) NMR studies.<sup>33</sup>

#### **1.4.2 Maturation Inhibitors**

Many anti-retroviral treatments have been developed to target many different stages of the HIV lifecycle. Maturation inhibitors are a new class of anti-HIV targets, which potently prevents infectious particle formation by blocking the CA-SP1 cleavage.<sup>45,46</sup>

A small molecule 3-O-(3'3'-dimenthylsuccinyl)betulinic acid (DSB) in Fig 1.8a, known as Bevirimat (BVM) or PA-457, is one of the early maturation inhibitors.<sup>47</sup> BVM was first isolated from natural herb leaves, *Syzygium calviflorum*<sup>48</sup>, and then its analogs<sup>31,46,49</sup> have been derived to increase the solubility while retaining the maturation inhibition properties. PF-46396 (PF96) 1-[2-(4-tertbutylphenyl)-2-(2,3-dihydro-1H-inden-2-ylamino)ethyl]-3-(trifluoromethyl)pyridin-2(1H)-one, Fig 1.8b, is another maturation inhibitor that was discovered by drug screening.<sup>50,51</sup> BVM and

PF96 were proposed to block the CA-SP1 cleavage and prevent the formation of mature virion by means of binding to the CA-SP1 junction as well as stabilizing the immature lattice.<sup>47,52</sup> Despite extensive studies, the MI binding sites and the maturation inhibition are poorly understood.

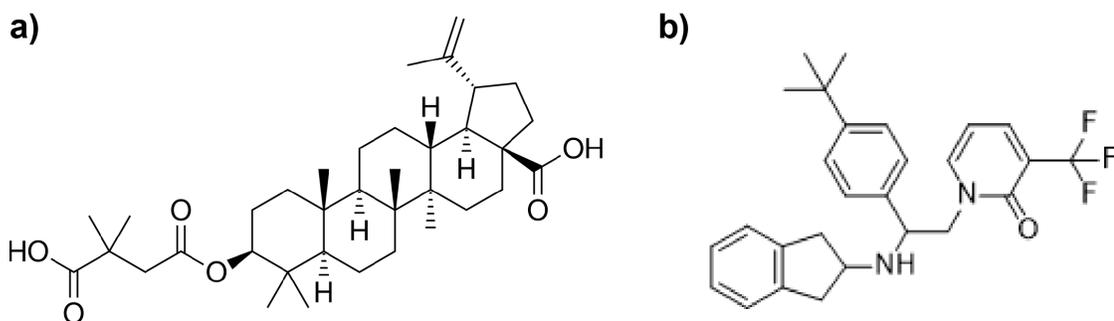


Figure 1.8 Chemical structures of Bevirimat molecule (a) and PF-46396 molecule (b).

#### 1.4.3 T8I Mutation in Spacer Peptide 1 (SP1)

BVM has been tested and managed to clinical trials, showing to exhibit great safety and efficacy.<sup>53</sup> However, it has been reported that sequence polymorphism in the SP1 (6-8 amino acids) is associated with BVM resistance<sup>54,55</sup>, implying possible BVM binding sites. For PF96, resistance mutations were recognized to reside in three regions, CA-SP1 junction, MHR and residue 201 in CA domain.<sup>51</sup> PF96-dependent MHR and SP1 mutants express a replication defect, which could be rescued by a second-site compensatory mutation, T8I mutation in SP1.<sup>45,51</sup> This T8I mutant is highly replication-defective and prevents partial CA-SP1 cleavage.<sup>51,56</sup> It was thus recognized to phenocopy the effect of MI binding by presumably interfering with the SP1 cleavage and stabilizing the immature lattice<sup>56</sup>. Therefore, in chapter 4, the

structure and dynamics of CA-SP1(T8I) construct were probed in the context of its being a maturation-arresting mutant to further our understanding on the MI bound state.

## REFERENCES

- (1) Joint United Nations Programme on HIV/AIDS (UNAIDS), 2017.
- (2) Sharp, P. M.; Hahn, B. H. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* **2011**, *1* (1), a006841.
- (3) Nkeze, J.; Li, L.; Benko, Z.; Li, G.; Zhao, R. Y. Molecular characterization of HIV-1 genome in fission yeast *Schizosaccharomyces pombe*. *Cell Biosci* **2015**, *5*, 47.
- (4) Freed, E. O. HIV-1 Replication. *Somatic Cell and Molecular Genetics* **2001**, *26*.
- (5) Briggs, J. A. G.; Riches, J. D.; Glass, B.; Bartonova, V.; Zanetti, G.; Krausslich, H. G. Structure and assembly of immature HIV. *P Natl Acad Sci USA* **2009**, *106* (27), 11090.
- (6) Freed, E. O. HIV-1 assembly, release and maturation. *Nat Rev Microbiol* **2015**, *13* (8), 484.
- (7) Tedbury, P. R.; Freed, E. O. HIV-1 Gag: An Emerging Target for Antiretroviral Therapy. *Curr Top Microbiol* **2015**, *389*, 171.
- (8) Fontana, J.; Jurado, K. A.; Cheng, N.; Ly, N. L.; Fuchs, J. R.; Gorelick, R. J.; Engelman, A. N.; Steven, A. C. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. *J Virol* **2015**, *89* (19), 9765.
- (9) Engelman, A.; Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol* **2012**, *10* (4), 279.
- (10) Wilen, C. B.; Tilton, J. C.; Doms, R. W. HIV: Cell Binding and Entry. *Csh Perspect Med* **2012**, *2* (8).

- (11) Campbell, E. M.; Hope, T. J. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol* **2015**, *13* (8), 471.
- (12) Matreyek, K. A.; Engelman, A. Viral and Cellular Requirements for the Nuclear Entry of Retroviral Preintegration Nucleoprotein Complexes. *Viruses-Basel* **2013**, *5* (10), 2483.
- (13) Flexner, C. HIV drug development: the next 25 years. *Nat Rev Drug Discov* **2007**, *6* (12), 959.
- (14) Arts, E. J.; Hazuda, D. J. HIV-1 Antiretroviral Drug Therapy. *Csh Perspect Med* **2012**, *2* (4).
- (15) Ballantyne, A. D.; Perry, C. M. Dolutegravir: First Global Approval. *Drugs* **2013**, *73* (14), 1627.
- (16) Bharat, T. A. M.; Menendez, L. R. C.; Hagen, W. J. H.; Lux, V.; Igonet, S.; Schorb, M.; Schur, F. K. M.; Krausslich, H. G.; Briggs, J. A. G. Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *P Natl Acad Sci USA* **2014**, *111* (22), 8233.
- (17) Pornillos, O.; Ganser-Pornillos, B. K.; Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **2011**, *469* (7330), 424.
- (18) Zhao, G. P.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J. Y.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **2013**, *497* (7451), 643.
- (19) Ganser-Pornillos, B. K.; Yeager, M.; Sundquist, W. I. The structural biology of HIV assembly. *Curr Opin Struct Biol* **2008**, *18* (2), 203.
- (20) Ganser-Pornillos, B. K.; Cheng, A.; Yeager, M. Structure of full-length HIV-1CA: a model for the mature capsid lattice. *Cell* **2007**, *131* (1), 70.
- (21) Mateu, M. G. The capsid protein of human immunodeficiency virus: intersubunit interactions during virus assembly. *Febs J* **2009**, *276* (21), 6098.

- (22) Gitti, R. K.; Lee, B. M.; Walker, J.; Summers, M. F.; Yoo, S.; Sundquist, W. I. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **1996**, *273* (5272), 231.
- (23) Gamble, T. R.; Vajdos, F. F.; Yoo, S. H.; Worthylake, D. K.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **1996**, *87* (7), 1285.
- (24) Lu, M. M.; Hou, G. J.; Zhang, H. L.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Perilla, J. R.; Langmead, C. J.; Hung, I.; Gor'kov, P. L. et al. Dynamic allostery governs cyclophilin A-HIV capsid interplay. *P Natl Acad Sci USA* **2015**, *112* (47), 14617.
- (25) Liu, C.; Perilla, J. R.; Ning, J. Y.; Lu, M. M.; Hou, G. J.; Ramalho, R.; Himes, B. A.; Zhao, G. P.; Bedwell, G. J.; Byeon, I. J. et al. Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat Commun* **2016**, *7*.
- (26) Byeon, I. J. L.; Hou, G. J.; Han, Y.; Suiter, C. L.; Ahn, J.; Jung, J.; Byeon, C. H.; Gronenborn, A. M.; Polenova, T. Motions on the Millisecond Time Scale and Multiple Conformations of HIV-1 Capsid Protein: Implications for Structural Polymorphism of CA Assemblies. *J Am Chem Soc* **2012**, *134* (14), 6455.
- (27) Mammano, F.; Ohagen, A.; Høglund, S.; Gottlinger, H. G. Role of the Major Homology Region of Human-Immunodeficiency-Virus Type-1 in Virion Morphogenesis. *Journal of Virology* **1994**, *68* (8), 4927.
- (28) Han, Y.; Ahn, J.; Concel, J.; Byeon, I. J. L.; Gronenborn, A. M.; Yang, J.; Polenova, T. Solid-State NMR Studies of HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2010**, *132* (6), 1976.
- (29) Gres, A. T.; Kirby, K. A.; KewalRamani, V. N.; Tanner, J. J.; Pornillos, O.; Sarafianos, S. G. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **2015**, *349* (6243), 99.
- (30) Pornillos, O.; Ganser-Pornillos, B. K.; Kelly, B. N.; Hua, Y.; Whitby, F. G.; Stout, C. D.; Sundquist, W. I.; Hill, C. P.; Yeager, M. X-ray structures of the hexameric building block of the HIV capsid. *Cell* **2009**, *137* (7), 1282.

- (31) Li, F.; Goila-Gaur, R.; Salzwedel, K.; Kilgore, N. R.; Reddick, M.; Matallana, C.; Castillo, A.; Zoumplis, D.; Martin, D. E.; Orenstein, J. M. et al. PA-457: A potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *P Natl Acad Sci USA* **2003**, *100* (23), 13555.
- (32) Zhou, J.; Yuan, X.; Dismuke, D.; Forshey, B. M.; Lundquist, C.; Lee, K. H.; Aiken, C.; Chen, C. H. Small-Molecule Inhibition of Human Immunodeficiency Virus Type 1 Replication by Specific Targeting of the Final Step of Virion Maturation. *Journal of Virology* **2003**, *78* (2), 922.
- (33) Han, Y.; Hou, G. J.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Lipton, A. S.; Burton, S.; Hung, I.; Gor'kov, P. L.; Gan, Z. H. et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2013**, *135* (47), 17793.
- (34) Krausslich, H. G.; Facke, M.; Heuser, A. M.; Konvalinka, J.; Zentgraf, H. The Spacer Peptide between Human-Immunodeficiency-Virus Capsid and Nucleocapsid Proteins Is Essential for Ordered Assembly and Viral Infectivity. *Journal of Virology* **1995**, *69* (6), 3407.
- (35) Pettit, S. C.; Henderson, G. J.; Schiffer, C. A.; Swanstrom, R. Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *Journal of Virology* **2002**, *76* (20), 10226.
- (36) Woodward, C. L.; Cheng, S. N.; Jensen, G. J. Electron Cryotomography Studies of Maturing HIV-1 Particles Reveal the Assembly Pathway of the Viral Core. *Journal of Virology* **2015**, *89* (2), 1267.
- (37) Datta, S. A. K.; Temeselew, L. G.; Crist, R. M.; Soheilian, F.; Kamata, A.; Mirro, J.; Harvin, D.; Nagashima, K.; Cachau, R. E.; Rein, A. On the Role of the SP1 Domain in HIV-1 Particle Assembly: a Molecular Switch? *Journal of Virology* **2011**, *85* (9), 4111.
- (38) Frank, G. A.; Narayan, K.; Bess, J. W.; Del Prete, G. Q.; Wu, X. W.; Moran, A.; Hartnell, L. M.; Earl, L. A.; Lifson, J. D.; Subramaniam, S. Maturation of the HIV-1 core by a non-diffusional phase transition. *Nat Commun* **2015**, *6*.

- (39) Wright, E. R.; Schooler, J. B.; Ding, H. J.; Kieffer, C.; Fillmore, C.; Sundquist, W. I.; Jensen, G. J. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *Embo J* **2007**, *26* (8), 2218.
- (40) Wang, M. Z.; Quinn, C. M.; Perilla, J. R.; Zhang, H. L.; Shirra, R.; Hou, G. J.; Byeon, I. J.; Suiter, C. L.; Ablan, S.; Urano, E. et al. Quenching protein dynamics interferes with HIV capsid maturation. *Nat Commun* **2017**, *8*.
- (41) Schur, F. K. M.; Obr, M.; Hagen, W. J. H.; Wan, W.; Jakobi, A. J.; Kirkpatrick, J. M.; Sachse, C.; Kräusslich, H.-G.; Briggs, J. A. G. An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* **2016**, *353* (6298), 506.
- (42) Schur, F. K. M.; Hagen, W. J. H.; Rumlova, M.; Ruml, T.; Muller, B.; Krausslich, H. G.; Briggs, J. A. G. Structure of the immature HIV-1 capsid in intact virus particles at 8.8 angstrom resolution. *Nature* **2015**, *517* (7535), 505.
- (43) Bayro, M. J.; Ganser-Pornillos, B. K.; Zdrozny, K. K.; Yeager, M.; Tycko, R. Helical Conformation in the CA-SP1 Junction of the Immature HIV-1 Lattice Determined from Solid-State NMR of Virus-like Particles. *J Am Chem Soc* **2016**, *138* (37), 12029.
- (44) Wagner, J. M.; Zdrozny, K. K.; Chrustowicz, J.; Purdy, M. D.; Yeager, M.; Ganser-Pornillos, B. K.; Pornillos, O. Crystal structure of an HIV assembly and maturation switch. *Elife* **2016**, *5*.
- (45) Adamson, C. S.; Ablan, S. D.; Boeras, I.; Goila-Gaur, R.; Soheilian, F.; Nagashima, K.; Li, F.; Salzwedel, K.; Sakalian, M.; Wild, C. T. et al. In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (Bevirimat). *J Virol* **2006**, *80* (22), 10957.
- (46) Nguyen, A. T.; Feasley, C. L.; Jackson, K. W.; Nitz, T. J.; Salzwedel, K.; Air, G. M.; Sakalian, M. The prototype HIV-1 maturation inhibitor, bevirimat, binds to the CA-SP1 cleavage site in immature Gag particles. *Retrovirology* **2011**, *8*, 101.
- (47) Keller, P. W.; Adamson, C. S.; Heymann, J. B.; Freed, E. O.; Steven, A. C. HIV-1 Maturation Inhibitor Bevirimat Stabilizes the Immature Gag Lattice. *Journal of Virology* **2011**, *85* (4), 1420.

- (48) Fujioka, T.; Kashiwada, Y.; Kilkuskie, R. E.; Cosentino, L. M.; Ballas, L. M.; Jiang, J. B.; Janzen, W. P.; Chen, I. S.; Lee, K. H. Anti-Aids Agents .11. Betulinic Acid and Platonic Acid as Anti-Hiv Principles from *Syzigium-Claviflorum*, and the Anti-Hiv Activity of Structurally Related Triterpenoids. *J Nat Prod* **1994**, *57* (2), 243.
- (49) Urano, E.; Ablan, S. D.; Mandt, R.; Pauly, G. T.; Sigano, D. M.; Schneider, J. P.; Martin, D. E.; Nitz, T. J.; Wild, C. T.; Freed, E. O. Alkyl Amine Bevirimat Derivatives Are Potent and Broadly Active HIV-1 Maturation Inhibitors. *Antimicrob Agents Ch* **2016**, *60* (1), 190.
- (50) Blair, W. S.; Cao, J.; Fok-Seang, J.; Griffin, P.; Isaacson, J.; Jackson, R. L.; Murray, E.; Patick, A. K.; Peng, Q. H.; Perros, M. et al. New Small-Molecule Inhibitor Class Targeting Human Immunodeficiency Virus Type 1 Virion Maturation. *Antimicrob Agents Ch* **2009**, *53* (12), 5080.
- (51) Waki, K.; Durell, S. R.; Soheilian, F.; Nagashima, K.; Butler, S. L.; Freed, E. O. Structural and Functional Insights into the HIV-1 Maturation Inhibitor Binding Pocket. *Plos Pathog* **2012**, *8* (11).
- (52) Keller, P. W.; Huang, R. K.; England, M. R.; Waki, K.; Cheng, N. Q.; Heymann, J. B.; Craven, R. C.; Freed, E. O.; Steven, A. C. A Two-Pronged Structural Analysis of Retroviral Maturation Indicates that Core Formation Proceeds by a Disassembly-Reassembly Pathway Rather than a Displacive Transition. *Journal of Virology* **2013**, *87* (24), 13655.
- (53) Smith, P. F.; Ogundele, A.; Forrest, A.; Wilton, J.; Salzwedel, K.; Doto, J.; Allaway, G. P.; Martin, D. E. Phase I and II study of the safety, virologic effect, and pharmacokinetics/pharmacodynamics of single-dose 3-O-(3',3'-dimethylsuccinyl)betulinic acid (bevirimat) against human immunodeficiency virus infection. *Antimicrob Agents Ch* **2007**, *51* (10), 3574.
- (54) Adamson, C. S.; Sakalian, M.; Salzwedel, K.; Freed, E. O. Polymorphisms in Gag spacer peptide 1 confer varying levels of resistance to the HIV-1 maturation inhibitor bevirimat. *Retrovirology* **2010**, *7*.
- (55) Van Baelen, K.; Salzwedel, K.; Rondelez, E.; Van Eygen, V.; De Vos, S.; Verheyen, A.; Steegen, K.; Verlinden, Y.; Allaway, G. P.; Stuyver, L. J. Susceptibility of Human Immunodeficiency Virus Type 1 to the Maturation Inhibitor Bevirimat Is Modulated by Baseline Polymorphisms in Gag Spacer Peptide 1. *Antimicrob Agents Ch* **2009**, *53* (5), 2185.

- (56) Fontana, J.; Keller, P. W.; Urano, E.; Ablan, S. D.; Steven, A. C.; Freed, E. O. Identification of an HIV-1 Mutation in Spacer Peptide 1 That Stabilizes the Immature CA-SP1 Lattice. *Journal of Virology* **2016**, *90* (2), 972.

## Chapter 2

### STRUCTURE AND DYNAMICS OF PROTEIN ASSEMBLIES BY MAGIC ANGLE SPINNING SOLID-STATE NMR

#### 2.1 Introduction

Numerous biophysical studies using mutational analysis<sup>1-3</sup>, X-ray crystallography<sup>4-6</sup>, cryo-EM<sup>7-9</sup>, solution NMR<sup>10,11</sup>, small-angle X-ray scattering (SAXS)<sup>12,13</sup>, Förster resonance energy transfer (FRET)<sup>14</sup>, and computational simulation<sup>15</sup> have contributed to structural characterization of proteins by providing a variety of information in regard to conformation and dynamics. Nevertheless, it is particularly challenging to obtain atomic-resolution structures of large proteins due to size, insolubility, their non-crystalline nature, intrinsic flexibility, or sample instability.

Magic angle spinning (MAS) NMR is a complementary technique to study complex biological systems, including protein aggregates, membrane proteins, amyloid fibrils, viral capsids and protein-ligand complexes.<sup>16,17</sup> It is capable of probing both global and local dynamic information on a wide range of timescales. Moreover, MAS NMR has been increasingly integrated with other techniques, such as cryo-EM, X-ray crystallography, as well as quantum mechanical (QM) calculations, and molecular dynamics (MD) simulations to be beneficial to atomic-level structure refinement and understanding of dynamic process.<sup>18-22</sup> In this chapter, comprehensive strategies for studying structure and dynamics of protein assemblies at atomic resolution by MAS NMR will be discussed.

## 2.2 Magic Angle Spinning NMR Spectroscopy

Magic angle spinning is fundamental for obtaining high resolution and sensitivity spectra on solid-state protein samples.<sup>23</sup> In solution, the rapid random tumbling of nuclear spins averages out anisotropic spin interactions, which results in sharp isotropic lines and high-resolution NMR spectra. In contrast, the signals in solid-state NMR spectra are very broad due to strong anisotropic or orientation-dependent interactions. For instance, chemical shift anisotropy, dipolar couplings and quadrupolar couplings are present in such a condensed phase. These anisotropic interactions contain a spatial component,  $3 \cos^2 \theta - 1$ . By rotating the sample about an axis oriented at  $54.74^\circ$  (magic angle) with respect to the magnetic field, the term  $3 \cos^2 \theta - 1$  reduces to zero over the rotor cycle, and the anisotropic interactions can be averaged out. The averaging is more efficient as the spinning frequency increases, illustrated in Fig 2.1.<sup>24</sup> The spinning frequency needs to be greater than or equal to the magnitude of the anisotropic interaction in order to fully suppress it. In practice, spinning frequencies of 5-62 kHz are achieved with different NMR probes and using rotors with diameters ranging from 7 to 1.3 mm. With the recent development of commercial probe heads, currently, the spinning of 110 kHz has been reached using a 0.7 mm rotor, which provides higher resolution and sensitivity for studying large proteins using proton-detection at ultrafast spinning frequency.<sup>25,26</sup>

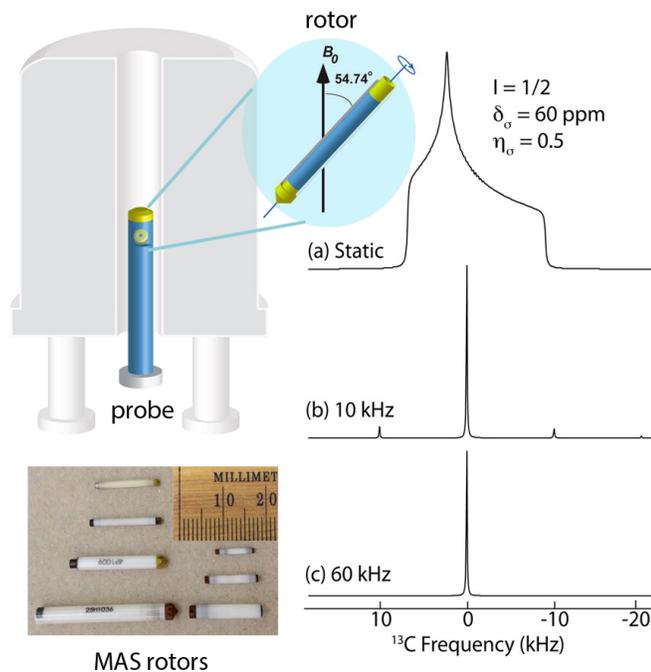


Figure 2.1 Schematic illustration of MAS NMR setup. The sample is packed into a rotor and placed into an NMR probe at the magic angle of  $54.74^\circ$  with respect to the magnetic field. The effect of MAS on NMR lineshapes for an  $I=1/2$  nucleus has been depicted in a-c. The broad static  $^{13}\text{C}$  powder pattern due to dipolar interaction (a) is split into a series of spinning sidebands (b). When the MAS frequency exceeds the magnitude of the anisotropic interaction, the interaction is averaged out into an isotropic peak (c). Figure is adapted from REFERENCES<sup>24</sup>. Permission for reuse in this dissertation is granted by American Chemical Society, copyright 2015.

### 2.3 Resonance Assignments

In the general protocol for protein structure elucidation by MAS NMR, the first essential step is performing site-specific resonance assignments by assigning correlated signals to specific atoms in a target protein. A protein with molecular weight of 20-50 kDa contains several hundred of amino acids, thousands of carbon atoms and hundreds of nitrogen atoms. Therefore, full resonance assignments for an entire protein are challenging and time-consuming. Various isotopic labeling schemes

and multidimensional NMR experiments are frequently used in a combination to achieve the site-specific resonance assignments accurately and rapidly. Since chemical shifts are sensitive to the local environment and global dynamics, they provide certain information of protein secondary structure and tertiary conformation. Protein secondary structures and torsion angles can be predicted by semi-empirical prediction programs (e.g. CSI 2.0) based on the isotropic chemical shifts of different atoms, which help to understand structure-function relationships and the behavior of proteins in complicated biological systems.

### 2.3.1 Isotopic Labeling Schemes

Isotopic labeling is required for NMR studies of proteins. In general, uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeling scheme is the most practical labeling method when one starts the NMR protein sample preparation if the protein is produced recombinantly from *E. coli* expression systems. Isotopically enriched chemicals such as uniformly  $^{13}\text{C}$ -labeled glucose and glycerol,  $^{15}\text{N}$ -labeled ammonium chloride and ammonium sulfate are used during protein growth and expression<sup>27</sup>, resulting in a set of intra-residue  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{13}\text{C}$ - $^{15}\text{N}$  correlations in the NMR spectra from all residues in the protein. However, a high degree of spectral crowding may exist due to the large number of resonances. In this case, the specific amino acid labeling and sparse labeling can be employed to reduce the amount of  $^{13}\text{C}$ -labeled sites and simplify the spectra, so as to increase the resolution.<sup>28</sup> Metabolically tailored  $^{13}\text{C}$  labeling methods based on  $[2\text{-}^{13}\text{C}]/[1,6\text{-}^{13}\text{C}]$ -glucose or  $[1,3\text{-}^{13}\text{C}]/[2\text{-}^{13}\text{C}]$ -glycerol have been applied to the structure determination of a number of proteins that have been studied by MAS NMR.<sup>29-32</sup> Perdeuteration is an alternative labeling method to greatly improve the spectral resolution, especially in the proton-detected experiments, by suppressing the proton-proton dipolar couplings.<sup>33,34</sup>

Moreover, the partial deuteration schemes, for instance reduced adjoining protonation (RAP)<sup>35</sup>, stereospecific isotopic labeling of methyl groups ILV<sup>36</sup>, and stereo-array isotope labeling (SAIL) have been exploited to reduce homonuclear dipolar couplings and transverse relaxation, which facilitate the assignment of aliphatic <sup>1</sup>H-<sup>13</sup>C correlations.<sup>17,37</sup>

### 2.3.2 Multidimensional NMR Experiments

A series of multidimensional (2D and 3D) homo- and heteronuclear correlation experiments comprise typical experimental protocols for protein resonance assignments by MAS NMR. The fundamental homonuclear correlation experiments for <sup>13</sup>C-<sup>13</sup>C correlation include spin-diffusion experiments (PDSD<sup>38</sup>, DARR<sup>39</sup>), radio frequency driven recoupling (RFDR) experiments<sup>40</sup> as well as combined supercycled R2<sub>n</sub><sup>v</sup> driven (CORD) experiments<sup>41</sup> that were developed by our lab. The CORD pulse scheme has higher efficiency than the spin-diffusion schemes for broadband recoupling at MAS frequencies ranging from 14 to 60 kHz, exhibiting uniform cross peak intensities across the entire correlation spectrum.<sup>41,42</sup> The 2D <sup>13</sup>C-<sup>13</sup>C correlation experiments are generally used to assign backbone and side chain carbon atoms in the same residue, as well as to obtain long distance correlations (Fig 2.2) by using longer mixing times. Heteronuclear correlation experiments have been developed and demonstrated experimentally on uniformly <sup>13</sup>C,<sup>15</sup>N-labeled proteins. By applying frequency selective heteronuclear polarization transfer, the specific N-C $\alpha$  and N-Co correlations are detected in the 2D NCA and NCO experiments<sup>43</sup> (Fig 2.2). The 2D/3D NCACX<sup>44</sup> experiments provide intra-residue correlations between nitrogen atom and the dipolar-coupled C $\alpha$  atoms, followed by the dipolar-based transfers to the side chain carbons, while the 2D/3D NCOCX<sup>44</sup> experiments are exploited to sequentially

assign resonances of neighboring residues. In NCOCX, the magnetization is transferred from the amide nitrogen in the residue (i) to the carbonyl carbon in the neighboring residue (i-1), followed by subsequent transfers to the  $\alpha$ -carbon and side-chain carbons in the prior residue (i-1). In addition, 3D CANCO(CX) and CONCA experiments are supplementary experiments used to achieve unambiguous sequential resonance assignments, by providing information of sequential correlations in the form of  $\text{Co}_{(i-1)}\text{-N}_i\text{-C}\alpha_i$  (Fig 2.2).

Besides the dipolar-coupling (through-space) based experiments mentioned above, the scalar-coupling (through-bond) based experiments are complementary for assigning resonances in highly dynamic regions of proteins, since J coupling interactions are less sensitive to molecular mobility.<sup>45,46</sup> Some MAS NMR J coupling-based sequences are adapted from solution NMR experiments, such as INADEQUATE (incredible natural abundance double quantum transfer experiment)<sup>47,48</sup>, COSY (correlation spectroscopy)<sup>49</sup>, and TOBSY (total through-bond correlation spectroscopy)<sup>50</sup>. In chapter 4, we show that the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  INADEQUATE spectra exhibit excellent resolution even in the highly mobile CTD-SP1 region, which is not detected in dipolar-based experiments.

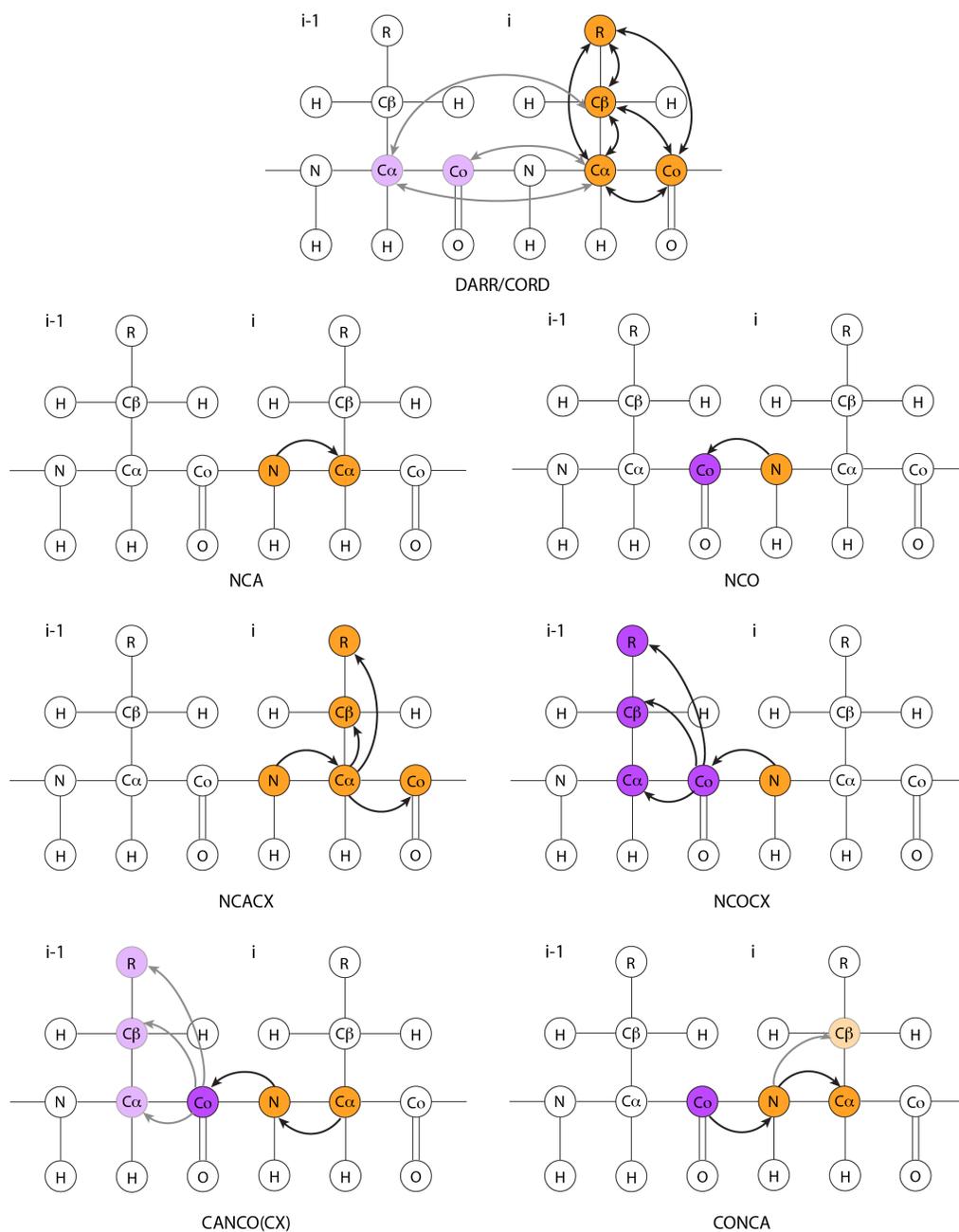


Figure 2.2 Magnetization transfer pathways of homonuclear and heteronuclear NMR experiments, including DARR/CORD, NCA, NCO, NCACX, NCOCX, CANCO(CX) and CONCA. Atoms in residue ( $i$ ) are labeled in orange, and the ones in the neighboring residue ( $i-1$ ) are labeled in purple. Strong transfers are marked with a black arrow, while weaker transfers or transfers that require longer mixing time are marked with gray arrows and lightly colored atoms.

Currently, the primary challenge in structure determination of large protein assemblies using solid-state NMR is to achieve high sensitivity and resolution as well as to reduce experimental acquisition times. Non-uniform sampling (NUS)<sup>51</sup> is one method to enhance sensitivity by reducing the number of points sampled in indirect dimensions with an appropriate NUS sampling schedule for data collection in multidimensional NMR experiments. The acquired points are chosen using strategically designed sampling schedules and are distributed nonlinearly. To attain time-domain sensitivity enhancements, the NUS schedules need to be random, weighted by a decaying function (e.g., exponential). The frequency-domain data are then calculated by time-domain signal reconstruction any suitable protocol, such as maximum entropy algorithms or iterative soft thresholding or  $l_1/l_2$ -norm minimization, rather than discrete Fourier transform.<sup>52,53</sup> The NUS method is especially beneficial to multi-dimensional heteronuclear-detected experiments for protein structure determination in challenging systems, such as disordered proteins<sup>54,55</sup>, low-sensitivity macromolecular assemblies and proteins studied by in-cell NMR<sup>56</sup>.

### **2.3.3 Secondary Structure Analysis**

In an NMR spectrum of a protein, chemical shifts are characteristic of different secondary structures, including helices, sheets and coils. Based on the chemical shifts determined from resonance assignments, protein secondary structures can be immediately derived. This step is essential for analyzing structural features of proteins. There are several programs, such as CSI 2.0<sup>57</sup>, CSI 3.0<sup>58</sup>, TALOS+<sup>59</sup>, TALOS-N<sup>60</sup>, and PLUQ<sup>61</sup> that have been developed for predicting the secondary structures and backbone torsion angles based on chemical shifts. In particular, CSI 2.0 and CSI 3.0

are based on Chemical Shift Index (CSI)<sup>62</sup>. The TALOS+ and TALOS-N packages are hybrid systems for empirical prediction that are based on an artificial neural network (ANN), and provide protein backbone and side-chain torsion angles using chemical shift assignments for protein sequence. Both CSI and TALOS programs require a series of sequential resonance assignments (at least 3 sequential residues) to identify the secondary structure accurately. When the complete resonance assignments for the entire protein sequence are unavailable, PLUQ<sup>61</sup> is an alternative program that provides rapid prediction of the amino acid type and secondary structure for the residues or short fragments by using the <sup>13</sup>C and <sup>15</sup>N chemical shifts in the Biological Magnetic Resonance Bank (BMRB) database.

For certain proteins, there are 3D structures that have been solved by X-ray crystallography or cryo-EM. One can use atomic coordinates in PDB files to predict NMR chemical shifts by programs such as SHIFTX2<sup>63</sup> and SPARTA+<sup>64</sup>. The calculated NMR chemical shifts can assist in the manual resonance assignments of large proteins, as well as aid in structure refinement and validation. However, it is still not reliable to quantify the predicted values due to the errors of the prediction programs.

## **2.4 Protein Dynamics by MAS NMR**

Proteins are dynamic rather than static entities. Motions in proteins modulate their functions. Characterization of protein dynamics is thus necessary for elucidating the biological mechanisms. MAS NMR is a powerful tool to probe protein dynamics at atomic resolution on a very wide range of timescales through diverse types of experiments.

### 2.4.1 NMR Timescale of Protein Motions

During biological processes, the participating biomacromolecules are undergoing constant fluctuations rather than remaining static in order to execute their functions. Proteins are flexible and play central roles in the cellular function. They experience both local as well as global motions on a broad range of timescales from femtoseconds to seconds.<sup>65</sup> As shown in Fig 2.3, local fluctuations are mainly associated with the timescales of femto- to nanoseconds and include bond vibrations, methyl group rotations as well as conformational changes in side chains and loop regions. Global motions, such as backbone fluctuations, domain motions and protein folding happen on slower timescales of nanoseconds to seconds.

Among the diverse methods that have been utilized to investigate protein motions, NMR spectroscopy is one of the most potent, since it can provide atomic-resolution information about assorted motions at timescales from picoseconds to seconds by probing specific nuclei (e.g.  $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^1\text{H}$ ,  $^2\text{H}$ ,  $^{19}\text{F}$ ,  $^{31}\text{P}$ ). Solution NMR has been extensively applied to protein dynamics studies in the liquid state.<sup>10,66,67</sup> In the past few decades, MAS solid-state NMR has been used to study protein dynamics in the cases where overall rotational tumbling is absent (“solid” state), such as protein assemblies, microcrystalline proteins and amyloid fibrils<sup>68,69</sup>. Several MAS NMR experimental methods to study dynamics are illustrated in Fig 2.3. Averaged anisotropic lineshapes, such as dipolar coupling lineshapes and chemical shift anisotropy (CSA) lineshapes are sensitive to the timescales from pico- to microseconds.  $T_1$  relaxation rates ( $^{13}\text{C}$  or  $^{15}\text{N}$ ) report on protein dynamics on the pico- to nanoseconds.<sup>70</sup> Spin-lattice relaxation in the rotating frame,  $T_{1\rho}$  is sensitive probe for measuring molecular motions on the timescales of micro- to milliseconds.<sup>71</sup> The relaxation experiments report dynamics on rates, and lineshape experiments report

dynamics on motional amplitudes and symmetries of motion. Both types of experiments are very useful for understanding the full picture of protein dynamics. Cross polarization (CP)<sup>72</sup> is a key approach to transfer magnetization from nuclei with high gyromagnetic ratio (e.g. <sup>1</sup>H) to nuclei with low gyromagnetic ratio (e.g. <sup>13</sup>C and <sup>15</sup>N) in close proximity through heteronuclear dipolar coupling interactions. CP experiments are sensitive to motions occurring on nano- to millisecond timescales. For slower motions from milliseconds to seconds, exchange NMR experiments, such as CODEX (center-band only detection of exchange)<sup>73,74</sup> can be executed. Additionally, the T<sub>1</sub> relaxation time constants of <sup>1</sup>H and <sup>13</sup>C generally correspond to timescale ranges from milliseconds to several seconds in protein samples.

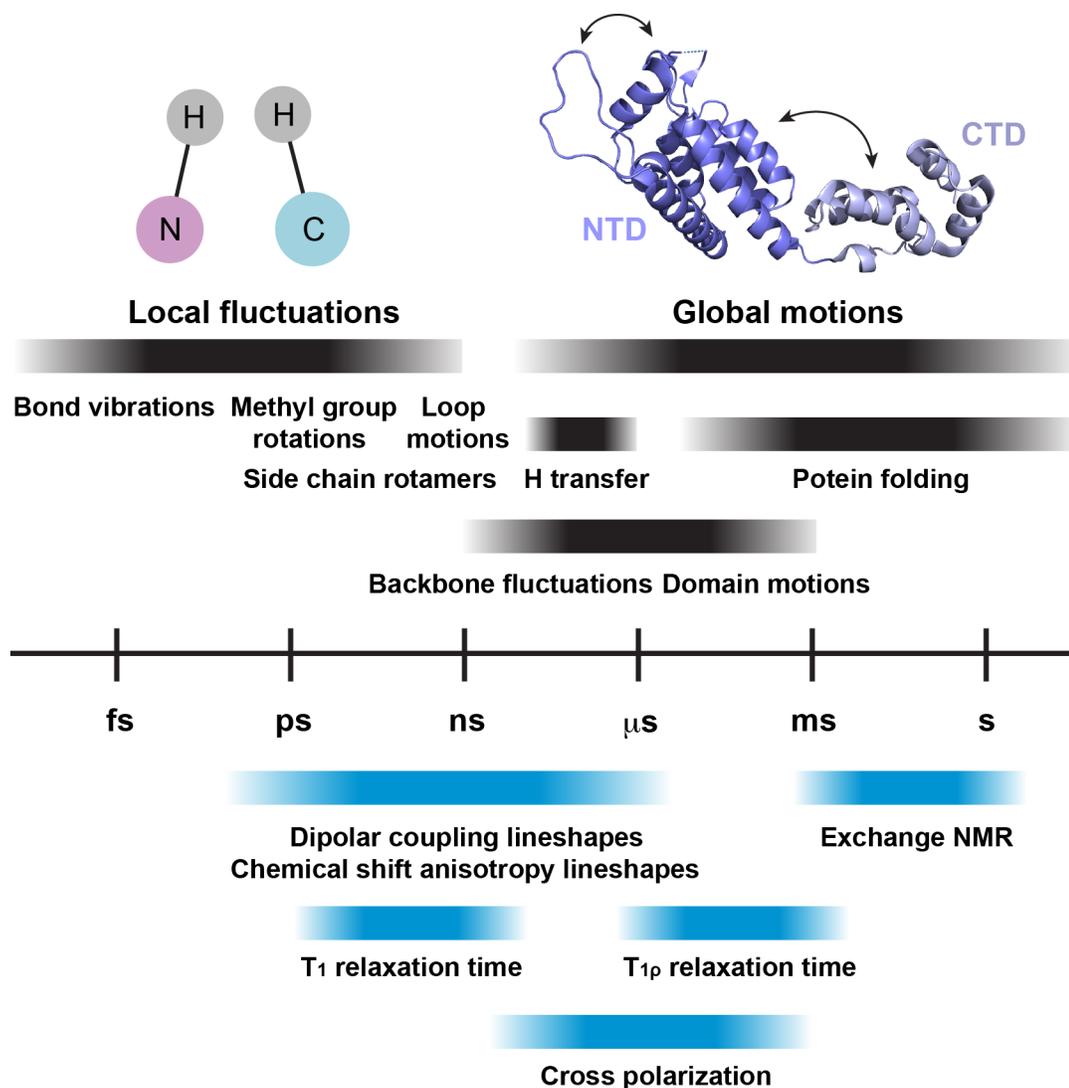


Figure 2.3 Timescales of protein dynamic processes and MAS NMR measurements of dynamics. The figure is adapted from REFERENCES<sup>65,68</sup>. Permissions for reuse in this dissertation is granted by Springer Nature, copyright 2007, and Elsevier B.V., copyright 2005.

#### 2.4.2 Experiments for Probing Motions on Milli- to Microsecond Timescales

The dynamics of milli- to microsecond timescale include protein backbone changes and domain motions, which play significant roles in biological events. The CP process (e.g.  $^1\text{H}$ - $^{13}\text{C}$ ,  $^1\text{H}$ - $^{15}\text{N}$ ,  $^{15}\text{N}$ - $^{13}\text{C}$ ) relies on dipolar coupling interactions, and

dipolar-coupling experiments are a major method to evaluate protein dynamics on milli- to microsecond timescales. In the 2D homo- and heteronuclear correlation spectra, the signals of residues in mobile regions exhibit peak broadening or missing peaks. For example, changes in peak intensity and linewidth in the 2D NCA spectra reveal protein backbone motions on milli- to microsecond timescales. In chapter 3 and 4, these experiments are performed to probe the dynamic behaviors of different HIV-1 protein assemblies so as to examine the relationship between protein dynamics and functions.

### **2.4.3 Experiments for Probing Motions on Micro- to Nanosecond Timescales**

Solid-state NMR spectroscopy provides anisotropic information that depends on the orientation of molecular frame with respect to the magnetic field, which is modulated by local motions. Lineshape measurements<sup>68</sup> such as heteronuclear  $^1\text{H}$ - $^{15}\text{N}/^1\text{H}$ - $^{13}\text{C}$  dipolar couplings and CSA lineshape experiments are commonly used to probe protein motions on micro- to nanosecond timescales by measuring dipolar couplings and CSA values.

The dipolar couplings between two nuclei depend on the internuclear distance and their motions. CSA is an orientation-dependent chemical shift tensor that arises from the asymmetric distribution of charge around nuclei, which can be affected by molecular motions. If the molecular motions are faster than the magnitude of anisotropic interactions, the dipolar couplings and CSA tensors are partially averaged, resulting in the reduction of the magnitude of the anisotropic interaction, and depending on the motional symmetry, the altered asymmetry parameter of the respective tensor.<sup>75</sup> To re-introduce these anisotropic features under magic angle spinning, suitable recoupling schemes have been developed. As will be introduced in

later chapters, two types of 3D RN-symmetry<sup>76</sup> dipolar recoupling experiments, DIPSHIFT<sup>77</sup> and PARS<sup>78</sup> experiment, have been employed to study dynamics of HIV-1 CA and CA-SP1 protein assemblies. CSA tensors can be extracted from a series of sidebands in slow MAS experiments as well as the 2D RN-symmetry based CSA experiments<sup>79</sup> that re-introduce CSA interaction under fast MAS frequencies.

The site-specific molecular motions affect the experimental lineshapes of dipolar coupling and CSA. The deviation of dipolar couplings can be quantified by an order parameter  $S_D$ , which is the ratio of measured dipolar coupling to the rigid limit value, and which ranges from 1 to 0 with one corresponding to the static site.<sup>19</sup> The rigid limit values of one-bond  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  dipolar coupling are 11.34<sup>80</sup> and 22.7<sup>81</sup> kHz, respectively, in the absence of any motions. For CSA interactions, CSA tensor can be expressed in a coordinate frame, and be described by three principal components (diagonal elements in the principal axis system). The parameters of anisotropy, reduced anisotropy and asymmetry are used to define the CSA line shapes, which can help determine the changes that are induced by motions on micro- to nanosecond timescales.<sup>79</sup>

## 2.5 Dynamic Nuclear Polarization Technique

In order to explore the application of biomolecular NMR to more challenging biological systems, it is essential to enhance spectral sensitivity and resolution, as well as shorten the experiment acquisition time. Dynamic nuclear polarization (DNP) is a method that provides significant sensitivity enhancements by transferring the magnetization from electron spins to nuclear spins. The electron spins are irradiated by microwave frequency pulses and have much larger polarizations.<sup>82-84</sup> DNP experiments are usually performed at cryogenic temperatures, 80-110 K, to improve

the relaxation process under high transfer efficiency. The full DNP pathways of a system containing  $^1\text{H}$  and  $^{13}\text{C}$  are illustrated in Fig 2.4<sup>85</sup>. The polarization is transferred from the electron spin (polarizing agent) directly and indirectly to the detected nucleus ( $^{13}\text{C}$ ), and a fraction of signal enhancements can be achieved. Theoretically, the maximum of signal enhancement is 658 folds for  $^1\text{H}$ , and 2617 folds for  $^{13}\text{C}$ .<sup>86</sup> When DNP MAS NMR spectroscopy is applied for structural biology, nitroxide radicals, such as TOTAPOL and AMUPol are commonly used as polarizing agents to gain high efficiency and stability.<sup>84,86</sup> With the combination of high-field DNP and MAS NMR, biological samples with lower concentration and larger size (e.g. membrane proteins<sup>87</sup>, amyloid fibrils<sup>88</sup> and protein assemblies<sup>89</sup>) can be studied in hours or days rather than several weeks.

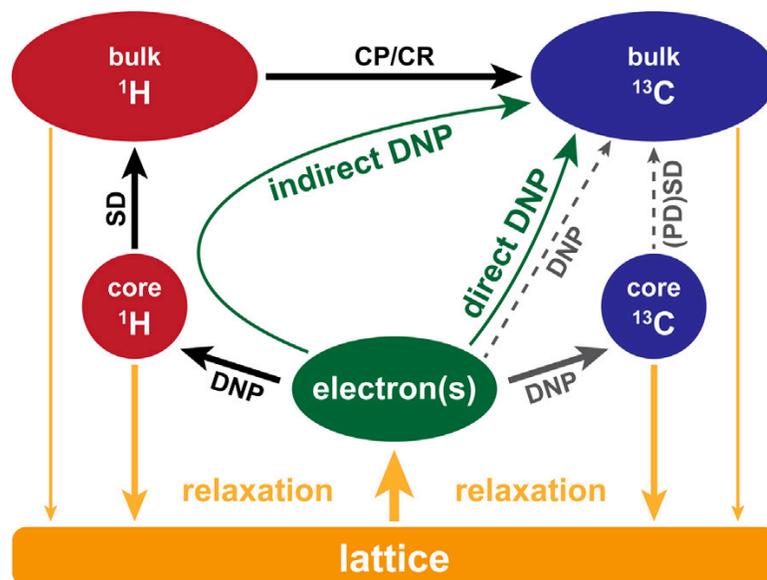


Figure 2.4 Schematic of DNP transfer pathways in a system containing  $^1\text{H}$  and  $^{13}\text{C}$ . Green arrows distinguish between indirect and direct DNP transfer.<sup>85</sup> Permissions for reuse in this dissertation is granted by Elsevier B.V., copyright 2017.

## REFERENCES

- (1) Li, Y.; Gupta, R.; Cho, J. H.; Raleigh, D. P. Mutational analysis of the folding transition state of the C-terminal domain of ribosomal protein L9: A protein with an unusual beta-sheet topology. *Biochemistry-Us* **2007**, *46* (4), 1013.
- (2) Nichols, N. M.; Evans, T. C. Mutational analysis of protein splicing, cleavage, and self-association reactions mediated by the naturally split Ssp DnaE intein. *Biochemistry-Us* **2004**, *43* (31), 10265.
- (3) Fontana, J.; Keller, P. W.; Urano, E.; Ablan, S. D.; Steven, A. C.; Freed, E. O. Identification of an HIV-1 Mutation in Spacer Peptide 1 That Stabilizes the Immature CA-SP1 Lattice. *Journal of Virology* **2016**, *90* (2), 972.
- (4) Pornillos, O.; Ganser-Pornillos, B. K.; Kelly, B. N.; Hua, Y.; Whitby, F. G.; Stout, C. D.; Sundquist, W. I.; Hill, C. P.; Yeager, M. X-ray structures of the hexameric building block of the HIV capsid. *Cell* **2009**, *137* (7), 1282.
- (5) Pornillos, O.; Ganser-Pornillos, B. K.; Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **2011**, *469* (7330), 424.
- (6) Gres, A. T.; Kirby, K. A.; KewalRamani, V. N.; Tanner, J. J.; Pornillos, O.; Sarafianos, S. G. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **2015**, *349* (6243), 99.
- (7) Zhao, G. P.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J. Y.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **2013**, *497* (7451), 643.
- (8) Holdgate, G.; Geschwindner, S.; Breeze, A.; Davies, G.; Colclough, N.; Temesi, D.; Ward, L. Biophysical methods in drug discovery from small molecule to pharmaceutical. *Methods Mol Biol* **2013**, *1008*, 327.
- (9) Costa, T. R. D.; Ignatiou, A.; Orlova, E. V. Structural Analysis of Protein Complexes by Cryo Electron Microscopy. *Methods Mol Biol* **2017**, *1615*, 377.

- (10) Kempf, J. G.; Loria, J. P. Protein dynamics from solution NMR - Theory and applications. *Cell Biochem Biophys* **2003**, 37 (3), 187.
- (11) Rosenzweig, R.; Kay, L. E. Solution NMR Spectroscopy Provides an Avenue for the Study of Functionally Dynamic Molecular Machines: The Example of Protein Disaggregation. *J Am Chem Soc* **2016**, 138 (5), 1466.
- (12) Mertens, H. D.; Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J Struct Biol* **2010**, 172 (1), 128.
- (13) Petoukhov, M. V.; Svergun, D. I. Applications of small-angle X-ray scattering to biomacromolecular solutions. *Int J Biochem Cell B* **2013**, 45 (2), 429.
- (14) Okamoto, K.; Sako, Y. Recent advances in FRET for the study of protein interactions and dynamics. *Current Opinion in Structural Biology* **2017**, 46, 16.
- (15) Perilla, J. R.; Goh, B. C.; Cassidy, C. K.; Liu, B.; Bernardi, R. C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. Molecular dynamics simulations of large macromolecular complexes. *Current Opinion in Structural Biology* **2015**, 31, 64.
- (16) van der Wel, P. C. A. New applications of solid-state NMR in structural biology. *Emerging Topics in Life Sciences* **2018**, 2 (1), 57.
- (17) Quinn, C. M.; Wang, M.; Polenova, T. NMR of Macromolecular Assemblies and Machines at 1 GHz and Beyond: New Transformative Opportunities for Molecular Structural Biology. *Methods Mol Biol* **2018**, 1688, 1.
- (18) Wang, M. Z.; Quinn, C. M.; Perilla, J. R.; Zhang, H. L.; Shirra, R.; Hou, G. J.; Byeon, I. J.; Suiter, C. L.; Ablan, S.; Urano, E. et al. Quenching protein dynamics interferes with HIV capsid maturation. *Nat Commun* **2017**, 8.
- (19) Lu, M. M.; Hou, G. J.; Zhang, H. L.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Perilla, J. R.; Langmead, C. J.; Hung, I.; Gor'kov, P. L. et al. Dynamic allostery governs cyclophilin A-HIV capsid interplay. *P Natl Acad Sci USA* **2015**, 112 (47), 14617.

- (20) Liu, C.; Perilla, J. R.; Ning, J. Y.; Lu, M. M.; Hou, G. J.; Ramalho, R.; Himes, B. A.; Zhao, G. P.; Bedwell, G. J.; Byeon, I. J. et al. Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat Commun* **2016**, *7*.
- (21) Fritz, M.; Quinn, C. M.; Wang, M. Z.; Hou, G. J.; Lu, X. Y.; Koharudin, L. M. I.; Struppe, J.; Case, D. A.; Polenova, T.; Gronenborn, A. M. Determination of accurate backbone chemical shift tensors in microcrystalline proteins by integrating MAS NMR and QM/MM. *Phys Chem Chem Phys* **2018**, *20* (14), 9543.
- (22) Fritz, M.; Quinn, C. M.; Wang, M. Z.; Hou, G. J.; Lu, X. G.; Koharudin, L. M. I.; Polenova, T.; Gronenborn, A. M. Toward Closing the Gap: Quantum Mechanical Calculations and Experimentally Measured Chemical Shifts of a Microcrystalline Lectin. *J Phys Chem B* **2017**, *121* (15), 3574.
- (23) Andrew, E. R.; Bradbury, A.; Eades, R. G. Nuclear Magnetic Resonance Spectra from a Crystal Rotated at High Speed. *Nature* **1958**, *182* (4650), 1659.
- (24) Polenova, T.; Gupta, R.; Goldbourn, A. Magic angle spinning NMR spectroscopy: a versatile technique for structural and dynamic analysis of solid-phase systems. *Anal Chem* **2015**, *87* (11), 5458.
- (25) Barbet-Massin, E.; Pell, A. J.; Retel, J. S.; Andreas, L. B.; Jaudzems, K.; Franks, W. T.; Nieuwkoop, A. J.; Hiller, M.; Higman, V.; Guerry, P. et al. Rapid Proton-Detected NMR Assignment for Proteins with Fast Magic Angle Spinning. *J Am Chem Soc* **2014**, *136* (35), 12489.
- (26) Struppe, J.; Quinn, C. M.; Lu, M. M.; Wang, M. Z.; Hou, G. J.; Lu, X. Y.; Kraus, J.; Andreas, L. B.; Stanek, J.; Lalli, D. et al. Expanding the horizons for structural analysis of fully protonated protein assemblies by NMR spectroscopy at MAS frequencies above 100 kHz. *Solid State Nucl Mag* **2017**, *87*, 117.
- (27) McIntosh, L. P.; Dahlquist, F. W. Biosynthetic Incorporation of <sup>15</sup>N and <sup>13</sup>C for Assignment and Interpretation of Nuclear Magnetic Resonance Spectra of Proteins. *Quarterly Reviews of Biophysics* **2009**, *23* (1), 1.

- (28) Verardi, R.; Traaseth, N. J.; Masterson, L. R.; Vostrikov, V. V.; Veglia, G. Isotope labeling for solution and solid-state NMR spectroscopy of membrane proteins. *Adv Exp Med Biol* **2012**, *992*, 35.
- (29) Hoogstraten, C. G.; Johnson, J. E. Metabolic labeling: Taking advantage of bacterial pathways to prepare spectroscopically useful isotope patterns in proteins and nucleic acids. *Concept Magn Reson A* **2008**, *32a* (1), 34.
- (30) Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. Fractional C-13 enrichment of isolated carbons using [1-C-13]- or [2-C-13]-glucose facilitates the accurate measurement of dynamics at backbone C-alpha and side-chain methyl positions in proteins. *J Biomol Nmr* **2007**, *38* (3), 199.
- (31) Hong, M.; Jakes, K. Selective and extensive C-13 labeling of a membrane protein for solid-state NMR investigations. *J Biomol Nmr* **1999**, *14* (1), 71.
- (32) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **2002**, *420* (6911), 98.
- (33) Akbey, U.; Lange, S.; Franks, W. T.; Linser, R.; Rehbein, K.; Diehl, A.; van Rossum, B. J.; Reif, B.; Oschkinat, H. Optimum levels of exchangeable protons in perdeuterated proteins for proton detection in MAS solid-state NMR spectroscopy. *J Biomol Nmr* **2010**, *46* (1), 67.
- (34) Chevelkov, V.; van Rossum, B. J.; Castellani, F.; Rehbein, K.; Diehl, A.; Hohwy, M.; Steuernagel, S.; Engelke, F.; Oschkinat, H.; Reif, B. H-1 detection in MAS solid-state NMR Spectroscopy of biomacromolecules employing pulsed field gradients for residual solvent suppression. *J Am Chem Soc* **2003**, *125* (26), 7788.
- (35) Asami, S.; Schmieder, P.; Reif, B. High Resolution H-1-Detected Solid-State NMR Spectroscopy of Protein Aliphatic Resonances: Access to Tertiary Structure Information. *J Am Chem Soc* **2010**, *132* (43), 15133.
- (36) Gans, P.; Hamelin, O.; Sounier, R.; Ayala, I.; Dura, M. A.; Amero, C. D.; Noirclerc-Savoye, M.; Franzetti, B.; Plevin, M. J.; Boisbouvier, J. Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic

Studies of High-Molecular-Weight Proteins. *Angew Chem Int Edit* **2010**, *49* (11), 1958.

- (37) Krausslich, H. G.; Facke, M.; Heuser, A. M.; Konvalinka, J.; Zentgraf, H. The Spacer Peptide between Human-Immunodeficiency-Virus Capsid and Nucleocapsid Proteins Is Essential for Ordered Assembly and Viral Infectivity. *Journal of Virology* **1995**, *69* (6), 3407.
- (38) Szeverenyi, N. M.; Sullivan, M. J.; Maciel, G. E. Observation of Spin Exchange by Two-Dimensional Fourier-Transform C-13 Cross Polarization-Magic-Angle Spinning. *Journal of Magnetic Resonance* **1982**, *47* (3), 462.
- (39) Takegoshi, K.; Nakamura, S.; Terao, T. C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* **2001**, *344* (5-6), 631.
- (40) Bennett, A. E.; Ok, J. H.; Griffin, R. G.; Vega, S. Chemical-Shift Correlation Spectroscopy in Rotating Solids - Radio Frequency-Driven Dipolar Recoupling and Longitudinal Exchange. *J Chem Phys* **1992**, *96* (11), 8624.
- (41) Hou, G.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T. Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J Magn Reson* **2013**, *232*, 18.
- (42) Yan, S.; Suiter, C. L.; Hou, G. J.; Zhang, H. L.; Polenova, T. Probing Structure and Dynamics of Protein Assemblies by Magic Angle Spinning NMR Spectroscopy. *Accounts Chem Res* **2013**, *46* (9), 2047.
- (43) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol Phys* **1998**, *95* (6), 1197.
- (44) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. Backbone and side-chain C-13 and N-15 signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 tesla. *ChemBiochem* **2001**, *2* (4), 272.
- (45) Bertini, I.; Emsley, L.; Felli, I. C.; Laage, S.; Lesage, A.; Lewandowski, J. R.; Marchetti, A.; Pierattelli, R.; Pintacuda, G. High-resolution and sensitivity

through-bond correlations in ultra-fast magic angle spinning (MAS) solid-state NMR. *Chem Sci* **2011**, *2* (2), 345.

- (46) Chen, L.; Kaiser, J. M.; Polenova, T.; Yang, J.; Rienstra, C. M.; Mueller, L. J. Backbone assignments in solid-state proteins using J-based 3D Heteronuclear correlation spectroscopy. *J Am Chem Soc* **2007**, *129* (35), 10650.
- (47) Lesage, A.; Auger, C.; Caldarelli, S.; Emsley, L. Determination of through-bond carbon-carbon connectivities in solid-state NMR using the INADEQUATE experiment. *J Am Chem Soc* **1997**, *119* (33), 7867.
- (48) Han, Y.; Hou, G. J.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Lipton, A. S.; Burton, S.; Hung, I.; Gor'kov, P. L.; Gan, Z. H. et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2013**, *135* (47), 17793.
- (49) Chen, L. L.; Kaiser, J. M.; Lai, J. F.; Polenova, T.; Yang, J.; Rienstra, C. M.; Mueller, L. J. J-based 2D homonuclear and heteronuclear correlation in solid-state proteins. *Magn Reson Chem* **2007**, *45*, S84.
- (50) Andronesi, O. C.; Mintzopoulos, D.; Struppe, J.; Black, P. M.; Tzika, A. A. Solid-state NMR adiabatic TOBSY sequences provide enhanced sensitivity for multidimensional high-resolution magic-angle-spinning H-1 MR spectroscopy. *Journal of Magnetic Resonance* **2008**, *193* (2), 251.
- (51) Paramasivam, S.; Suiter, C. L.; Hou, G. J.; Sun, S. J.; Palmer, M.; Hoch, J. C.; Rovnyak, D.; Polenova, T. Enhanced Sensitivity by Nonuniform Sampling Enables Multidimensional MAS NMR Spectroscopy of Protein Assemblies. *J Phys Chem B* **2012**, *116* (25), 7416.
- (52) Hyberts, S. G.; Arthanari, H.; Wagner, G. Applications of Non-Uniform Sampling and Processing. *Top Curr Chem* **2012**, *316*, 125.
- (53) Hoch, J. C.; Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S. Nonuniform Sampling and Maximum Entropy Reconstruction in Multidimensional NMR. *Accounts Chem Res* **2014**, *47* (2), 708.

- (54) Kazimierczuk, K.; Stanek, J.; Zawadzka-Kazimierczuk, A.; Kozminski, W. High-Dimensional NMR Spectra for Structural Studies of Biomolecules. *Chemphyschem* **2013**, *14* (13), 3015.
- (55) Roche, J.; Ying, J. F.; Maltsev, A. S.; Bax, A. Impact of Hydrostatic Pressure on an Intrinsically Disordered Protein: A High-Pressure NMR Study of alpha-Synuclein. *Chembiochem* **2013**, *14* (14), 1754.
- (56) Sakakibara, D.; Sasaki, A.; Ikeya, T.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; Walchli, M. et al. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **2009**, *458* (7234), 102.
- (57) Hafsa, N. E.; Wishart, D. S. CSI 2.0: a significantly improved version of the Chemical Shift Index. *J Biomol Nmr* **2014**, *60* (2-3), 131.
- (58) Hafsa, N. E.; Arndt, D.; Wishart, D. S. CSI 3.0: a web server for identifying secondary and super-secondary structure in proteins using NMR chemical shifts. *Nucleic Acids Res* **2015**, *43* (W1), W370.
- (59) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS plus : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol Nmr* **2009**, *44* (4), 213.
- (60) Shen, Y.; Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol Nmr* **2013**, *56* (3), 227.
- (61) Fritzscheing, K. J.; Yang, Y.; Schmidt-Rohr, K.; Hong, M. Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *J Biomol Nmr* **2013**, *56* (2), 155.
- (62) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The Chemical-Shift Index - a Fast and Simple Method for the Assignment of Protein Secondary Structure through Nmr-Spectroscopy. *Biochemistry-Us* **1992**, *31* (6), 1647.
- (63) Han, B.; Liu, Y. F.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *J Biomol Nmr* **2011**, *50* (1), 43.

- (64) Shen, Y.; Bax, A. SPARTA plus : a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J Biomol Nmr* **2010**, *48* (1), 13.
- (65) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450* (7172), 964.
- (66) Göbl, C.; Tjandra, N. Application of Solution NMR Spectroscopy to Study Protein Dynamics. *Entropy* **2012**, *14* (3), 581.
- (67) Kovermann, M.; Rogne, P.; Wolf-Watz, M. Protein dynamics and function from solution state NMR spectroscopy. *Q Rev Biophys* **2016**, *49*, e6.
- (68) Krushelnitsky, A.; Reichert, D. Solid-state NMR and protein dynamics. *Prog Nucl Mag Res Sp* **2005**, *47* (1-2), 1.
- (69) Schanda, P.; Ernst, M. Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules. *Prog Nucl Mag Res Sp* **2016**, *96*, 1.
- (70) Hologne, M.; Faelber, K.; Diehl, A.; Reif, B. Characterization of dynamics of perdeuterated proteins by MAS solid-state NMR. *J Am Chem Soc* **2005**, *127* (32), 11208.
- (71) Lewandowski, J. R.; Sass, H. J.; Grzesiek, S.; Blackledge, M.; Emsley, L. Site-Specific Measurement of Slow Motions in Proteins. *J Am Chem Soc* **2011**, *133* (42), 16762.
- (72) Pines, A.; Shattuck, T. W. C-13 Proton Nmr Cross-Polarization Times in Solid Adamantane. *J Chem Phys* **1974**, *61* (3), 1255.
- (73) deAzevedo, E. R.; Hu, W. G.; Bonagamba, T. J.; Schmidt-Rohr, K. Centerband-only detection of exchange: Efficient analysis of dynamics in solids by NMR. *J Am Chem Soc* **1999**, *121* (36), 8411.
- (74) Li, W. B.; McDermott, A. Investigation of slow molecular dynamics using R-CODEX. *Journal of Magnetic Resonance* **2012**, *222*, 74.
- (75) Hong, M.; Su, Y. C. Structure and dynamics of cationic membrane peptides and proteins: Insights from solid-state NMR. *Protein Sci* **2011**, *20* (4), 641.

- (76) Levitt, M. H. Symmetry in the design of NMR multiple-pulse sequences. *J Chem Phys* **2008**, *128* (5).
- (77) Hou, G. J.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. H-1-C-13/H-1-N-15 Heteronuclear Dipolar Recoupling by R-Symmetry Sequences Under Fast Magic Angle Spinning for Dynamics Analysis of Biological and Organic Solids. *J Am Chem Soc* **2011**, *133* (46), 18646.
- (78) Hou, G. J.; Lu, X. Y.; Vega, A. J.; Polenova, T. Accurate measurement of heteronuclear dipolar couplings by phase-alternating R-symmetry (PARS) sequences in magic angle spinning NMR spectroscopy. *J Chem Phys* **2014**, *141* (10).
- (79) Hou, G. J.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. Recoupling of chemical shift anisotropy by R-symmetry sequences in magic angle spinning NMR spectroscopy. *J Chem Phys* **2012**, *137* (13).
- (80) Yao, L. S.; Vogeli, B.; Ying, J. F.; Bax, A. NMR Determination of Amide N-H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements. *J Am Chem Soc* **2008**, *130* (49), 16518.
- (81) Alkaraghoul, A. R.; Koetzle, T. F. Neutron-Diffraction Study of L-Phenylalanine Hydrochloride. *Acta Crystallogr B* **1975**, *31* (Oct15), 2461.
- (82) Carver, T. R.; Slichter, C. P. Polarization of Nuclear Spins in Metals. *Physical Review* **1953**, *92* (1), 212.
- (83) Overhauser, A. W. Polarization of Nuclei in Metals. *Physical Review* **1953**, *92* (2), 411.
- (84) Akbey, U.; Franks, W. T.; Linden, A.; Orwick-Rydmark, M.; Lange, S.; Oschkinat, H. Dynamic nuclear polarization enhanced NMR in the solid-state. *Top Curr Chem* **2013**, *338*, 181.
- (85) Lilly Thankamony, A. S.; Wittmann, J. J.; Kaushik, M.; Corzilius, B. Dynamic nuclear polarization for sensitivity enhancement in modern solid-state NMR. *Prog Nucl Magn Reson Spectrosc* **2017**, *102-103*, 120.

- (86) Sauvee, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P. Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency. *Angew Chem Int Ed Engl* **2013**, *52* (41), 10858.
- (87) Bajaj, V. S.; Hornstein, M. K.; Kreischer, K. E.; Sirigiri, J. R.; Woskov, P. P.; Mak-Jurkauskas, M. L.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G. 250GHz CW gyrotron oscillator for dynamic nuclear polarization in biological solid state NMR. *Journal of Magnetic Resonance* **2007**, *189* (2), 251.
- (88) Bayro, M. J.; Debelouchina, G. T.; Eddy, M. T.; Birkett, N. R.; MacPhee, C. E.; Rosay, M.; Maas, W. E.; Dobson, C. M.; Griffin, R. G. Intermolecular Structure Determination of Amyloid Fibrils with Magic-Angle Spinning and Dynamic Nuclear Polarization NMR. *J Am Chem Soc* **2011**, *133* (35), 13967.
- (89) Gupta, R.; Lu, M.; Hou, G.; Caporini, M. A.; Rosay, M.; Maas, W.; Struppe, J.; Suiter, C.; Ahn, J.; Byeon, I. J. et al. Dynamic Nuclear Polarization Enhanced MAS NMR Spectroscopy for Structural Analysis of HIV-1 Protein Assemblies. *J Phys Chem B* **2016**, *120* (2), 329.

## Chapter 3

### STRUCTURE CHARACTERIZATION OF HIV-1 CA TUBULAR AND SPHERICAL ASSEMBLIES BY MAS NMR

#### 3.1 Introduction

*In vivo*, approximately 1000-1500 copies of HIV-1 capsid (CA) protein arrange into a conical lattice to protect the viral genome in a mature virion.<sup>1,2</sup> The conical capsids exhibit structural polymorphism with varying sizes and shapes. *In vitro*, CA proteins assemble into a variety of morphologies, such as cones, spheres, and tubes under different assembly conditions.<sup>2,3</sup> The relationships between various morphologies, the stability of the capsid, and the viral infectivity remain elusive. A recent all-atom model of the mature capsid, derived from cryo-EM and MD simulations, showed that one common stoichiometry in a conical capsid is 216 hexamers and 12 pentamers, which are located at the end of the high-curvature lattice to close the structure<sup>1</sup>. The distribution of the 12 pentamers in the capsid lattice affects the overall size and symmetry of the capsid. Therefore, it is important to investigate the role of CA pentamers in capsid assembly and viral replication. While the studies of pentameric and hexameric units in the context of the overall capsid are challenging, model CA mutant constructs that assemble into pentamers and hexamers have been developed.<sup>4</sup>

In this chapter, I examined two covalently cross-linked CA constructs (Fig 3.1), A14C/E45C/W184A/M185A and N21C/A22C/W184A/M185A (NL4-3 strain), which form the hexamer and pentamer units, respectively. One limitation is that these

hexamer and pentamer are flat structures,<sup>2</sup> unlike the native capsid lattice exhibiting curvature, but still, understanding of the structural and dynamic features of these building blocks is instructive.

I prepared and investigated tubular assembly of cross-linked CA hexamers, as well as a spherical co-assembly formed by a mixture of CA hexamers and CA pentamers. The CA A14C/E45C/W184A/M185A protein forms cross-linked hexamers and assembles into tubes in the presence of 1-2.4 M NaCl. The cross-linked hexamer tubular assemblies exhibit increased stability and homogeneity compared to CA WT tubular assemblies, as illustrated by X-ray studies.<sup>4</sup> To further characterize the structure of cross-linked tubular assemblies, MAS NMR experiments were conducted for resonance assignments and conformational investigation. Interestingly, CA cross-linked hexamers and pentamers can co-assemble into a sphere morphology at a molar ratio of 10:1. <sup>13</sup>C-<sup>13</sup>C correlation MAS NMR experiments were employed to probe the structural difference between the pentamers in the context of the co-assembly and CA hexamers.

This work was in collaboration with Dr. Caitlin Quinn and Brent Runge in our lab, Professor Jinwoo Ahn and Professor Angela Gronenborn from the University of Pittsburgh, and Professor Christopher Aiken from Vanderbilt University. My contributions were the preparation of CA A14C/E45C/W184A/M185A and CA N21C/A22C/W184A/M185A samples for MAS NMR studies, performing MAS NMR experiments and the data analysis.

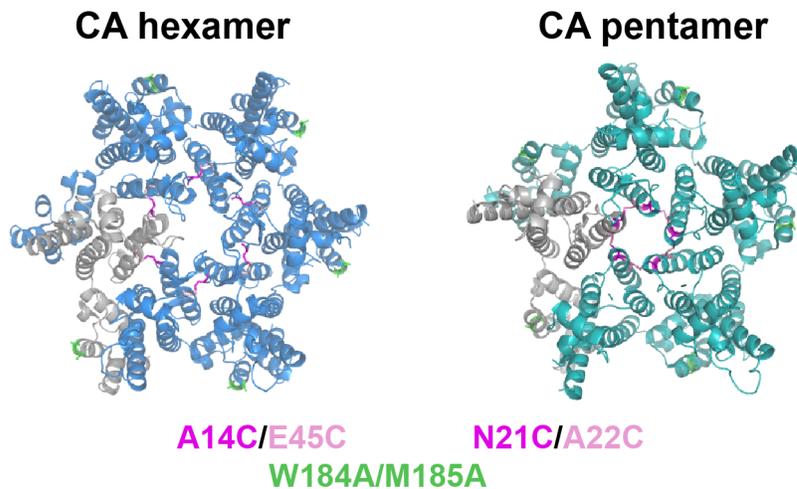


Figure 3.1 Structures of CA hexamer (A14C/E45C/W184A/M185A) and CA pentamer (N21C/A22C/W184A/M185A), NL4-3 strain. Mutation sites, cysteine (pink) and alanine (green), are labeled as sticks in 3D structures, showing cross-linked subunits. PDB: 3H4E and 3P05.<sup>4,5</sup>

## 3.2 Methods and Experiments

### 3.2.1 Materials

Common chemicals were purchased from Fisher Scientific or Sigma Aldrich. Chromatography columns were purchased from GE Healthcare.  $^{15}\text{NH}_4\text{Cl}$  and  $\text{U-}^{13}\text{C}_6$ -glucose were purchased from Cambridge Laboratories, Inc. The Rosetta (DE3) competent cells used for protein expression were purchased from Novagen. The HIV-1 CA cross-linked hexamer (A14C/E45C/W184A/M185A, NL4-3 strain) and pentamer (N21C/A22C/W184A/M185A, NL4-3 strain) constructs were prepared in the laboratory of our collaborator Dr. Christopher Aiken from Vanderbilt University Medical Center.

### 3.2.2 Expression and Purification of CA Cross-Linked Hexamer and Pentamer Proteins

Natural abundance,  $^{15}\text{N}$ -enriched, and uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -enriched CA cross-linked hexamer and pentamer proteins were prepared using the same protocol, as reported previously with modifications<sup>6,7</sup>. The gene was amplified and ligated into pET21 vector using NdeI and XhoI sites, and *E. coli* Rosetta (DE3) cells were transformed with the vector.

To express natural abundance CA cross-linked hexamer and pentamer proteins, cells were pre-cultured in 25 mL Luria-Bertani (LB) medium (25 g/L), and grown at 37 °C overnight. Cells were transferred into 1 L of LB medium, and grown at 37 °C until cells reached 0.7 OD. Then, cells were induced with 0.5 mM IPTG and expressed at 25 °C for 16-18 h. For  $^{15}\text{N}$ -labeled, and uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA cross-linked hexamer and pentamer proteins, cells were cultured in 2 L of Luria-Bertani (LB) medium, and grown at 37 °C until cells reached 1.2-1.4 OD. Following a centrifugation at RFC 4000 g at 4 °C for 30 min, cell pellets were washed with M9 medium that contains no nitrogen and carbon sources. Then cells were resuspended in 1 L of M9 medium that contains 2 g/L  $^{15}\text{NH}_4\text{Cl}$ , 2 g/L U- $^{13}\text{C}_6$ -glucose, basal vitamins,  $\text{CaCl}_2$ ,  $\text{MgSO}_4$ . After incubation at 25 °C for 1 h, cells were grown to 1.0-1.2 OD, induced with 0.8 mM IPTG and expressed at 25 °C for 16-18 h. To harvest, following a centrifugation at RFC 4000 g at 4 °C for 30 min, cells were resuspended in a 25 mM sodium phosphate buffer (pH 7.0) and stored at -80 °C.

CA A14C/E45C/W184A/M185A and CA N21C/A22C/W184A/M185A monomer proteins were purified by the same protocol. Cells were thawed and sonicated at 35% of amplitude with 10 s on and 10 s off for 20 min in an ice bath. The lysed cells were centrifuged at 27,000 g at 4 °C for 1 h. The pH of the supernatant was

adjusted to 5.8 with acetic acid, and the conductivity was reduced to below 2.5 ms/cm with de-ionized water. Following another centrifugation at 27,000 g at 4 °C for 1 h, the final supernatant was loaded onto a cation exchange column (HiTrap SP HP, 5 mL) and eluted with a 0-1 M NaCl gradient in buffer containing 25 mM sodium phosphate (pH 5.8), 1 mM DTT, and 0.02% NaN<sub>3</sub>. Concentrated protein fractions were further purified using a size-exclusion column (HiLoad 26/600 Superdex 75 pg) equilibrated with a buffer containing 25 mM sodium phosphate (pH 7.5), 1 mM DTT, and 0.02% NaN<sub>3</sub>. The purity of the two proteins (monomer) was assessed by SDS-PAGE gel after running through each column (Fig 3.2). The typical yield for natural abundance and <sup>15</sup>N-labeled CA protein is 30-40 mg/L, for uniformly <sup>13</sup>C, <sup>15</sup>N-enriched CA protein is 10-15 mg/L.

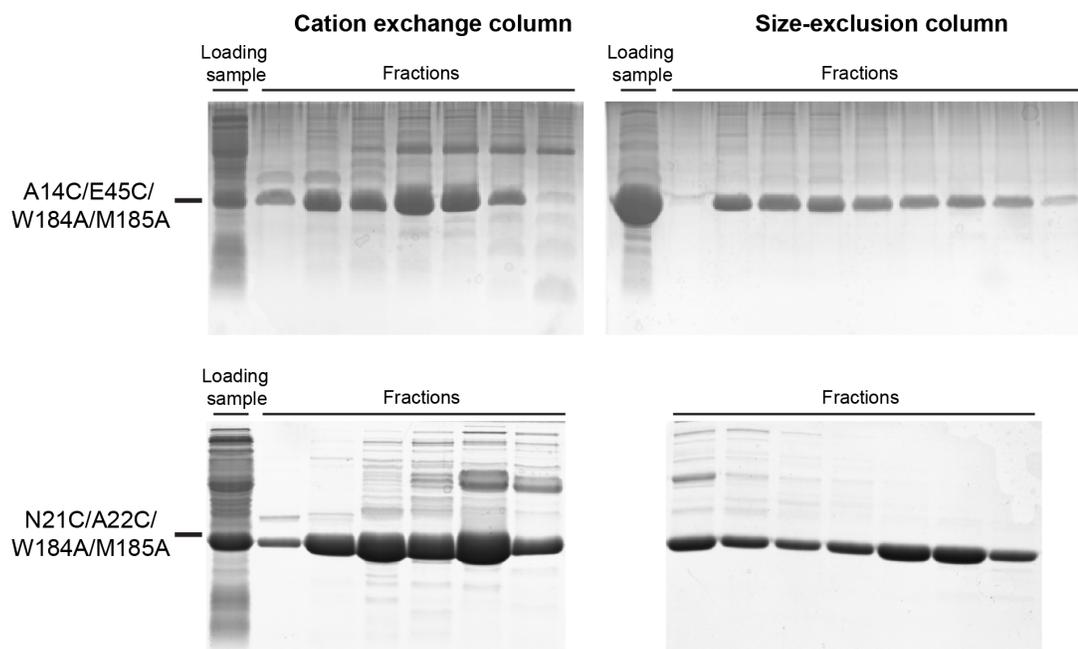


Figure 3.2 Purification of monomer (25.6 kDa) CA A14C/E45C/W184A/M185A and N21C/A22C/W184A/M185A proteins assessed by SDS-PAGE gel.

### 3.2.3 Preparation of Cross-Linked CA Hexamer and Pentamer Assemblies

CA A14C/E45C/W184A/M185A and CA N21C/A22C/W184A/M185A yield cross-linked CA hexamers and CA pentamers, respectively, upon the formation of disulfide bonds between cysteines in the two neighboring CA monomers, as shown in Fig 3.1. Redox reagents, such as DTT, can prevent the formation of disulfide bonds.<sup>8</sup> Therefore, CA hexamer proteins were buffer exchanged into 25 mM sodium phosphate buffer (pH 7.5) without DTT or other redox reagents overnight to achieve complete cross-linking. CA cross-linked hexamers were assembled from the concentrated protein solution (26 mg/mL) in 25 mM phosphate buffer (pH 7.5, no DTT) upon the addition of highly concentrated NaCl (2 M or 4.8 M) in 50 mM Tris buffer (pH 8.0) at a volume ratio of 1:1, followed by incubation at 37 °C for 1 h and 4 °C overnight. The final NaCl concentration in assemblies was 1 M or 2.4 M. CA cross-linked pentamers were assembled using the same protocol as above. To examine the oligomerization in the CA cross-linked hexamers and pentamers, the free protein solution and protein assemblies were studied by non-reducing SDS-PAGE gel (Fig 3.3). TEM images of CA cross-linked assemblies are shown in section 3.2.6. In free pentamer (Fig 3.3, Lane 1) and free hexamer protein solutions (Fig 3.3, Lane 3), although hexamer or pentamer existed as the major species, other oligomers, such as tetramer, trimer, dimer, and monomer, are present. By adding 2.4 M NaCl buffer, CA cross-linked pentamer assemblies (Fig. 3.3, Lane 2) and hexamer assemblies (Fig. 3.3, Lane 4) were pelleted, and the pellet contained a smaller population of other oligomers. Specifically, CA hexamer assemblies with 2.4 M NaCl only contained less than 1 % of other oligomers (Fig 3.3, Lane 4).

To prepare solutions of clean cross-linked hexamer and cross-linked pentamer proteins, further purification steps are required. Purified CA monomer proteins first

need to be cross-linked by buffer exchange into 25 mM sodium phosphate buffer (pH 7.5-8.0) without DTT overnight, and then the clean cross-linked hexamer (~150 kDa) or pentamer (~125 kDa) can be separated from other oligomers by a size-exclusion column (HiLoad 16/600 Superdex 200 pg). This protocol for preparing clean cross-linked hexamer and cross-linked pentamer proteins was developed with Brent Runge in our lab.

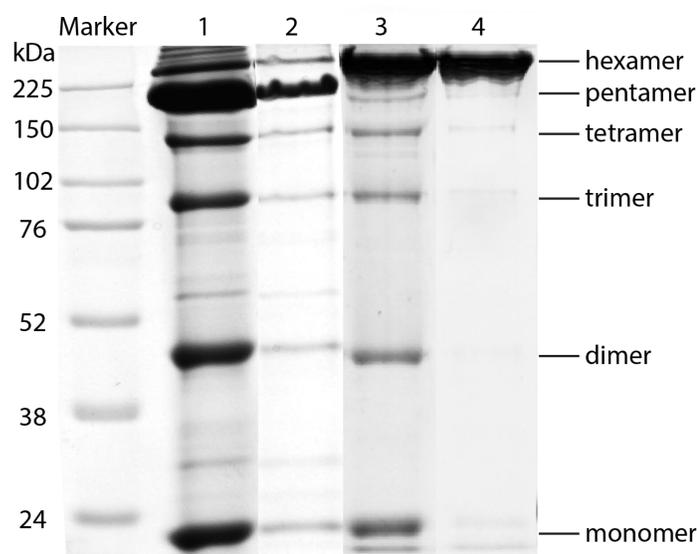
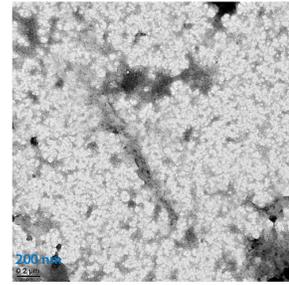
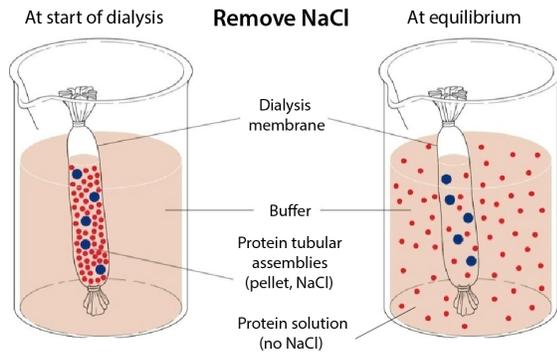


Figure 3.3 Oligomerization of CA cross-linked hexamer and pentamer assemblies were characterized by non-reducing SDS-PAGE gel with molecular weight (MW) marker. Lane 1 is CA cross-linked pentamer protein solution (no DTT); Lane 2 is CA cross-linked pentamer assembly (pelleted with 2.4 M NaCl); Lane 3 is CA cross-linked hexamer protein solution (no DTT); Lane 4 is CA cross-linked hexamer assembly (pelleted with 2.4 M NaCl). The mismatch of MW marker and protein is due to the fact the current MW marker was not calibrated for this large oligomer assay, and CA hexamer and pentamer run slower than the MW maker.

### **3.2.4 The Reversible Assembly of Cross-Linked CA Hexamers and the Effect of NaCl**

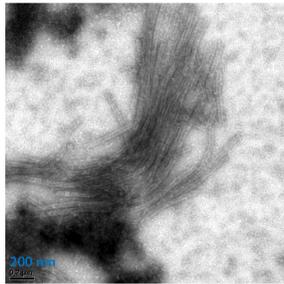
To investigate the assembly process of the cross-linked CA hexamer further, dialysis trials that remove NaCl from protein assemblies were performed, as shown in Fig 3.4a. 10 mg/mL CA hexamer tubular assemblies (1 M or 2.4 M NaCl) were packed in a dialysis bag and dialyzed against 25 mM sodium phosphate buffer (pH 7.5) overnight. As illustrated by TEM images in Fig 3.4b, CA hexamer assemblies that contain 1 M NaCl disassembled into protein solution, remaining as single cross-linked hexamers. By adding NaCl assembly buffer, the individual cross-linked hexamers re-assembled into homogenous tubes only with 2.4 M NaCl, not with 1.0 M NaCl buffer. The detailed TEM characterization is described in section 3.2.6. Conclusively, CA hexamer assemblies prepared in 1 M NaCl undergo a disassembly upon removal of salt and a re-assembly in the presence of 2.4 M NaCl. In contrast, CA hexamer assemblies with 2.4 M NaCl do not disassemble completely during the dialysis process. Oligomerization during these events was characterized by non-reducing SDS-PAGE gel (Fig 3.5). By comparing Lane 1 and Lane 4 of CA cross-linked hexamer protein solution samples before and after dialysis process, it was concluded that much purer CA cross-linked hexamers existed in protein solution after removing NaCl. As shown in Lane 5, the re-assemblies contained pure hexamers without any other oligomers.

a) Schematic of dialysis

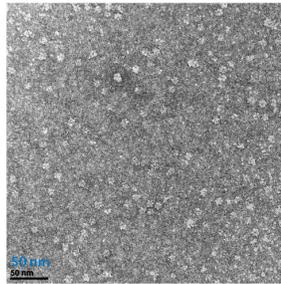


No re-assemblies  
5 mg/mL protein, 1.0 M NaCl

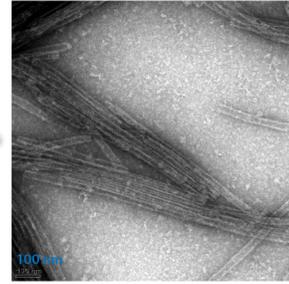
b) CA hexamer re-assembly



Hexamer assemblies  
10 mg/mL protein, 1.0 M NaCl



Single cross-linked hexamers  
10 mg/mL protein



Hexamer re-assemblies  
5 mg/mL protein, 2.4 M NaCl

Figure 3.4 (a) Schematic diagram of dialysis process to remove NaCl from CA cross-linked assemblies. (b) TEM images of CA hexamer illustrate the process of disassembly and re-assembly.

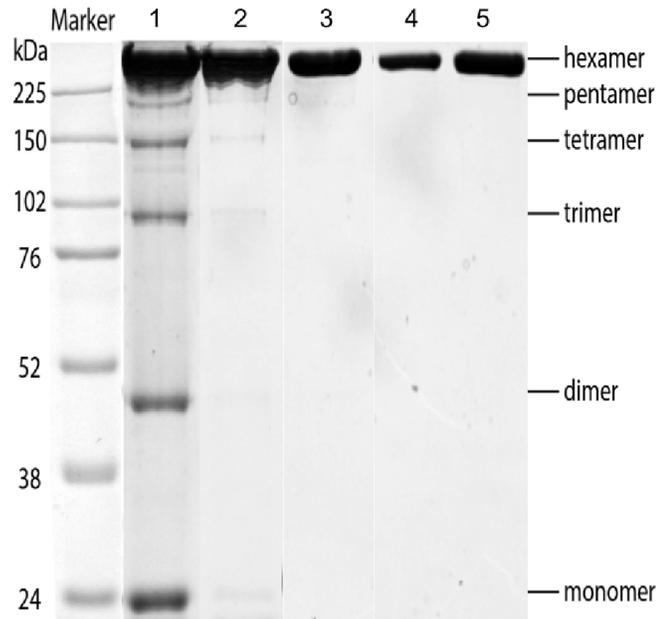


Figure 3.5 Oligomerization of CA cross-linked hexamer assemblies was characterized by non-reducing SDS-PAGE gel with molecular weight (MW) marker. Lane 1 is CA hexamer protein solution (no DTT); Lane 2 is CA hexamer assembly pellet (2.4 M NaCl); Lane 3 is CA hexamer assembly pellet (1 M NaCl); Lane 4 is CA cross-linked hexamer protein solution after removing NaCl; Lane 5 is CA hexamer re-assembly pellet with 2.4 M NaCl.

### 3.2.5 Preparation of CA Cross-Linked Hexamer/Pentamer Co-Assemblies

To prepare CA hexamer/pentamer co-assemblies, CA pentamers ( $U\text{-}^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled) were pre-assembled in 25 mM phosphate buffer containing 2.4 M NaCl (pH 6.5), followed by incubation at 37 °C for one hour and overnight at 4 °C. CA hexamers (natural abundance) protein solution was then added to the CA pentamer assemblies with a final molar ratio of 10:1 and 20:1 (hexamers:pentamers), respectively. After incubating the mixed protein solution on a rotating platform at 25 °C for 1 h, CA hexamer and pentamer proteins were co-assembled with a 2.4 M NaCl, 25 mM phosphate solution (pH 6.5) and incubated overnight at 4 °C. The CA

hexamer/pentamer co-assemblies of 20:1 molar ratio formed tubes (Fig 3.6a, bottom). Under a 10:1 ratio, the co-assemblies form mixture of tubes and spheres, and the spherical assemblies were displayed in Fig 3.6a, top. The detailed TEM characterization was described in section 3.2.6. Co-sedimentation assay of CA hexamer/pentamer (10:1) co-assemblies was employed, which indicates that both pentamer and hexamer exist in the pellet with approximately 60 % co-assembly efficiency.

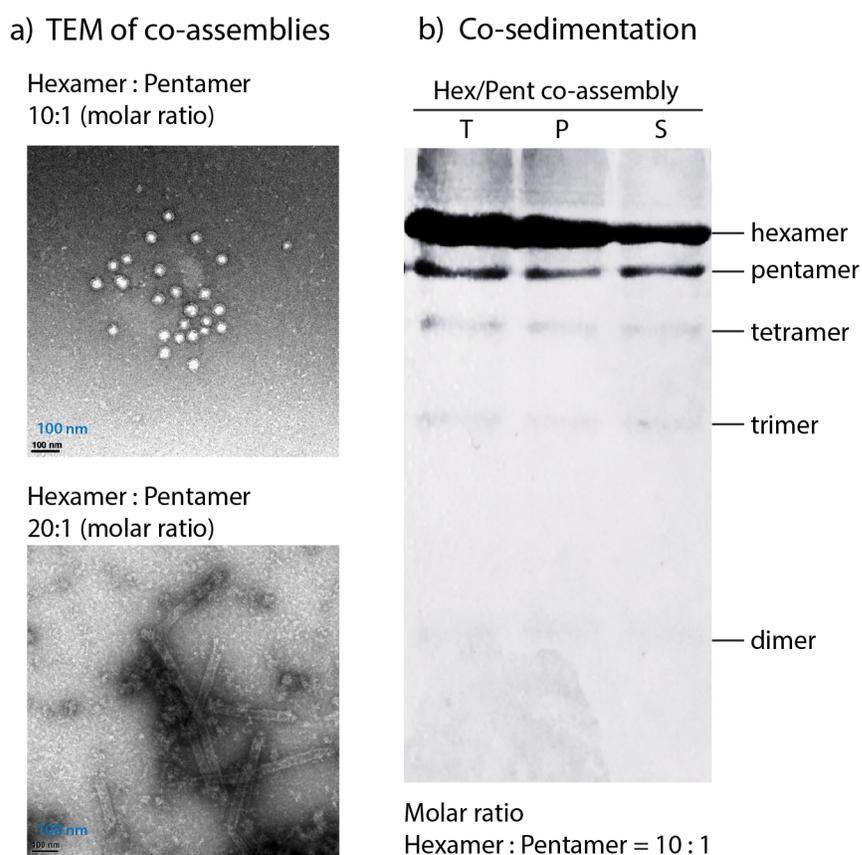


Figure 3.6 (a) TEM images of CA hexamer/pentamer co-assemblies, molar ratio of 10:1 and 20:1. (b) Co-sedimentation assay of CA hexamer/pentamer (10:1) co-assemblies. T: total, P: pellet, S: supernatant.

### 3.2.6 Transmission Electron Microscopy (TEM)

The morphologies of CA cross-linked free protein and assemblies were characterized by transmission electron microscopy (TEM). TEM images were collected at the Delaware Biotechnology Institute (DBI) in the Bioimaging Center. The images were acquired with a Zeiss Libra 120 transmission electron microscope operating at 120 kV, and equipped with a Gatan Ultrascan 1000 2k x 2k charge-coupled device camera. CA tubular and conical assemblies were stained with uranyl acetate (0.5-1% w/v), deposited onto 400 mesh, formvar/carbon-coated copper grids, and dried for 45 min in the air. The copper grids were glow discharged prior to staining, so that the assemblies are uniformly spread on the grid surface and adhere to it.

As shown in Fig 3.7, CA cross-linked hexamer protein assembled into tubes in the presence of 1 M and 2.4 M NaCl. CA cross-linked pentamer protein assembled into spherical particles in the presence of 1 M and 2.4 M NaCl, although the efficiency of the assembly was much lower than that for the CA hexamer protein.

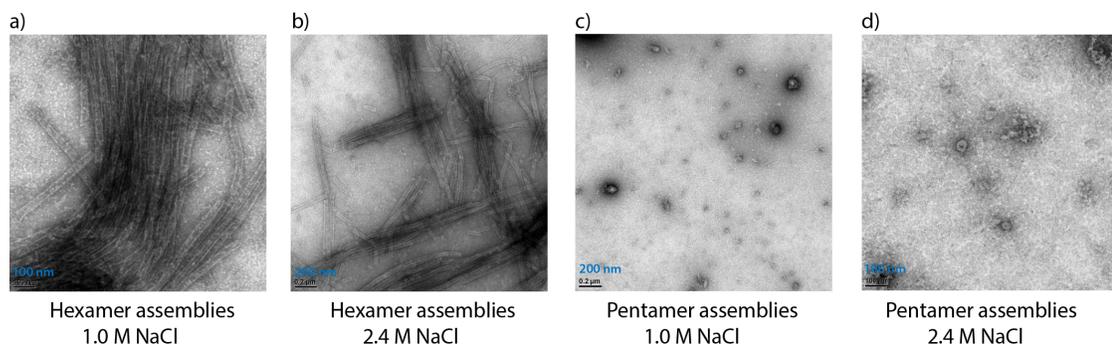


Figure 3.7 TEM images of CA cross-linked hexamer tubes (a-b) and pentamer spheres (c-d) with 1 M and 2.4 M NaCl.

### 3.2.7 MAS NMR Experiments

$U\text{-}^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA A14C/E45C/W184A/M185A proteins were assembled from a 26 mg/mL protein solution in 25 mM phosphate buffer (pH 7.5, no DTT) that was added to the same volume of 4.8 M NaCl in 50 mM Tris buffer (pH 8.0), followed by incubation at 37 °C for 1 h and 4 °C overnight. The final NaCl concentration in assemblies was 2.4 M. 53 mg of the CA hexamer tubular assemblies were pelleted at 10,000 g and packed into a Bruker thin wall 3.2 mm rotor. The co-assemblies of CA hexamer and pentamer co-assemblies containing 2.4 M NaCl buffer (pH 6.5) (preparation method was mentioned in section 3.2.5) were pelleted at 12,000 g, and 53 mg of the CA co-assemblies were packed into a Bruker thin wall 3.2 mm rotor.

All two-dimensional  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  spectra were acquired on a Bruker 20.0 T spectrometer operating at Larmor frequencies of 850.4 MHz ( $^1\text{H}$ ), 213.8 MHz ( $^{13}\text{C}$ ) and 86.2 MHz ( $^{15}\text{N}$ ), with a Bruker 3.2 mm E-free HCN probe at a sample temperature of  $4 \pm 1$  °C.  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts were REFERENCESd with respect to the external standards adamantane and ammonium chloride, respectively. MAS NMR spectra were collected at the MAS frequency of  $14.000 \pm 0.002$  kHz, regulated by a Bruker MAS controller. For all experiments, the typical pulse lengths were 2.5-2.8  $\mu\text{s}$  ( $^1\text{H}$ ), 3.5-3.8  $\mu\text{s}$  ( $^{13}\text{C}$ ), and 3.7-4.4  $\mu\text{s}$  ( $^{15}\text{N}$ ).  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  CP was performed with a linear amplitude ramp of 90-110%, with the  $^1\text{H}$  radio frequency (rf) field of 80-85 kHz matched to Hartmann-Hahn conditions at the first spinning sideband. Typical  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  contact times were 0.8-1.1 ms and 1.5-2 ms, respectively. Typical double cross polarization (DCP) power levels of 30-42 kHz on  $^{13}\text{C}$  and 55-60 kHz on  $^{15}\text{N}$  were used with optimized  $^{15}\text{N}$ - $^{13}\text{C}$  DCP contact time of 5-5.5 ms. SPINAL-64 decoupling ( $\sim 100$  kHz) was applied during the evolution and acquisition periods. The  $^1\text{H}$  field strength during CORD was 12.5 kHz, and the CORD

mixing time was 50 ms. For the co-assemblies of the natural abundance hexamers with U-<sup>13</sup>C, <sup>15</sup>N labeled pentamer (10:1 ratio), the <sup>13</sup>C-<sup>13</sup>C 2D CORD spectrum was acquired with non-uniform sampling (NUS<sup>9</sup>) with 40% sampling and a 2T<sub>2</sub> bias (144 complex points).

All spectra were processed in TopSpin and in NMRpipe<sup>10</sup> and analyzed using SPARKY<sup>11</sup> and CCPNMR<sup>12</sup>. For 2D and 3D data sets, 30°, 45°, 60° or 90° shifted sine bell apodization followed by a Lorentzian-to Gaussian transformation was applied in both dimensions. Forward linear prediction to twice the number of the original data points was used in the indirect dimension in some data sets followed by zero filling to twice the total number of points. Spectra were assigned by comparison with CA resonance assignments as reported previously.<sup>3</sup>

### **3.3 Results**

#### **3.3.1 Resonance Assignments of CA Cross-linked Tubular Assemblies**

HIV-1 CA hexamer protein (A14C/E45C/W184A/M185A) contains 231 amino acids and 4 mutations compared with CA wild type, see sequence in Fig 3.8.

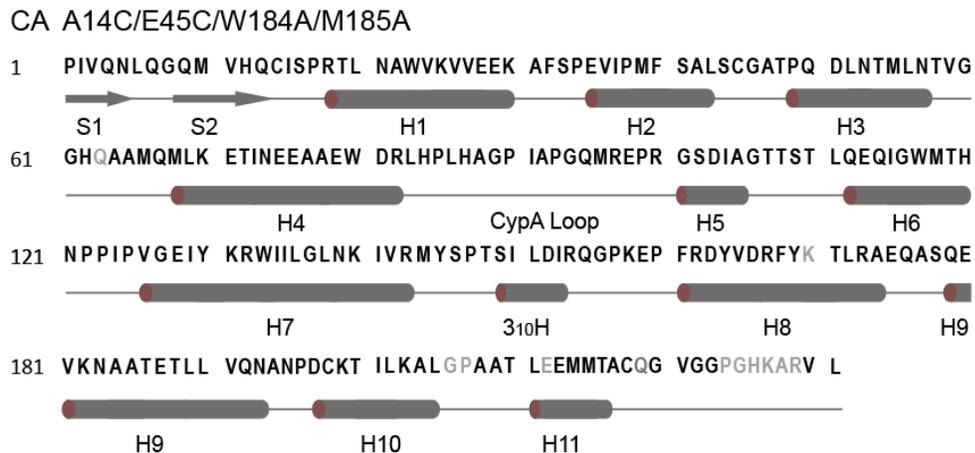


Figure 3.8 The amino acid sequence of CA A14C/E45C/W184A/M185A mutant. The residues that contain ambiguous backbone resonance assignments are labeled in gray.

2D homo- and hetero-nuclear spectra were collected for the cross-linked hexamer protein assembled in tubes. The spectra exhibit remarkably high resolution and sensitivity. A total of 89% backbone (N, C $\alpha$ , Co) resonance assignments were completed in CA cross-linked hexamer assemblies by combining 2D  $^{13}\text{C}$ - $^{13}\text{C}$  CORD, NCA, NCO, NCACX, and NCOCX spectra, see Fig 3.9. The residues with ambiguous assignments are located in several dynamic regions of loop 3/4, loop 10/11 and CTD tail (Fig 3.8). Additional 3D data sets are required to confirm their assignments. The unambiguous resonance assignments are listed in Table 3.1, and the predicted secondary structure for the CA cross-linked hexamer protein is identical to the CA WT protein.

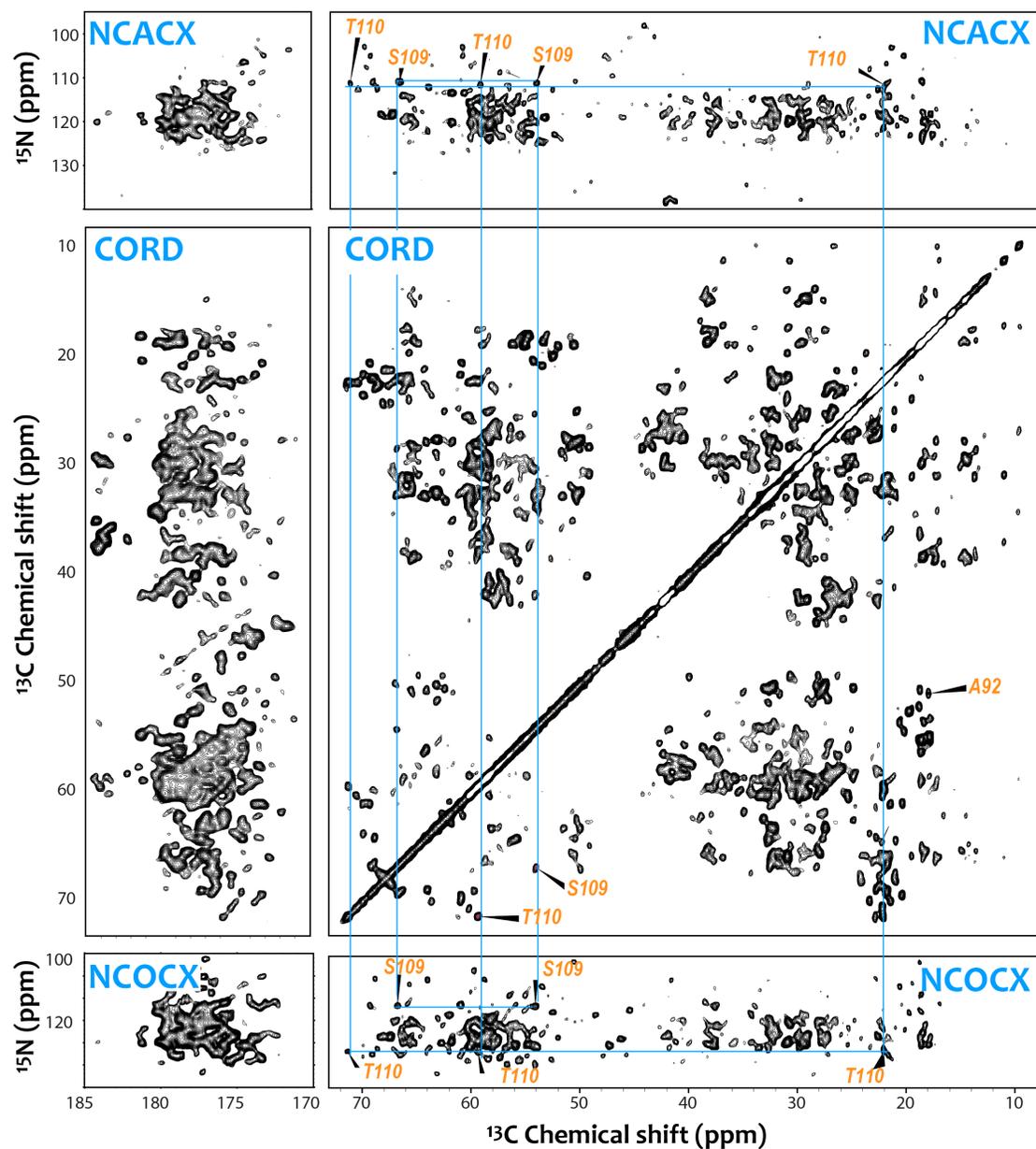


Figure 3.9 2D CORD, NCACX and NCOCX MAS NMR spectra for CA cross-linked hexamer tubular assemblies, showing assignments for S109-T110. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz; the data were processed with 60° shifted sine bell apodization process in both dimensions.

Table 3.1 Resonance assignments for CA cross-linked hexamer assemblies, including backbone N, Co, C $\alpha$ , sidechain C $\beta$ , C $\gamma$ /C $\gamma$ 2, C $\delta$ /C $\delta$ 2, C $\epsilon$ , and C $\zeta$ . CSI 2.0 <sup>13</sup> secondary structure prediction: C-coils, H-helix, and B-sheets.

Residue		N	Co	C $\alpha$	C $\beta$	C $\gamma$	C $\gamma$ 2	C $\delta$	C $\delta$ 2	C $\epsilon$	C $\zeta$	CSI 2.0
1	P	139.0	174.5	62.4	33.1	24.9		50.3				C
2	I	121.4		62.7	38.3	28.9		15.1				B
3	V	124.3		58.7	35.8	21.5	18.6					B
4	Q	120.8	179.5	54.2	29.9							C
5	N	125.7	175.2	50.8	39.1	177.1						C
6	L	116.6		56.4	42.4	26.1		22.4				C
7	Q	119.1	178.2	60.0	29.0	33.5						C
8	G	110.5	175.9	47.2								C
9	Q	121.1		54.6	29.5	34.3		180.0				C
10	M	114.1		55.1	35.1	31.3						B
11	V		180.3	60.5	36.0	21.5	19.3					B
12	H	121.4	174.6	58.2	32.7			115.7				B
13	Q			54.0	32.5	33.4						B
14	C	121.6	174.1	58.0	33.5							B
15	I	121.9	175.7	62.0	38.8	29.2	16.6	14.6				C
16	S	122.8	174.5	56.6	64.2							C
17	P	137.0	178.2	65.6	32.4	27.9		50.4				H
18	R	115.6	179.5	59.6	30.0	27.8		43.5			159.3	H
19	T	119.1	176.2	67.2	68.4	22.1						H
20	L	122.0	178.1	58.2	40.8	26.4		24.7				H
21	N	115.0	175.8	55.8	38.6	177.2						H
22	A	121.1	180.0	54.9	18.4							C
23	W	120.5	175.5	58.2	30.0	111.4		127.6				H
24	V	118.1	176.4	65.7	31.1	24.1	22.0					H
25	K	118.8	178.5	60.0	31.9	25.4		29.8		41.6		H
26	V	120.3	177.5	65.8	30.8	23.7	21.7					H
27	V	120.0	178.1	65.5	30.7	22.4						H
28	E	119.4	176.6	58.6	29.8	37.4		183.4				H
29	E	116.1		58.4	30.9	36.7		183.6				C
30	K	120.8	176.2	55.6	32.8	25.3		27.1		41.6		C
31	A	123.5	177.7	52.5	17.6							C
32	F	113.6	175.8	58.1	38.1							C

33	S	114.4	176.8	58.2	63.1							H
34	P		178.3	66.6	31.6	28.1		50.8				H
35	E	115.6	178.5	58.3	29.2	37.3						H
36	V	114.9	179.0	66.3	32.1	22.4	21.7					H
37	I	115.7	176.2	67.2	34.3	30.1	19.0	13.3				H
38	P	134.6		64.8	30.6	28.1		49.1				H
39	M		178.9	58.4	32.7					15.8		H
40	F	121.1	179.6	61.9	37.9							H
41	S	114.2	175.7	61.8	62.7							H
42	A	120.7	181.2	54.3	20.1							H
43	L	117.6	175.8	56.4	41.8	27.3		22.7				C
44	S	109.8	174.5	57.5	63.6							C
45	C	124.3		60.2	31.7							C
46	G	116.4	174.8	45.7								C
47	A	119.9	178.5	53.4	20.6							C
48	T	107.6	174.5	59.7	69.0	21.9						C
49	P	132.4		67.0	30.6	28.3		49.9				H
50	Q	115.5	178.6	60.4	29.4	34.3						H
51	D	116.8	178.0	57.6	41.3							H
52	L	120.5	176.7	57.5	41.7			27.1				H
53	N	118.6	178.6	56.5	38.2							H
54	T	118.9	176.2	67.2	68.5	24.2						H
55	M	119.6	179.6	60.2	33.1	32.8				17.9		H
56	L	121.6	180.3	58.6	41.6	26.9						H
57	N	115.7	176.5	54.8	37.4	177.3						H
58	T	108.2	174.9	63.6	69.9	23.4						C
59	V	121.4	177.2	62.6	32.4	21.9	21.5					C
60	G	117.0	177.2	45.3								C
61	G	106.9	174.5	47.3								C
62	H	115.4	177.2	57.1	28.9	131.4				138		C
63	Q											H
64	A	121.5	178.7	54.7	18.2							H
65	A	121.2	180.2	55.2	18.7							H
66	M	113.3	178.5	56.6	31.9	29.6				17.4		H
67	Q	121.4	178.8	58.4	27.4	32.6		181.4				H
68	M	119.4	178.2	58.4	32.6					17.9		H
69	L	122.1	179.5	58.8	41.8	27.5						H
70	K	119.1	178.6	60.0	31.9	25.4		29.0		42.1		H

71	E	118.6	178.3	59.9	29.8	36.9		183.4				H
72	T	118.5	176.8	67.2	68.5	22.1						H
73	I	122.5	177.0	65.9	38.1	29.4	18.5	14.5				H
74	N	116.4	178.6	56.0	37.7	175.9						H
75	E	121.3	179.6	59.3	29.4	36.2		183.9				H
76	E	120.7	177.5	58.9	29.3	37.5		184.2				H
77	A	123.5	179.4	54.8	17.7							H
78	A	119.1	181.3	54.6	17.8							H
79	E	121.1	178.1	58.5	28.9	34.9		183.6				H
80	W	120.9	179.7	61.8	28.0	110.4		126.4			113.3	H
81	D	118.9	178.8	57.4	39.7							H
82	R	120.4	178.3	59.1	30.7	28.2		43.3			159.3	H
83	L	116.5	176.2	55.3	42.7	25.3		22.5				C
84	H	116.9	171.1	53.0	27.5	130.2		118.7				C
85	P			63.1	32.2	27.4		50.1				C
86	V	121.1	176.8	62.0	32.5	21.1						C
87	H	125.0	174.5	55.6	29.3	130.9		118.3		137.4		C
88	A	127.6	177.7	52.3	18.8							C
89	G	109.4	171.6	44.7								C
90	P	132.6	176.6	63.3	32.1	27.4		49.9				C
91	I	125.1	174.7	59.8	39.0	28.9	17.9	15.6				C
92	A	130.0	176.1	50.8	18.0							C
93	P	132.4		64.1	31.6	27.7		50.4				C
94	G	111.1	174.2	45.4								C
95	Q	119.1	175.3	53.9	30.1	33.6		179.5				C
96	M	122.0	176.4	55.5	34.6	31.2						C
97	R	124.6	175.2	55.4	30.2	27.2		43.8			159.3	C
98	E	123.3		54.7	29.3	36.7		183.6				C
99	P	134.2	177.1	62.6	30.3	26.4		49.3				C
100	R	120.5	178.6	53.1	30.9	26.6		43.6			158.9	C
101	G	116.7	175.9	48.5								C
102	S	114.2	176.9	60.7	62.0							C
103	D	122.5	176.1	56.8	41.3	179.9						C
104	I	125.5	172.6	65.8	36.8	30.8	17.0	11.0				C
105	A	111.7	177.1	50.5	18.7							C
106	G	102.0	173.5	45.6								C
107	T	113.1	177.3	64.0	68.8	22.8						C
108	T	106.1	173.2	60.9	69.4	20.4						C

109	S	112.0	174.1	54.1	66.8							C
110	T	112.1	175.8	59.4	71.3	22.1						C
111	L	122.6	178.6	57.9	42.0	26.5		25.3	22.4			H
112	Q	115.9	178.4	59.6	27.8	35.2						H
113	E	120.4		59.3	29.6	37.4		184.1				H
114	Q	116.4	177.7	60.2	28.3	32.6		179.6				H
115	I	115.8	179.7	65.7	38.3	29.1	17.3	14.2				H
116	G	110.5	176.7	47.3								H
117	W	121.5	178.4	62.0	29.0	110.6		125.8 127.5		118.3 138.5		H
118	M	116.4	177.4	59.9	35.3	33.1				14.8		H
119	T	103.9	173.7	60.9	69.9	21.4						H
120	H	124.2	172.5	57.2	28.7	131.1						C
121	N	119.4	172.3	49.3	40.0	177.7						C
122	P	138.5		63.4	32.8			50.3				C
123	P	132.7		63.1	32.6	27.5		50.4				C
124	I	123.6	176.0	59.4	38.3	27.6	17.3	14.7				C
125	P	136.4	176.7	63.2	29.1	27.2		51.3				C
126	V	115.0	176.7	65.7	30.9	23.6	19.5					H
127	G	107.1	174.6	47.5								H
128	E	121.8	179.8	58.4	29.4	35.9						H
129	I	120.6	176.1	65.3	38.3	28.3	18.3	14.9				H
130	Y	116.4	176.0	56.0	36.3	128.3						H
131	K	116.5	177.8	60.4	31.8	25.3		29.8		41.4		H
132	R	117.1	178.3	60.2	29.9	28.3		44.3			159.5	H
133	W	122.0	178.7	58.4	29.5	110.8		124.8		138.8 119.5	114.2	H
134	I	118.5	178.3	65.2	38.3	29.8	18.6	15.0				H
135	I	120.8	177.7	66.3	37.6	30.9	17.3	13.7				H
136	L	120.9	181.2	58.8	41.5	26.9						H
137	G	106.0	175.9	47.7								H
138	L	123.7	178.7	57.7	42.2	26.0		25.4	22.8			H
139	N	118.6	178.2	56.9	38.3	176.1						H
140	K	117.5	178.7	60.1	32.3	25.4		29.9		42.2		H
141	I	119.6	178.3	65.2	37.7	29.7	18.1	14.7				H
142	V	119.6	177.2	65.7	32.3	21.9						H
143	R	119.8	178.2	59.7	29.9	26.7		41.7				H
144	M		178.0	57.6	32.8	29.9						H
145	Y	117.4		59.7	38.1							H

146	S	118.5	176.7	56.5	64.6							C
147	P			63.3	32.7	28.6		51.6				C
148	T	116.3	173.4	62.0	70.2	22.4						B
149	S		174.2	56.2	64.3							B
150	I		175.4	60.3	39.2	29.1	17.6	15.8				B
151	L	118.0	178.4	56.7	41.8	27.7		25.8	23.5			B
152	D	114.9	175.6	54.1	41.8							B
153	I	121.8	174.0	58.6	34.0	26.5	17.1	9.6				B
154	R	125.2	175.3	54.0	33.2	27.0		43.2			159.4	B
155	Q	127.0	176.9	55.7	29.8	32.5		179.8				B
156	G	117.1	174.1	44.9								C
157	P		177.2	65.6	32.4	27.7		50.4				C
158	K	115.5	175.6	54.1	32.6	25.0		30.0		41.9		C
159	E	125.6	174.5	53.4	32.2	35.0		184.1				C
160	P		177.4	63.1	32.6	28.6		51.5				C
161	F	128.1	177.1	62.6	39.1							C
162	R	116.6	179.4	60.0	30.5	27.2		43.6				H
163	D	121.5		57.2	40.0	179.2						H
164	Y	123.9	176.2	59.3	38.2	129.4		132.5		118.3		H
165	V	116.0		66.4	31.0	25.2						H
166	D	120.4		57.7	41.2	180.1						-
167	R	120.5										-
168	F	124.8	177.1	62.5	39.2	138.1						H
169	Y	114.4	179.6	63.4	38.1							H
170	K											H
171	T	114.8	176.4	66.3	68.9	21.6						H
172	L	123.4	178.8	57.8	42.5	26.0						H
173	R	116.1	179.6	59.1	30.2	27.1						H
174	A	119.4	178.5	53.5	18.4							H
175	E	118.4	177.8	57.1	29.4	36.4		182.3				C
176	Q	124.4										C
177	A	124.9	177.3	52.1	20.6							C
178	S	119.2	178.0	57.5	64.8							C
179	Q	116.6	177.8	58.7	28.7	33.2						H
180	E	118.1	179.7	60.0	29.8							H
181	V	120.7	178.0	66.5	30.9	24.8	22.4					H
182	K			61.2	32.6			29.7				H
183	N	117.7	175.8	56.9	37.5							H

184	A	122.0	177.2	53.6	19.4							H
185	A	118.9	176.9	54.7	17.9							H
186	T	111.8	176.5	66.7	69.2	21.8						H
187	E	123.2		60.0	29.6							H
188	T	113.4	175.2	64.7	69.6	22.2						H
189	L		178.7	57.8	42.2							H
190	L	118.1	178.0	58.4	41.0							H
191	V	115.4	177.3	65.7	32.1	21.6						H
192	Q	115.4	177.2	58.3								H
193	N	113.8	175.2	52.6	38.3							C
194	A	124.2	175.5	52.2	18.7							C
195	N	119.4	174.3	51.6	35.5							C
196	P	131.7	178.6	66.8	32.6	28.2		50.2				H
197	D	115.7	178.3	57.1	40.4	179.7						H
198	C	121.1	176.0	62.2	27.7							H
199	K	121.8	177.7	61.1	32.4	24.7		29.8		42.4		H
200	T	111.8	176.4	66.6	69.0	21.8						H
201	I	123.4	179.0	64.6	38.3	29.4	18.6	14.3				H
202	L	119.3	178.5	57.5	41.9	27.7		25.6	23.3			H
203	K	120.0		59.4	30.8	25.8		29.2				H
204	A	119.9	178.9	53.3	18.5							H
205	L	118.3		57.3	41.1	26.4						C
206	G											-
207	P			64.3				50.4				-
208	G	110.3	173.6	45.6								C
209	A	123.8	177.8	52.5	19.8							C
210	T	113.6	175.3	60.2	70.6	22.2						C
211	L	124.5	178.6	58.5	40.5	27.2		25.0				H
212	E											H
213	E		177.8	59.3	29.5	37.6						H
214	M	120.4	176.7	60.2	33.5							H
215	M	117.2	178.4	58.8	32.4							H
216	T	115.2	176.5	66.4	68.9	21.8						H
217	A	123.8	178.8	56.0	18.6							H
218	C	112.2		59.2	29.1							C
219	Q											-
220	G	115.4		47.4								-
221	V	121.5	176.7	64.2	32.7	22.2	21.0					-

222	G	117.1		44.9								C
223	G	112.4	174.2	45.6								-
224	P											-
225	G											-
226	H											-
227	K											-
228	A											-
229	R											-
230	V	119.8		65.9	30.9	21.8						C
231	L	116.4	178.7	56.3	42.4	26.4						C

### 3.3.2 The Effect of Cross-Linking on Structure of CA Tubular Assemblies

Disulfide bonds introduce increased stability of the CA hexameric lattice by cross-linking two cysteines. To understand the effect of cross-linking on the structure of CA assemblies, 2D NCA and CORD spectra for CA cross-linked hexamer assemblies were compared with those for CA WT assemblies. Multiple chemical shift changes were observed in the superposition of homo- and hetero-nuclear experiments, Fig 3.10. The chemical shift perturbations were calculated for each resolved residue by adding differences of backbone N and C $\alpha$  resonances. The standard deviation (STD) of chemical shift difference was 0.6 ppm, indicating the sizeable effects on the local structure throughout the CA domain, arising from the mutations as well as cross-linking. 1.2 ppm (2\*STD) was used as the cutoff to identify the residues with significant perturbations, Fig 3.11. The most pronounced chemical shift differences are seen in H2, H3 (NTD-NTD interface) and loop 8/9, in the immediate vicinity of mutations. Furthermore, other chemical shift differences suggest that the mutations induce allosteric conformational changes in the beta-hairpin, loop 1/2, CypA loop and H8 (CTD-CTD interface). For instance, the large perturbations observed in residues E175 and S178 (loop 8/9) have been identified as critical factors for CA dimerization<sup>14</sup>, which is likely due to the W184A and W185A mutations. These observed chemical shift differences are also consistent with the structural changes observed in the X-ray crystal structure of CA cross-linked hexamer protein by Pornillos group<sup>4</sup>: conformational differences were seen in this cross-linked hexamer construct at NTD-NTD intra-hexamer and CTD-CTD inter-hexamer interfaces.



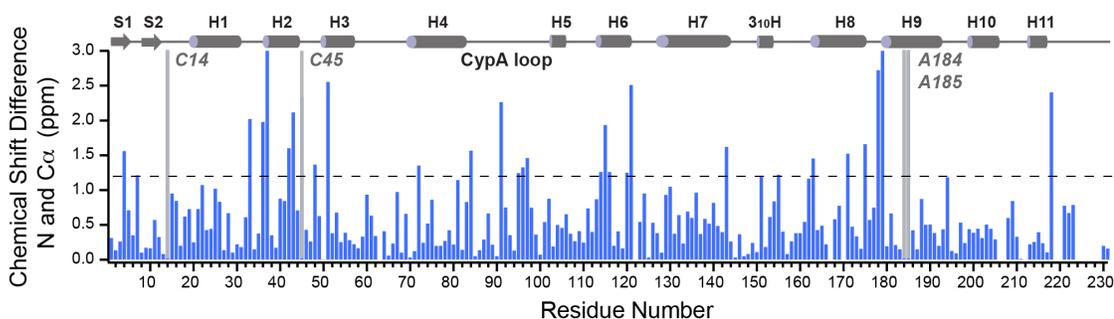


Figure 3.11 Chemical shift changes between CA and CA cross-linked hexamer: sum of  $^{13}\text{C}\alpha$ ,  $^{15}\text{N}$  chemical shift perturbations (CSP). Note that a CSP > 1.2 ppm is considered a significant perturbation, while a CSP < 0.6 ppm (1\*STD) is not considered major given the mutations as well as experimental and systematic error.

### 3.3.3 The Effect of Morphology on MAS NMR Spectra: Pentamers vs. Hexamers

*In vivo*, hexamers and pentamers are thought to be the building blocks for HIV-1 mature capsid lattice. For characterization of the capsid assembly and disassembly processes in the viral lifecycle, it is essential to analyze the structural and dynamic signatures of these pentameric and hexameric building blocks in the conical capsid core. While this is a difficult task in the context of the native capsid, in our work we turned attention to the model systems. To do so, covalently cross-linked units of CA hexamers (A14C/E45C/W184A/M185A) and pentamers (N21C/A22C/W184A/M185A) were employed in assemblies studied by MAS NMR. Under high salt conditions, CA cross-linked hexamer (A14C/E45C/W184A/M185A) proteins efficiently assembled into tubes. These tubes yielded high-quality MAS NMR spectra (Fig 3.12). We also co-assembled cross-linked hexamers and cross-linked pentamers in a 10:1 ratio, albeit the morphology of the resulting sample appears to be mixed. To examine the pentamer structure in the context of these model co-assemblies, I

prepared mixed labeled samples with uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled cross-linked pentamers and natural abundance cross-linked hexamers. This sample was subjected to MAS NMR experiments. As illustrated in Fig 3.12 and 3.13,  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra of CA pentamers showed high resolution albeit the sensitivity was very low because of small amounts of isotopically labeled pentamers in the sample (less than 10%). The narrow lines suggested that CA pentamers have significant local order.

The superposition of the  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra of the cross-linked pentamer/hexamer co-assemblies and the cross-linked hexamer assemblies (Fig 3.13) illustrates that the chemical shifts are very similar, suggesting no major changes in secondary or tertiary structures, although several residue-specific chemical shift and peak intensity differences are observed. It is not surprising that the pronounced perturbations are associated with residues near the mutation sites, such as A42, A47, and T48 near the E45C change. Moreover, sizable differences are also present in the residues distal from the mutation sites, which indicate conformational changes in both NTD and CTD for these two assemblies. For example, the perturbed resonance of A217 possibly indicates conformational changes in the NTD-CTD interface associated with H7 and H11 in neighboring subunits<sup>4</sup>. In addition, the large chemical shift difference of T110 suggests a dramatically different assembly conformation associated with H5 and loop 5/6 in the hexamer and pentamer assemblies.

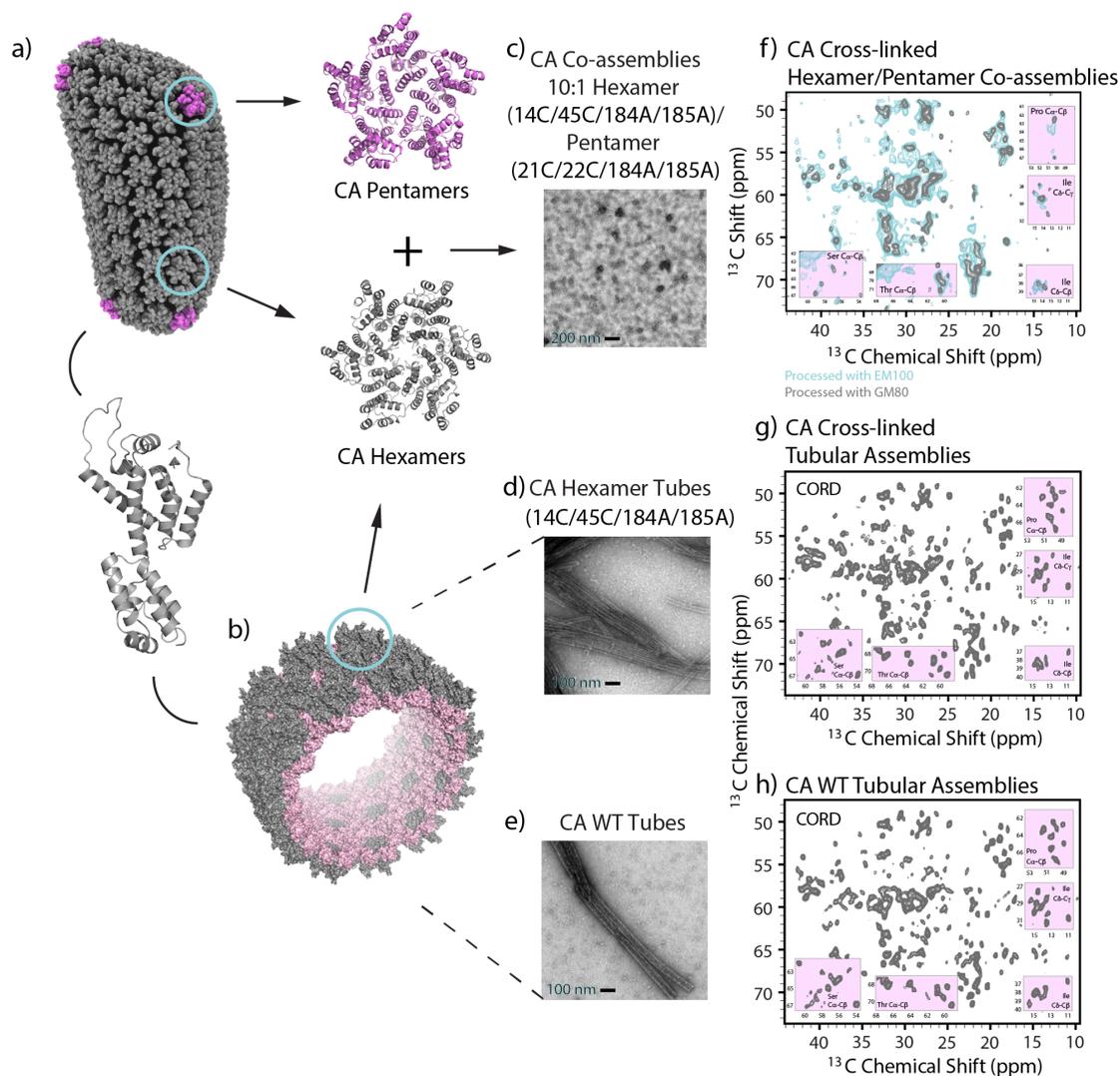


Figure 3.12 Structure of a HIV-1 conical capsid (a) and a tubular capsid (b), and their building blocks. TEM images and  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra (aliphatic region) of CA hexamer and pentamer co-assemblies (c, f), CA cross-linked tubular assemblies (d, g), and CA WT tubular assemblies (e, h). All the spectra were acquired at 20.0 T and the MAS frequency of 14 kHz; (f) were processed with Lorentz-to-Gauss (GM) 80 (gray) and Exponential-Multiply (EM) 100 (blue) in both dimensions, and (g, h) were process with  $45^\circ$  shifted sine bell apodization in both dimensions.

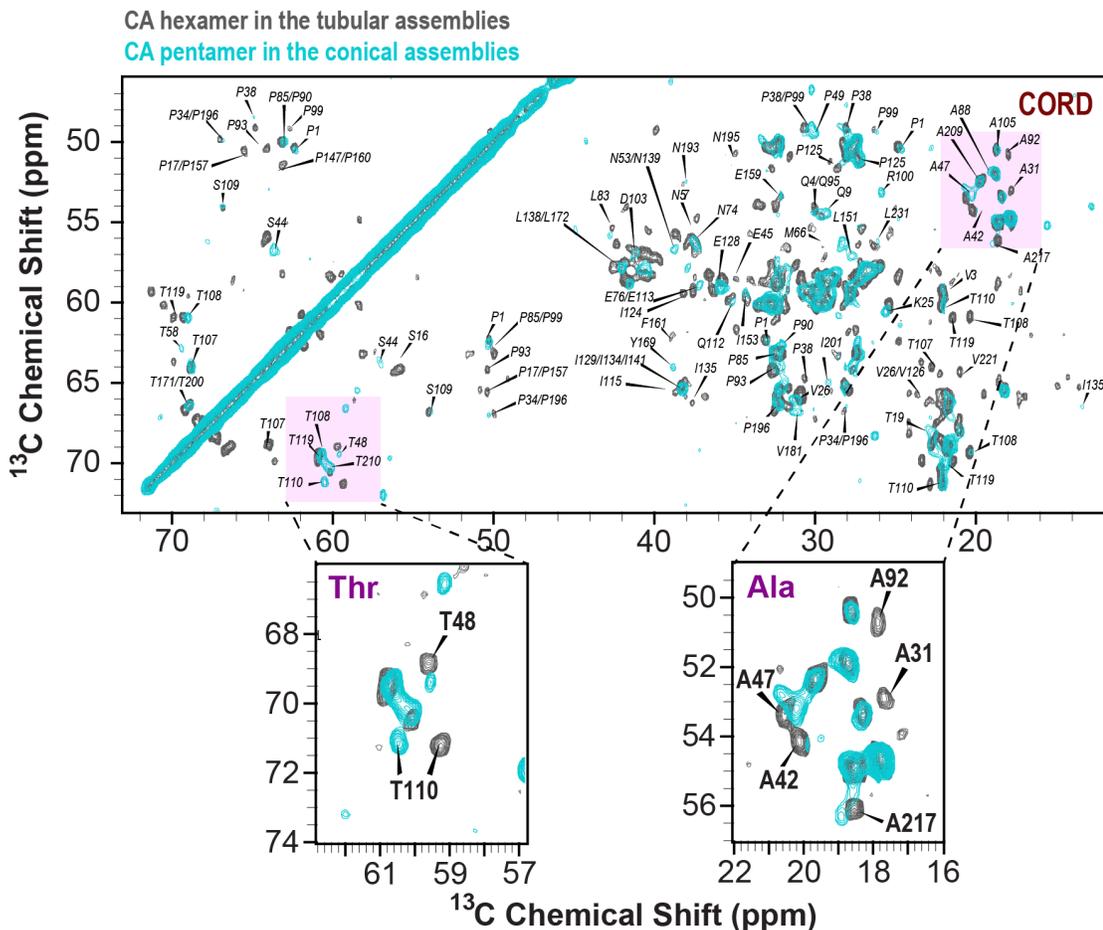


Figure 3.13 Superposition of 2D CORD (aliphatic region) MAS NMR spectra for CA hexamer (gray) in tubular assemblies and CA cross-linked pentamer (cyan) in co-assemblies with CA cross-linked hexamer. The expansions around  $\text{C}\alpha\text{-C}\beta$  regions of Thr and Ala are shown to illustrate chemical shift perturbations. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz, processed with Lorentz-to-Gauss (GM) 80 in both dimensions. The spectra of CA cross-linked pentamer in co-assemblies with the CA cross-linked hexamer were acquired with non-uniformly sampling (NUS).

### 3.4 Conclusions and Future Outlook

In this chapter, we discussed the sample preparation for the CA cross-linked pentamer and cross-linked hexamer proteins. CA cross-linked hexamer formed

homogenous tubular assemblies. Co-assemblies of CA cross-linked hexamers and pentamers were prepared as models to study the structure and dynamics of the HIV-1 capsid.

I have completed 89% of the backbone assignments of CA cross-linked hexamer tubular assemblies on the basis of multiple MAS NMR data sets. The chemical shifts revealed conformational differences in a number of residues in these assemblies compared with CA WT tubular assemblies. In tubular assemblies, we observed that the effect of cross-linking was pronounced on the inter-hexamer interfaces, but not on the protein secondary structure and tertiary structure. I have also prepared MAS NMR samples of co-assemblies of U-<sup>13</sup>C, <sup>15</sup>N cross-linked pentamers with natural abundance cross-linked hexamers and acquired 2D <sup>13</sup>C-<sup>13</sup>C correlations spectra. We observed multiple chemical shift perturbations suggesting conformational differences in the hexamers vs. pentamers.

Due to the low amount of isotopically labeled materials in co-assembly (less than 10%), the sensitivity of the current data sets were limited. Fast and/or ultrafast MAS approaches and <sup>1</sup>H detection may improve spectra quality and will be pursued in the future. To understand the HIV-1 capsid assembly and stability, further investigations will be performed on dynamic properties of cross-linked pentamers as well as cross-linked hexamers. Additionally, taking advantage of the stability and reversible assembly into tubes, CA cross-linked hexamers can be used to study different aspects of the HIV-1 function, for instance, the binding of non-polar capsid assembly inhibitor molecules (e.g., benzodiazepines and benzoimidazoles)<sup>15</sup>, and interaction of restriction factors (e.g., tripartite motif isoform 5 $\alpha$ , TRIM5 $\alpha$ )<sup>16-18</sup> with capsid assemblies.

## REFERENCES

- (1) Zhao, G. P.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J. Y.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **2013**, *497* (7451), 643.
- (2) Ganser-Pornillos, B. K.; Cheng, A.; Yeager, M. Structure of full-length HIV-1CA: a model for the mature capsid lattice. *Cell* **2007**, *131* (1), 70.
- (3) Han, Y.; Ahn, J.; Concel, J.; Byeon, I. J. L.; Gronenborn, A. M.; Yang, J.; Polenova, T. Solid-State NMR Studies of HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2010**, *132* (6), 1976.
- (4) Pornillos, O.; Ganser-Pornillos, B. K.; Kelly, B. N.; Hua, Y.; Whitby, F. G.; Stout, C. D.; Sundquist, W. I.; Hill, C. P.; Yeager, M. X-ray structures of the hexameric building block of the HIV capsid. *Cell* **2009**, *137* (7), 1282.
- (5) Pornillos, O.; Ganser-Pornillos, B. K.; Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **2011**, *469* (7330), 424.
- (6) Han, Y.; Hou, G. J.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Lipton, A. S.; Burton, S.; Hung, I.; Gor'kov, P. L.; Gan, Z. H. et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2013**, *135* (47), 17793.
- (7) Sun, S. J.; Han, Y.; Paramasivam, S.; Yan, S.; Siglin, A. E.; Williams, J. C.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. Solid-State NMR Spectroscopy of Protein Complexes. *Methods Mol Biol* **2012**, *831*, 303.
- (8) Sevier, C. S.; Kaiser, C. A. Formation and transfer of disulphide bonds in living cells. *Nature Reviews Molecular Cell Biology* **2002**, *3*, 836.

- (9) Hoch, J. C.; Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S. Nonuniform sampling and maximum entropy reconstruction in multidimensional NMR. *Acc Chem Res* **2014**, *47* (2), 708.
- (10) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995**, *6* (3), 277.
- (11) Goddard, T. D.; Kneller, D. G. University of California, San Francisco, 2004.
- (12) Stevens, T. J.; Fogh, R. H.; Boucher, W.; Higman, V. A.; Eisenmenger, F.; Bardiaux, B.; van Rossum, B. J.; Oschkinat, H.; Laue, E. D. A software framework for analysing solid-state MAS NMR data. *Journal of biomolecular NMR* **2011**, *51* (4), 437.
- (13) Hafsa, N. E.; Wishart, D. S. CSI 2.0: a significantly improved version of the Chemical Shift Index. *J Biomol Nmr* **2014**, *60* (2-3), 131.
- (14) Gres, A. T.; Kirby, K. A.; KewalRamani, V. N.; Tanner, J. J.; Pornillos, O.; Sarafianos, S. G. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **2015**, *349* (6243), 99.
- (15) Lemke, C. T.; Titolo, S.; von Schwedler, U.; Goudreau, N.; Mercier, J. F.; Wardrop, E.; Faucher, A. M.; Coulombe, R.; Banik, S. S. R.; Fader, L. et al. Distinct Effects of Two HIV-1 Capsid Assembly Inhibitor Families That Bind the Same Site within the N-Terminal Domain of the Viral CA Protein. *Journal of Virology* **2012**, *86* (12), 6643.
- (16) Sanz-Ramos, M.; Stoye, J. P. Capsid-binding retrovirus restriction factors: discovery, restriction specificity and implications for the development of novel therapeutics. *J Gen Virol* **2013**, *94*, 2587.
- (17) Gamble, T. R.; Vajdos, F. F.; Yoo, S.; Worthylake, D. K.; Houseweart, M.; Sundquist, W. I.; Hill, C. P. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **1996**, *87* (7), 1285.
- (18) Zhao, G. P.; Ke, D. X.; Vu, T.; Ahn, J.; Shah, V. B.; Yang, R. F.; Aiken, C.; Charlton, L. M.; Gronenborn, A. M.; Zhang, P. J. Rhesus TRIM5 alpha Disrupts the HIV-1 Capsid at the Inter-Hexamer Interfaces. *Plos Pathog* **2011**, *7* (3).



## Chapter 4

### INVESTIGATION OF HIV-1 MATURATION

#### 4.1 Introduction

In HIV-1 replication cycle, capsid maturation is a critical process where viral particles assemble and become infectious.<sup>1,2</sup> Maturation inhibitors, such as BVM can prevent the capsid maturation.<sup>3,4</sup> Despite extensive studies, both the maturation mechanism and its inhibition by small molecules remain poorly understood. In this chapter, we applied an integrated MAS NMR and molecular dynamics (MD) simulations approach to investigate the atomic-level details of the maturation and the influence of maturation inhibitors. By examining the final-step maturation intermediate, CA-SP1, we identified a helix-coil equilibrium in the CTD tail-SP1 subdomain potentially acting as a conformational switch during the maturation. CA-SP1(T8I) mutant was also investigated because T8I mutation phenocopies the MI-bound state. The T8I mutation in SP1 was found to introduce dramatic dynamics and structural changes, with molecular motions being suppressed, and CTD-SP1 junction residues showing higher helical structure propensity. Our results suggest that modulation of protein dynamics may be a determining factor in capsid maturation.

This work was conducted in our lab by Dr. Caitlin Quinn, Dr. Huilan Zhang, Dr. Guangjin Hou and Dr. Christopher Suiter, and by our collaborators, Professor Juan Perilla from the University of Delaware, Dr. In-Ja Byeon, Professor Peijun Zhang and Professor Angela Gronenborn from the University of Pittsburgh, the group of Dr. Eric Freed from the National Cancer Institute, Professor Klaus Schulten from the

University of Illinois. My contributions were the preparation of CA-SP1(T8I) samples for MAS NMR studies, performing MAS NMR experiments and the data analysis. The study has been published: Wang, M. Z.; Quinn, C. M.; Perilla, J. R.; Zhang, H. L.; Shirra, R.; Hou, G. J.; Byeon, I. J.; Suiter, C. L.; Ablan, S.; Urano, E. et al. Quenching protein dynamics interferes with HIV capsid maturation. *Nat Commun* **2017**, 8.

## **4.2 Methods and Experiments**

### **4.2.1 Materials**

Common chemicals were purchased from Fisher Scientific or Sigma Aldrich. Chromatography columns were purchased from GE Healthcare.  $^{15}\text{NH}_4\text{Cl}$  and  $\text{U-}^{13}\text{C}_6$ -glucose were purchased from Cambridge Laboratories, Inc. The Rosetta (DE3) competent cells used for protein expression were purchased from Novagen. The HIV-1 CA-SP1(T8I), NL4-3 protein construct was prepared in the laboratory of our collaborator, Dr. Eric O. Freed at the National Cancer Institute.

### **4.2.2 Expression and Purification of CA-SP1 Proteins**

Uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -enriched CA-SP1(T8I) proteins were expressed as reported previously with modifications.<sup>5,6</sup> The CA-SP1(T8I) gene was amplified and ligated into pET21 vector using NdeI and XhoI sites, and *E. coli* Rosetta (DE3) cells were transformed with the vector. A single colony was inoculated in 5 mL of LB medium at 37 °C for 8-10 h until cells were grown to 1.5 OD. Following a centrifugation at RFC 4000 g at 4 °C for 5 min, cells were resuspended in 50 mL of fresh M9 medium as a seeding culture, grown overnight at 37 °C. M9 seeding culture was transferred into 1 L of M9 medium that contains 2 g/L  $^{15}\text{NH}_4\text{Cl}$ , 2 g/L  $\text{U-}^{13}\text{C}_6$ -glucose, trace elements solution,  $\text{CaCl}_2$ ,  $\text{MgSO}_4$ , biotin and thiamin. Cells were grown to 1.0-1.2 OD, induced

with 0.8 mM IPTG and expressed at 18 °C for 16 h. Cells were harvested by a centrifugation at RFC 4000 g at 4 °C for 1 h, resuspended in a 25 mM sodium phosphate buffer (pH 7.0) and then stored at -80 °C.

Uniformly <sup>13</sup>C,<sup>15</sup>N-enriched CA-SP1(T8I) proteins were purified by adding 200 mM benzamidin (working concentration is 1 mM) in thawed cells as a protease inhibitor to prevent CA-SP1 cleavage. After sonication at 35% of amplifier power with 20 s on and 40 s off for 5 min, the lysed cells were centrifuged at 27,000 g at 4 °C for 1 h. The pH of the supernatant was adjusted to 5.8 with acetic acid, and the conductivity was reduced to below 2.5 ms/cm. Following another centrifugation at 27,000 g at 4 °C for 1 h, the final supernatant was loaded onto a cation exchange column (HiTrap SP HP, 5 mL) and eluted with a 0-1 M NaCl gradient in a buffer containing 25 mM sodium phosphate (pH 5.8), 1 mM DTT, and 0.02% NaN<sub>3</sub>. Concentrated protein fractions were further purified using a size-exclusion column (HiLoad 26/600 Superdex 75 pg) equilibrated with a buffer containing 25 mM sodium phosphate (pH 7.0), 1 mM DTT, and 0.02% NaN<sub>3</sub>.

CA-SP1(T8I) tubes were pre-assembled from a 26 mg/mL protein solution in a 25 mM phosphate buffer (pH 5.5) containing 2.4 M NaCl, followed by incubation at 37 °C for 1 h and 4 °C overnight. The CA-SP1(T8I) assemblies were pelleted at 10,000 g and packed into a Bruker 3.2 mm rotor.

#### **4.2.3 Transmission Electron Microscopy (TEM)**

The morphologies of CA and CA-SP1 assemblies were characterized by transmission electron microscopy (TEM). TEM images were collected at the Delaware Biotechnology Institute (DBI) in the Bioimaging Center. The images were acquired with a Zeiss Libra 120 transmission electron microscope operating at 120 kV, and

equipped with a Gatan Ultrascan 1000 2k x 2k charge-coupled device camera. CA and CA-SP1 assemblies were stained with uranyl acetate (0.5-1% w/v), deposited onto 400 mesh, formvar/carbon-coated copper grids, and dried for 45 min in the air. The copper grids were glow discharged prior to staining, so that the tubular assemblies were uniformly spread on the grid surface and adhered to it.

CA proteins, including CA wild type (NL4-3 and HXB2 strains) and CA(A92E) NL4-3 assembled into tubes (Fig 4.1, bottom)<sup>7</sup> that recapitulated the mature capsid lattice. CA-SP1 proteins, including CA-SP1(T8I) NL4-3, CA(A92E)-SP1 NL4-3<sup>5</sup> and CA-SP1 HXB2<sup>5</sup> assembled into tubes (Fig 4.1, top) under identical conditions that used for CA assemblies<sup>8</sup>.

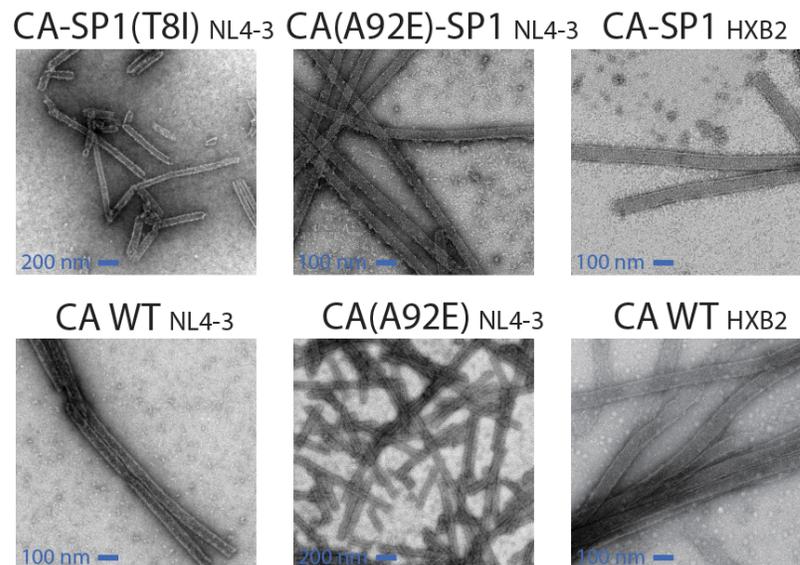


Figure 4.1 TEM images<sup>9</sup> of tubular assemblies of CA<sup>7</sup> and CA-SP1<sup>5</sup> variants.

#### 4.2.4 MAS NMR Experiments

MAS NMR experiments were performed on Bruker 20.0 T and 14.1 T narrow bore Avance III spectrometers outfitted with 3.2 mm E-Free HCN probes. MAS NMR spectra were collected at a magic angle spinning frequency of  $14.000 \pm 0.005$  kHz, and sample temperature of  $4 \pm 0.1$  °C throughout the experiments using a Bruker temperature controller.  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts were REFERENCED with respect to the external standards adamantane and ammonium chloride, respectively.

MAS NMR spectra of tubular assemblies of CA and CA-SP1 were acquired at 20.0 T, with Larmor frequencies of 850.4 MHz ( $^1\text{H}$ ), 213.8 MHz ( $^{13}\text{C}$ ) and 86.2 MHz ( $^{15}\text{N}$ ). The typical  $90^\circ$  pulse lengths were 2.4-3.1  $\mu\text{s}$  ( $^1\text{H}$ ), 3.2-3.9  $\mu\text{s}$  ( $^{13}\text{C}$ ), and 3.5-4.0  $\mu\text{s}$  ( $^{15}\text{N}$ ). The  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  CP employed a linear amplitude ramp of 90-110% on  $^1\text{H}$ , and the center of the ramp matched to Hartmann-Hahn conditions at the first spinning sideband, with contact time of 1.1-1.5 and 1.5-1.8 ms, respectively. For 2D  $^{13}\text{C}$ - $^{13}\text{C}$  CORD<sup>10</sup> spectra, the CORD mixing time was 50 ms, and  $^1\text{H}$  field strength during CORD was 12.5-13.5 kHz. For 2D NCA, 2D NCACX, and 2D NCOCX, band-selective magnetization transfer from  $^{15}\text{N}$  to  $^{13}\text{C}$  contact time was 4.3-5.5 ms. SPINAL-64 decoupling<sup>11</sup> (80-96 kHz) was used during the evolution and acquisition periods. For 3D RN-symmetry based DIPSHIFT and PARS experiments<sup>12,13</sup>,  $\text{R}12_1^4$  symmetry sequences with an RF field strength of 84 kHz were applied for reintroducing  $^1\text{H}$ - $^{15}\text{N}$  dipolar couplings. Some experiments of CA-SP1(T8I) assemblies were performed at 14.1 T, with Larmor frequencies of 600.8 MHz ( $^1\text{H}$ ), 150.8 MHz ( $^{13}\text{C}$ ) and 60.8 MHz ( $^{15}\text{N}$ ). The typical  $90^\circ$  pulse lengths were 2.8  $\mu\text{s}$  ( $^1\text{H}$ ), 3.6  $\mu\text{s}$  ( $^{13}\text{C}$ ), and 4.5  $\mu\text{s}$  ( $^{15}\text{N}$ ). The  $^1\text{H}$ - $^{13}\text{C}$  and  $^1\text{H}$ - $^{15}\text{N}$  CP employed a linear amplitude ramp of 90-110% on  $^1\text{H}$ , and the center of the ramp matched to Hartmann-Hahn conditions at the first spinning sideband, with contact time of 0.8 and 1.4 ms,

respectively. The CORD mixing time was 50 ms, and  $^1\text{H}$  field strength during CORD was 14 kHz. Band-selective magnetization transfer from  $^{15}\text{N}$  to  $^{13}\text{C}$  contact time was 5.0-6.5 ms. SPINAL-64 decoupling (89 kHz) was used during the evolution and acquisition periods.

All spectra were processed in TopSpin and with NMRpipe<sup>14</sup>; spectra analysis were performed using SPARKY<sup>15</sup> and CCPNMR<sup>16</sup>. For 2D and 3D data sets, 30°, 45°, 60° or 90° shifted sine bell apodization followed by a Lorentzian to Gaussian transformation was applied in both dimensions. Forward linear prediction to twice the number of the original data points was used in the indirect dimension in some data sets followed by zero filling to twice the total number of points.

#### **4.2.5 Simulations of NMR Dipolar Lineshapes and Calculations of Dipolar Order Parameters**

The SIMPSON software package (versions 1.1.2 and 2)<sup>17</sup> was used to complete the numerical simulations of  $^1\text{H}$ - $^{15}\text{N}$  dipolar lineshapes. For RN-DIPSHIFT experiments, 986 pairs of  $\{\alpha, \beta\}$  angles were generated with the ZCW algorithm, and 5  $\gamma$  angles that resulted in a total of 4930 angle triplets were used to calculate a powder average, as reported previously<sup>7</sup>. 320 pairs of  $\{\alpha, \beta\}$  angles were generated with the REPULSION algorithm, and 16  $\gamma$  angles that resulted in a total of 5120 angle triplets were applied to calculate a powder average for RN-PARS experiments. The experimental and processing parameters (i.e., MAS frequency, Larmor frequency, RF field strength, number of T1 points, zero-filling, line broadening, finite pulse lengths, etc.) were utilized for extracting the best-fit dipolar parameters with simulations. For all the 3D RN-DIPSHIFT/PARS spectra, a series of home-written C++ programs and shell scripts were used to process automatic extractions of dipolar lineshapes for the

residues in CypA loop (H84-S102), iteratively with manual inspections, to ensure the correctness of the assignments and extractions.

#### **4.2.6 Molecular Dynamics (MD) Simulation Trajectories of CA-SP1 Assemblies**

All-atom MD simulation trajectories of CA CTD-SP1 WT and CA CTD-SP1(T8I)<sup>9</sup> were carried out by our collaborators, Dr. Juan R. Perilla from the University of Delaware and Dr. Klaus Schulten from the University of Illinois.

#### **4.2.7 Calculations of Chemical Shifts from MD Trajectories**

For WT CA CTD-SP1, 7394 frames were extracted from 7-24  $\mu$ s MD trajectory. For CA CTD-SP1(T8I) mutant, 5860 frames were extracted from 0-15  $\mu$ s MD trajectory. SHIFTX2<sup>18</sup> was used to calculate  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{13}\text{C}\gamma$ , and  $^{15}\text{N}$  chemical shifts for each frame and each single chain. The STRIDE program was used to predict the secondary structures for every frame from atomic coordinates<sup>19</sup>. We categorized the secondary structure outputs into four types: helix,  $\beta$ -sheet, coil, and turn. The helix classification includes an  $\alpha$ -helix, a  $3_{10}$  helix, and a  $\pi$  helix. The  $\beta$ -sheet classification includes an extended configuration and an isolated  $\beta$ -sheet hydrogen bond formation.

### **4.3 Results**

#### **4.3.1 Resonance Assignments of CA-SP1(T8I) Assemblies**

CA-SP1(T8I) protein tubular assemblies yield high-quality MAS NMR spectra (Fig 4.2) similar to those in resolution and sensitivity of CA and CA-SP1 WT assemblies<sup>5,8</sup>. Site-specific resonance assignments of CA-SP1(T8I) protein assemblies were performed from a set of 2D and 3D homo- and heteronuclear correlation spectra,

including 2D CORD/direct-CORD, 2D/3D NCACX, 2D/3D NCOCX, 3D CONCA, and 2D J-INADEQUATE spectra.

From the superposition of dipolar-based  $^{13}\text{C}$ - $^{13}\text{C}$  spectra of CA and CA-SP1(T8I) tubular assemblies (Fig 4.2), strong and well-defined peaks were observed in both CA CTD tail and SP1 region. In contrast, the resonances of residues in the CTD tail and SP1 were absent in dipolar-based spectra in the previous work on CA-SP1 HXB2 and CA(A92E)-SP1 NL4-3.<sup>5</sup> This indicates that T8I mutation induces the decreased dynamics in the CTD tail and SP1 region, compared to WT SP1 proteins (section 4.3.2). A full stretch of sequential walk is shown for the CTD tail (CA: G225-L231) and SP1 region (SP1: A1-M14) by several types of MAS NMR spectra in Fig 4.3, and the site-specific chemical shifts of each residue is listed in Table 4.1.

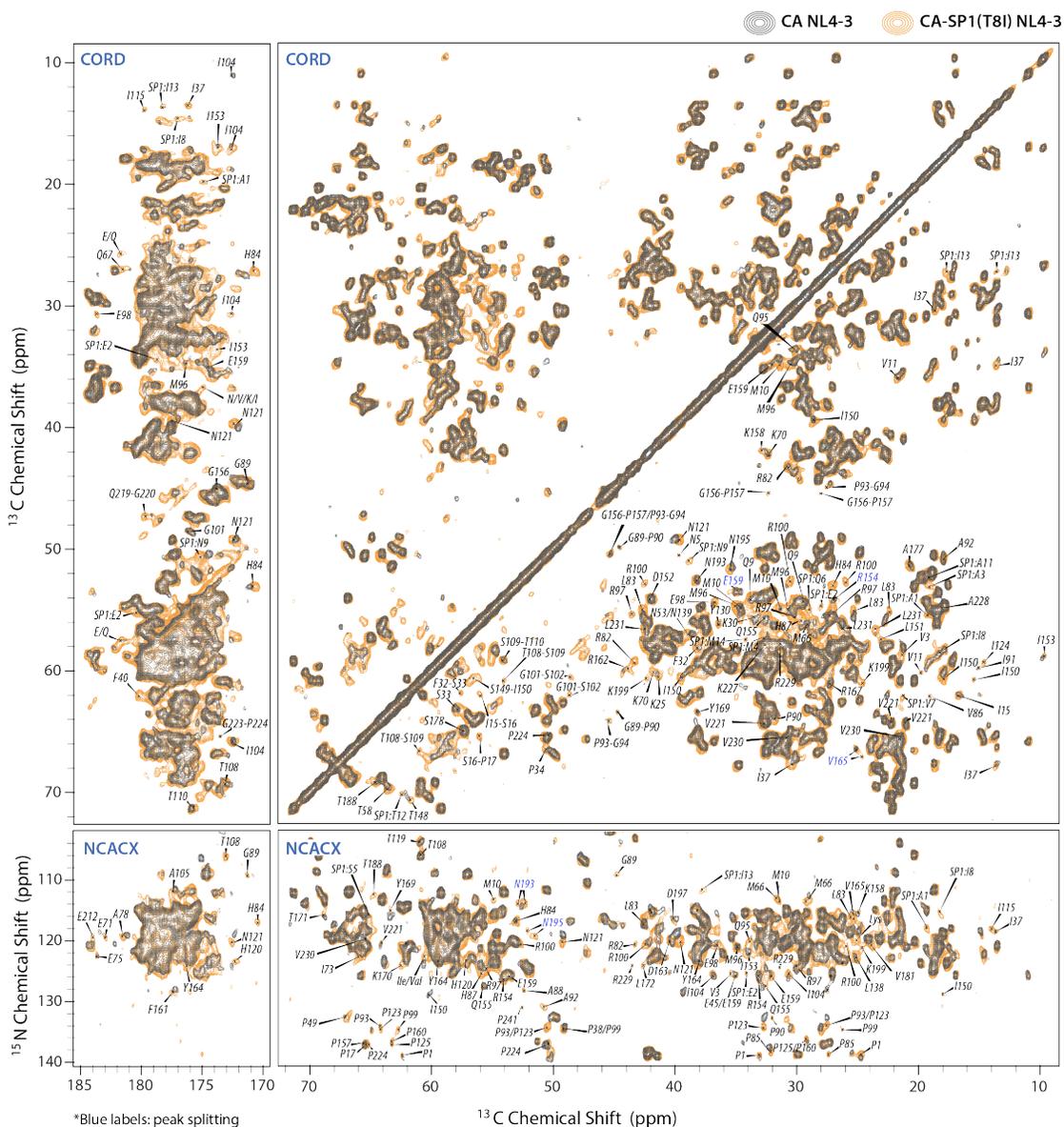
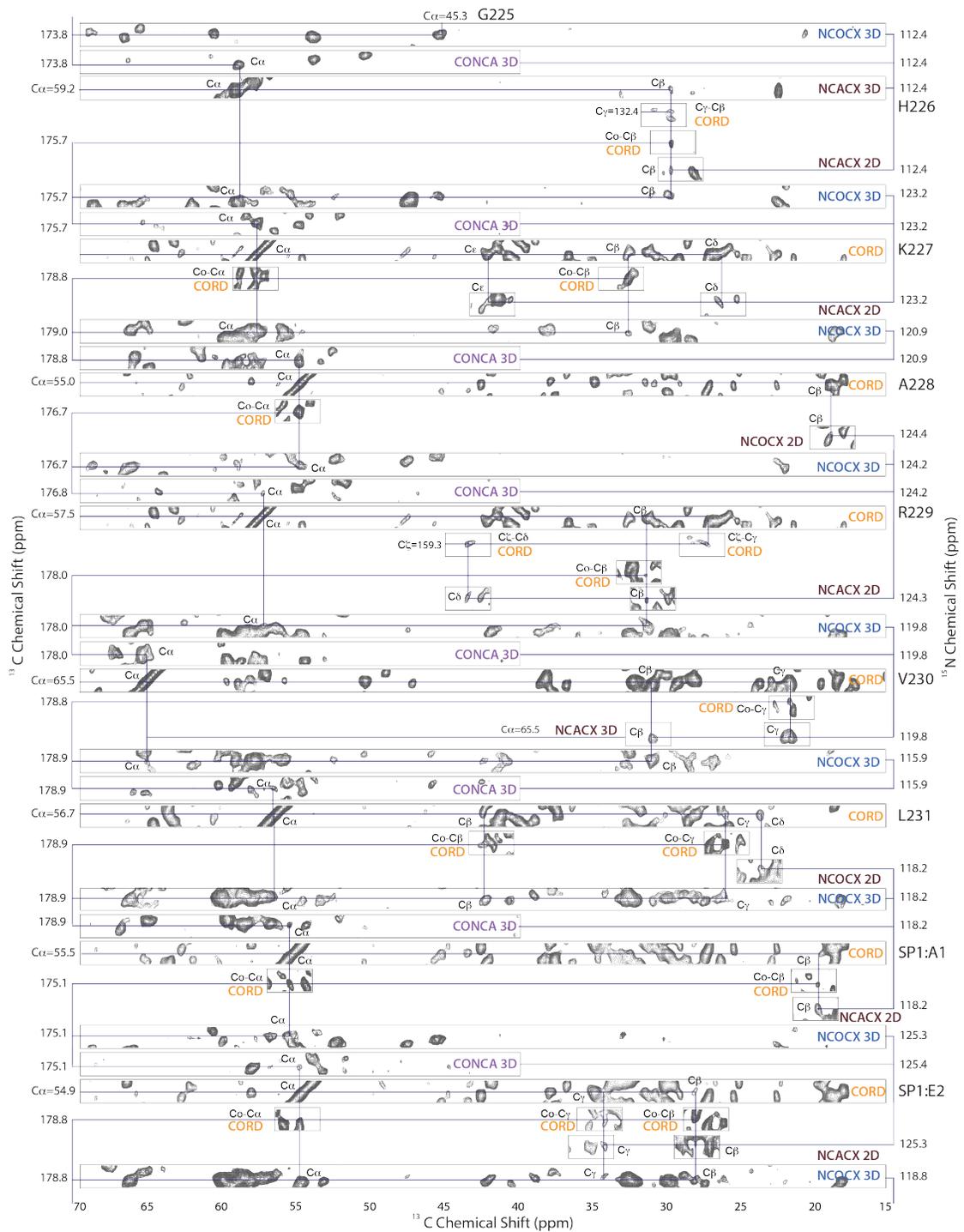
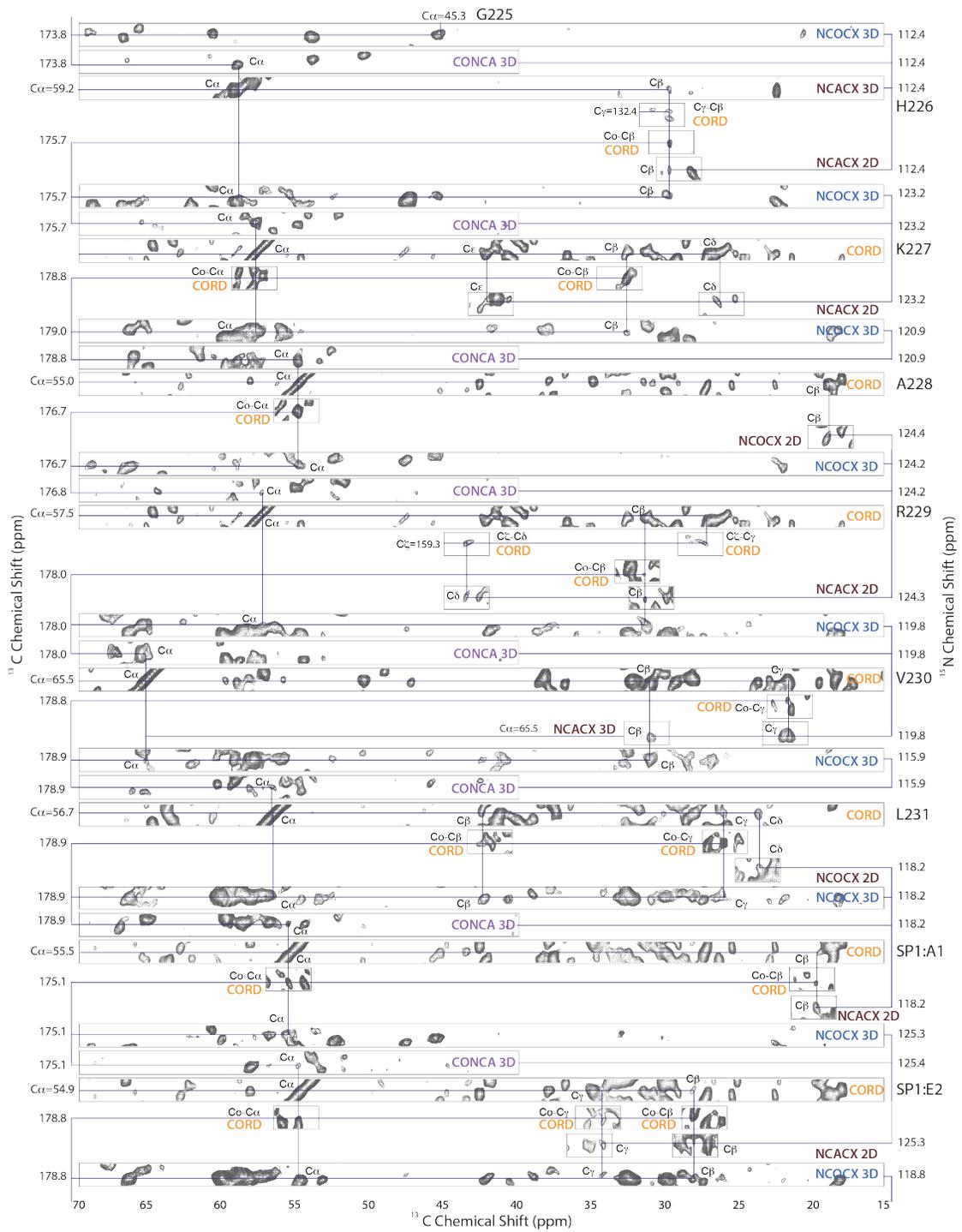
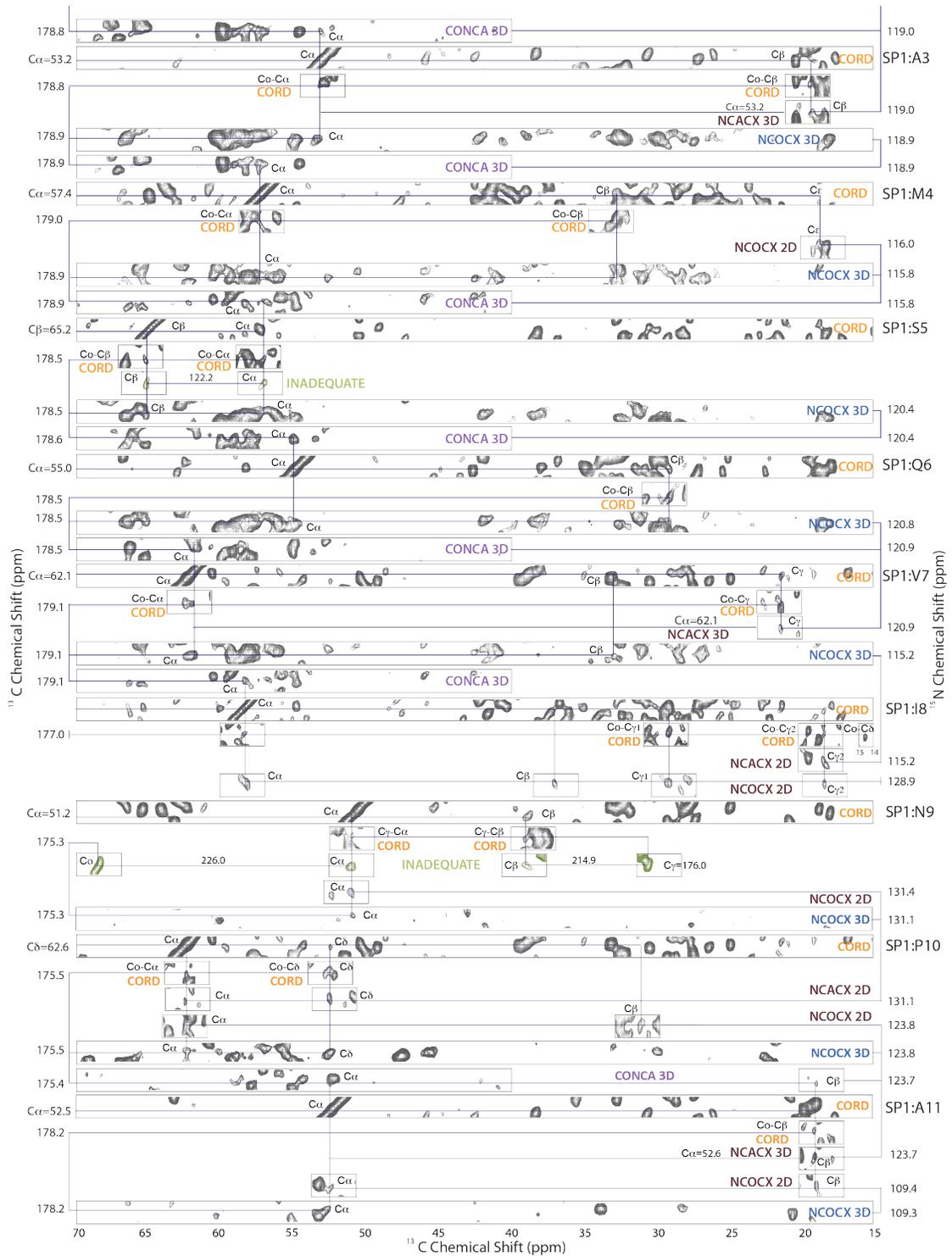


Figure 4.2 Superposition of dipolar-based spectra of HIV-1 CA and CA-SP1(T8I) tubular assemblies.<sup>9</sup> 2D CORD (top) and NCACX (bottom) MAS NMR spectra for CA NL4-3 (gray) and CA-SP1(T8I) NL4-3 (orange) illustrate the multiple chemical shift perturbations and peak intensity differences. The residues that performed peak splitting are labeled in blue. The spectra were acquired at 20.0 T and the MAS frequency of 14 kHz, with 60° shifted sine bell apodization process in both dimensions. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.







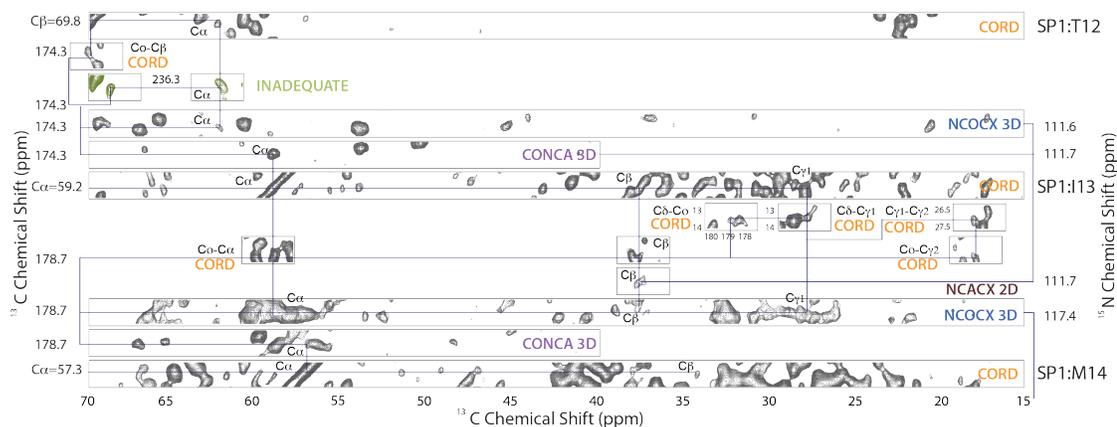


Figure 4.3 A stretch of sequential walk for the CTD-SP1 region (residues CA: G225-L231 to SP1: A1-M14) of CA-SP1(T8I) tubular assemblies.<sup>9</sup> These residues were assigned by a combination of various types of MAS NMR spectra, including 2D CORD (orange labels), 2D/3D(NUS) NCACX (brown/blue labels, respectively), 2D/3D NCOCX (brown/blue labels, respectively), 3D CONCA (purple labels), and 2D direct-INADEQUATE (green labels). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

Table 4.1 Chemical shifts of the CTD-SP1 region (CA residues 221-231, SP1 residues 1-14) of CA-SP1(T8I) assemblies determined by MAS NMR experiments. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

Sub-domain	Residue	N	C $\alpha$	C $\beta$	C $\gamma$	C $\delta$	C $\epsilon$	C $\zeta$
CA	V221	120.53	176.98	64.29	32.72	22.18/ 21.04	--	--
	G222	117.30	173.75	45.13	--	--	--	--
	G223	112.59	173.82	45.23	--	--	--	--
	P224	137.09	177.61	65.34	32.38	28.05	50.41	--
	G225	116.27	173.80	45.35	--	--	--	--
	H226	112.55	175.70	59.31	29.88	132.41	--	--
	K227	123.19	179.00	57.90	32.59	26.16	29.89	42.09
	A228	120.89	176.71	54.88	18.09	--	--	--
	R229	124.17	177.96	57.54	31.37	27.10	43.63	--
	V230	119.79	178.86	65.47	31.27	21.72	--	--
	L231	115.90	178.92	56.83	42.34	26.23	23.47	--
SP1	A1	118.24	175.09	55.70	19.55	--	--	--
	E2	125.35	178.80	54.89	28.09	34.11	--	--
	A3	118.88	178.90	53.14	19.10	--	--	--
	M4	118.89	179.03	57.42	32.54	--	--	18.64
	S5	115.76	178.53	57.05	65.2	--	--	--
	Q6	120.47	178.52	54.99	29.04	--	--	--
	V7	120.50	179.06	62.10	33.09	21.22	--	--
	I8	115.29	177.05	58.52	36.85	29.17/ 18.26	14.6	--
	N9	128.80	175.30	51.10	39.13	176	--	--
	P10	131.10	175.52	62.57	31.21	28.78	52.57	--
	A11	123.70	178.19	52.54	18.92	--	--	--
	T12	109.60	174.40	62.20	69.82	--	--	--
	I13	111.60	178.65	59.22	37.75	27.59/ 17.87	13.77	--
	M14	117.40	--	57.33	34.16/ 33.69	--	--	--

### 4.3.2 Conformation of SP1 in the Wild Type and T8I Mutant

SP1 plays an essential role in the HIV maturation with its highly dynamic behavior; however, its structure and dynamics remained elusive. Our collaborators, Dr. Peijun Zhang's group, analyzed a WT-like construct, CA(A92E)-SP1 protein tubes using Cryo-EM.<sup>9</sup> It was found that the SP1 domain did not present any detectable electron density, suggesting a conformationally or dynamically disordered structure rather than a stable helical structure. In a solution NMR investigation of WT CTD-SP1 protein from Dr. Angela Gronenborn's group, SP1 showed a weak helical propensity with the dominant coil conformation, as shown in Table 4.2. These results were in agreement with our findings on WT and WT-like CA-SP1 protein assemblies from MAS NMR<sup>5</sup>.

Comparison of MAS NMR spectra of CA-SP1 WT and CA-SP1(T8I) assemblies reveals that peaks of the CTD tail and SP1 are present in dipolar-based correlation spectra of SP1 T8I assemblies, but absent those for the WT CA assemblies (Fig 4.4, right panels). This result indicates that CTD tail and SP1 in the CA-SP1(T8I) protein are rigid on the microsecond timescale. Conversely, the same region of CA-SP1 WT undergoes substantial motions on the timescale of microseconds. In J-based experiments<sup>20</sup>, only residues that remain mobile on the micro- to millisecond timescales exhibit signals. The resonances of SP1 were detected in both WT and T8I mutant, demonstrated in <sup>13</sup>C-<sup>13</sup>C direct-INADEQUATE superposition spectra (Fig 4.5). Interestingly, the weak peak intensity and the missing peaks were observed for the CTD (H226, R229) tail and SP1 (Q6, N9, T12, I13) region in the CA-SP1(T8I) with respect to the CA-SP1 WT. This result indicates that T8I greatly reduces the mobility in the CTD tail and SP1 region on the micro- to millisecond timescale.

Furthermore, the analysis of the MAS NMR spectra indicated that for CA-SP1(T8I) assemblies, T8I not only stabilizes the CTD tail and SP1 region, but also induces a significant helical conformation for residues 226-231 (CA CTD) and 1-6 (SP1), determined by the MAS NMR chemical shifts (Table 4.2).

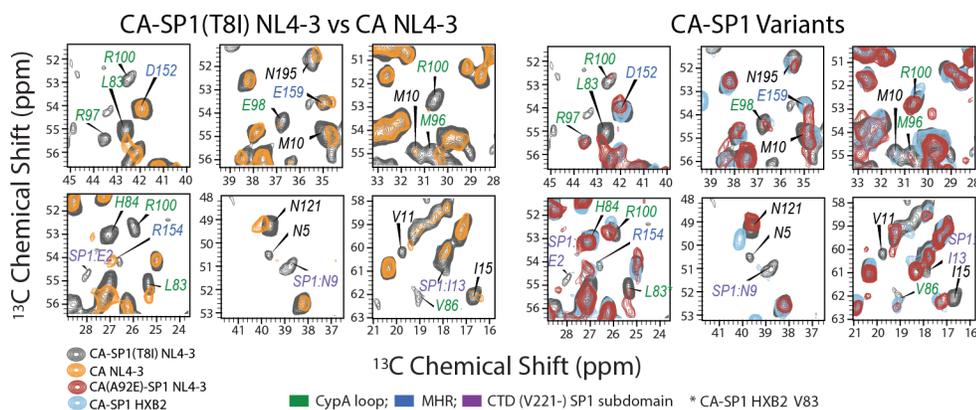


Figure 4.4 Superposition of 2D  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra for CA and CA-SP1 variants, highlighting SP1 resonances and residues exhibiting chemical shift or peak intensity changes. Left expansions: CA NL4-3 (orange) and CA-SP1(T8I) NL4-3 (gray). Right expansions: CA(A92E)-SP1 NL4-3 (red), CA-SP1 HXB2 (blue) and CA-SP1(T8I) NL4-3 (gray). Chemical shift perturbations are present in several residues of the CTD (V221)-SP1 tail, CypA loop, MHR, and NTD  $\beta$ -hairpin. All spectra were acquired at 20.0 T and an MAS frequency of 14 kHz. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.



Table 4.2 CSI 2.0<sup>21</sup> and TALOS-N<sup>22</sup> secondary structure prediction based on experimental chemical shifts for CA-SP1(T8I) NL4-3, CA(A92E)-SP1 NL4-3, and CA-SP1 HXB2. Designations for C-coil and H-helix are as described in Materials and Methods. "--" indicates predictions that are less reliable due to missing chemical shifts. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

		CA-SP1(T8I)		CA(A92E)-SP1		CA-SP1		CA CTD-SP1 NL4-3	
		NL4-3		NL4-3		HXB2		solution NMR	
Sub-domain	Residue	CSI 2.0	TALOS -N	CSI 2.0	TALOS -N	CSI 2.0	TALOS -N	CSI 2.0	TALOS -N
CA	V221	C	C	C	C	C	C	C	C
	G222	C	C	C	C	C	C	C	C
	G223	C	C	C	C	C	C	C	C
	P224	C	C	--	--	--	C	C	C
	G225	C	C	--	--	--	C	C	C
	H226	H	H	C	--	C	C	C	C
	K227	H	H	C	C	C	--	C	C
	A228	H	H	C	--	--	--	C	C
	R229	H	H	C	H	C	--	C	C
	V230	H	H	H	H	--	--	C	C
	L231	H	H	H	H	C	--	C	C
SP1	A1	H	H	H	H	H	--	C	C
	E2	H	H	H	H	H	H	C	H
	A3	H	H	C	--	H	--	C	H
	M4	H	H	C	--	--	--	C	C
	S5	H	H	C	--	--	--	C	C
	Q6	H	C	C	--	--	--	C	C
	V7	C	C	--	--	--	--	C	C
	I8/T8	C	C	C	--	C	--	C	C
	N9	C	C	C	--	--	--	C	C
	P10	C	C	--	--	--	--	H	C
	A11	C	C	C	--	C	--	H	C
	T12	C	C	C	--	--	--	H	C
	I13	C	C	C	--	--	--	C	C
	M14	C	C	C	--	--	--	C	C

### 4.3.3 Allosteric Conformational and Dynamic Changes in CA Induced by T8I Mutant

In addition to the residues in the CTD tail and SP1 subdomain, in the vicinity of the mutation sites, the T8I mutation also remarkably affects the conformation and dynamics in other regions of CA, such as NTD  $\beta$ -hairpin (e.g. M10, V11), CypA loop (e.g., L83, H84, R100), and MHR (e.g., R154, E159). The differences of chemical shift and peak intensity were observed in Fig 4.2 and Fig 4.4 (left panels).

As shown in Fig 4.6, in order to understand the reason of these distal changes, three pairs of CA-SP1 and corresponding CA tubular assemblies were quantitatively analyzed by calculating the sum of backbone  $^{13}\text{C}\alpha$  and  $^{15}\text{N}$  chemical shift perturbations (CSPs). In the tubular assemblies of CA-SP1 HXB2 (panels II) and CA(A92E)-SP1 NL4-3 (panels III) pairs, a number of residues exhibit sizeable (greater than 0.3 ppm) chemical shift perturbations, revealing that SP1 modulates the structure of CA domain. Remarkably, in the comparison of CA WT and CA-SP1(T8I) (panel I), a large number of residues have significant chemical shift perturbations (1-1.5 ppm), indicating conformational changes that are induced by this maturation-inhibiting T8I mutation. These conformational effects are linked to a change in the overall dynamics of the CA domain as well. Dynamic changes on the timescale of micro- to millisecond were investigated by the deviation of peak intensities in 2D NCA spectra (panel IV). Increased peak intensities are indicative of lower mobility throughout the CA domain in CA-SP1(T8I) protein assemblies. This overall reduced dynamics is also supported by the smaller number of resonances in J-based correlation MAS NMR spectra of tubular CA-SP1(T8I) assemblies, compared to those of CA(A92E)-SP1 (Fig 4.5). The regions that are affected by the presence of the T8I mutation include the N-terminal  $\beta$ -hairpin, the Cyclophilin A-binding loop (CypA

loop, loop 4/5), loop 6/7 in the NTD, as well as the MHR, loops 9/10 and 10/11 in the CTD. Among these regions, the conformation and dynamics of CypA loop is perturbed by the greatest extent by the T8I mutation.

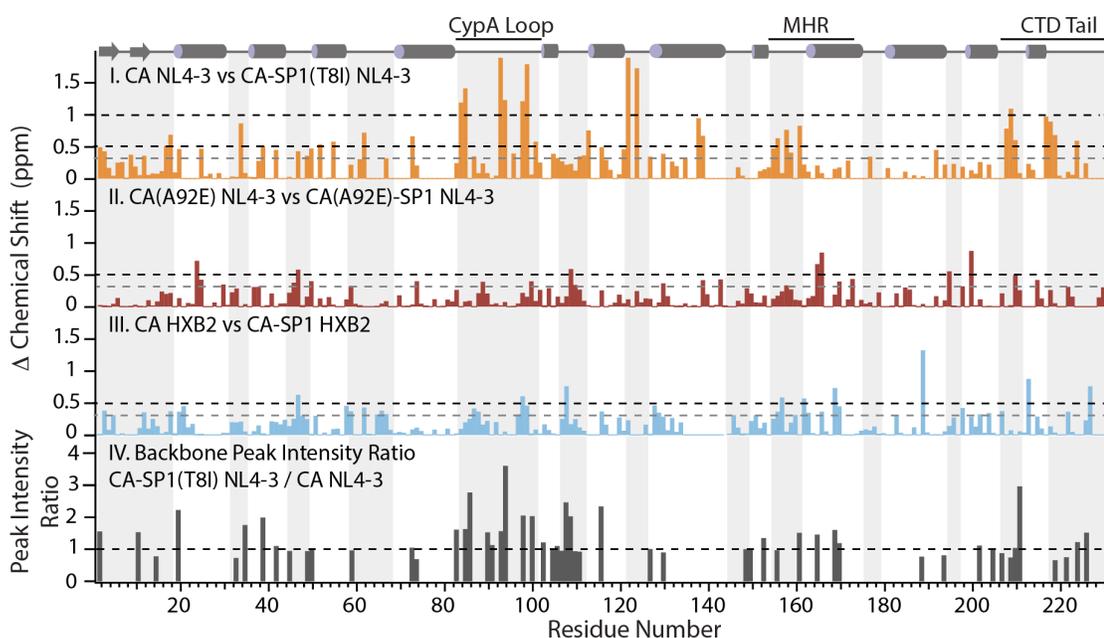


Figure 4.6 Chemical shift changes between CA and CA-SP1 plotted along the linear amino acid sequence I, II, III: sum of  $^{13}\text{C}\alpha$ ,  $^{15}\text{N}$  chemical shift perturbations (CSP): CA NL4-3 and CA-SP1(T8I) NL4-3 (orange), CA(A92E) NL4-3 and CA(A92E)-SP1 NL4-3 (brown), CA HXB2 and CA-SP1 HXB2 (blue). Note that a CSP  $> 0.5$  ppm is considered a significant perturbation, while a CSP  $< 0.3$  ppm (dashed gray) is negligible within experimental and systematic error. IV: Plot (grey) of the  $^{13}\text{C}\alpha$  -  $^{15}\text{N}$  backbone peak intensity ratio (CA-SP1(T8I) NL4-3/CA NL4-3) vs. residue number (for non-overlapping peaks with resonance assignments in the 2D NCA MAS NMR spectra). Residues with peak intensity ratio  $> 1$  indicate attenuated motions on micro- to millisecond timescales. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

In the HIV-1 replication cycle, the CypA loop in the assembled capsid is highly flexible and plays an important role by modulating loop dynamics in the HIV-1's escape from CypA dependence<sup>7</sup>. To detect the loop dynamics on the timescale of nano- to microseconds, <sup>1</sup>H-<sup>15</sup>N dipolar lineshapes and order parameters were recorded and analyzed for CA-SP1(T8I) mutant and CA WT assemblies (Fig 4.7). On these timescales, significantly attenuated CypA loop dynamics were observed in CA-SP1(T8I) mutant, as is obvious from the broader <sup>1</sup>H-<sup>15</sup>N lineshapes and increased dipolar order parameters. Surprisingly, this reduced mobility in CypA loop is of the same order of magnitude as the decreased loop dynamics that were previously observed in the A92E and G94D escape mutants of CA<sup>7</sup>. A92E and G94D mutations have been found to cause large loss of infectivity *in vivo* that is rescued upon inhibition of Cyclophilin A by Cyclosporin. However, in contrast, CA-SP1 WT assemblies showed moderately increased dynamics in CypA loop, compared to the corresponding CA assemblies<sup>5</sup>, which possibly suggests that NTD-NTD contacts may be modulated by this T8I mutation<sup>23,24</sup>.

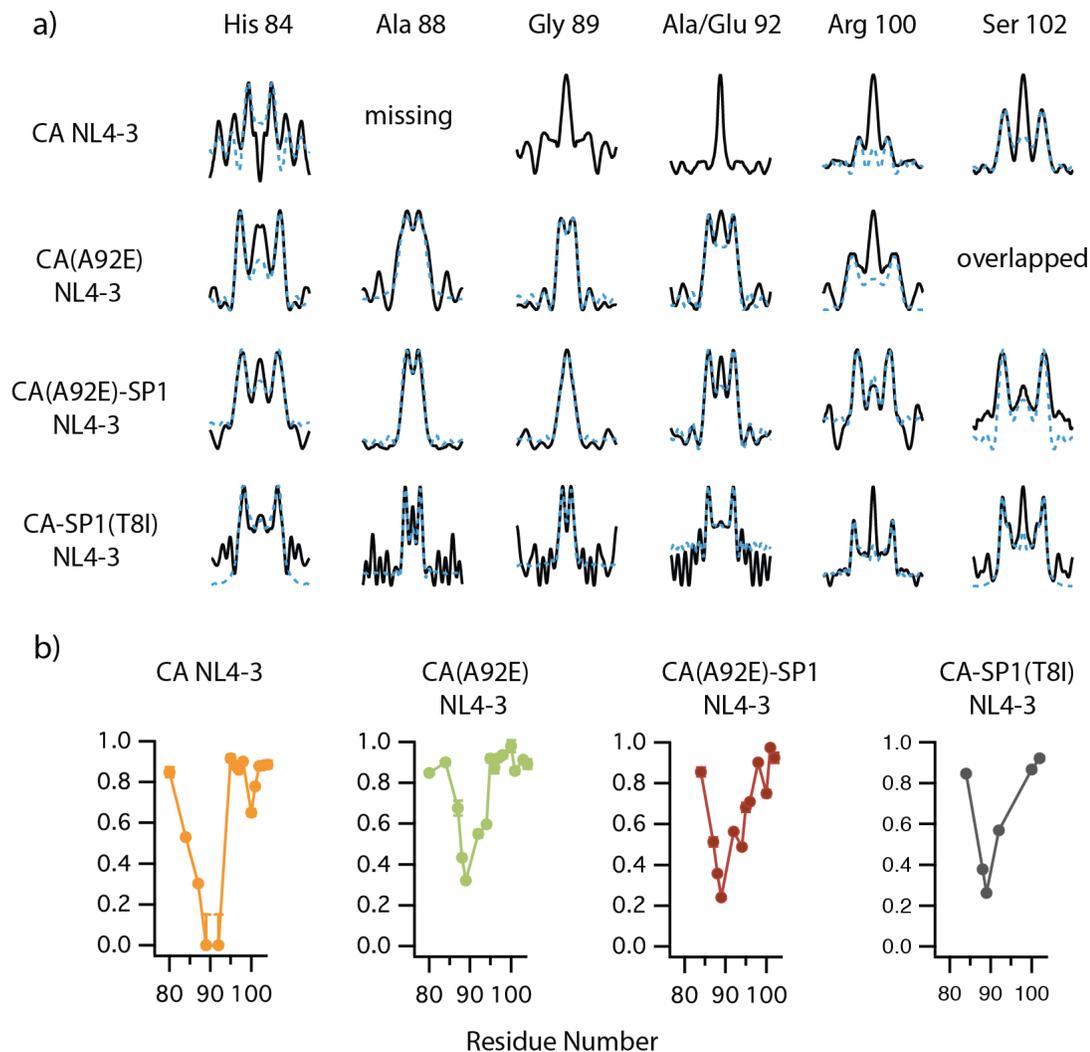


Figure 4.7 Comparison of nano-to-microsecond timescale dynamics for CypA loop residues observed in tubular assemblies of HIV-1 CA and CA-SP1 variants. (a) Experimental (solid black) and simulated (dashed blue)  $^1\text{H}$ - $^{15}\text{N}$  dipolar lineshapes for CypA loop residues in (top to bottom): CA NL4-3, CA NL4-3 A92E, CA-SP1 NL4-3 A92E and CA-SP1 NL4-3 T8I. (b)  $^1\text{H}$ - $^{15}\text{N}$  Dipolar order parameters are plotted vs. residue number for CypA loop residues in CA and CA-SP1 tubular assemblies. The CA-SP1(T8I) mutant exhibits the same order of magnitude attenuation of loop dynamics as previously observed in the CA(A92E) escape mutant<sup>7</sup>.

The N-terminal  $\beta$ -hairpin is stabilized in the CA-SP1(T8I) assemblies relative to WT CA assemblies. The conformation of the N-terminal residues in the CA-SP1 lattice is debated currently: it has been hypothesized that the  $\beta$ -hairpin is present only in the mature CA<sup>23</sup>, while other studies support that it is formed when cleavage between the MA and CA domains occurs<sup>25,26</sup>. The  $\beta$ -hairpin conformation is critical to the central intra-hexamer pore size and possibly the capsid's permeability to nucleotides<sup>27</sup>. CSI analysis of MAS NMR chemical shifts indicates that the N-terminal residues form a  $\beta$ -hairpin in all CA-SP1 assemblies inspected here (Table 4.3), implying that full capsid maturation is not required for the formation of the  $\beta$ -hairpin. Significant chemical shift perturbations are also observed in loop 6/7, which is at the trimer contact of the immature lattice<sup>24</sup>, suggesting a possibly altered trimer interface contacts induced by T8I mutation. Remarkably, the dramatic chemical shift changes and altered dynamics are present in the NTD of CA-SP1(T8I), while CA-SP1 HXB2 and CA(A92E)-SP1 show much smaller chemical shift changes in the NTD. These findings indicate that the T8I mutation has a more pronounced influence on the NTD conformation compared to moderate changes arising from the presence of the SP1 peptide alone. Taken together, the MAS NMR results demonstrate that the NTD conformation in assembled CA-SP1(T8I) is distinct from that in WT CA-SP1 as well as WT CA assemblies.

Table 4.3 Secondary structure predictions of the NTD determined from MAS NMR experiments in CA-SP1 assemblies. CSI 2.0<sup>28</sup> and TALOS-N<sup>22</sup> programs were used to predict secondary structures by CSI and torsion angles, respectively. "--" The predictions are less reliable due to the missing chemical shifts. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

Residue	CA-SP1(T8I) NL4-3		CA(A92E)-SP1 NL4-3		CA-SP1 HXB2	
	CSI 2.0	TALOS-N	CSI 2.0	TALOS-N	CSI 2.0	TALOS-N
P1	C	C	C	--	C	C
I2	B	C	B	--	B	C
V3	B	B	B	B	B	B
Q4	B	B	C	B	C	B
N5	C	C	C	C	C	C
L6/16	H	H	C	C	C	H
Q7	C	H	C	H	C	H
G8	C	H	H	H	C	H
Q9	C	C	C	C	C	C
M10	C	B	B	B	B	B
V11	B	B	B	B	B	B
H12	B	B	B	B	B	B
Q13	B	B	B	B	B	B
A14	C	B	B	B	B	B
I15	C	C	C	C	C	C
S16	C	C	C	C	C	C
P17	H	H	H	H	H	H
R18	H	H	H	H	H	H
T19	H	H	H	H	H	H
L20	H	H	H	H	H	H

Even though it is not surprising to observe conformational influence in the C-terminal domain of CA, in the vicinity of the mutation site, more pronounced chemical shift changes are observed in CA-SP1(T8I) mutant assemblies due to the presence of the T8I mutation, compared to CA-SP1 WT assemblies. The affected residues are mainly located in functionally relevant regions, including the highly conserved MHR

(residues 153-172), and loops 9/10 and 10/11. Specifically, the T8I mutation of the SP1 subdomain introduces increased conformational and dynamic heterogeneity (multiple conformers are detected) into the MHR (e.g., R154, E159, V165), compared to CA as well as other CA-SP1 assemblies under investigation (Fig 4.2 and 4.4). The MHR is reported to be involved in stabilizing intra-hexameric interactions in the immature lattice<sup>24</sup>. In addition, replication defects resulting from PF96-dependent resistance mutations in the MHR can be compensated by mutations in the CTD tail or SP1 subdomain (such as T8I), suggesting that an allosteric ‘cross-talk’ between the SP1 and MHR<sup>29</sup> may exist. Chemical shift perturbations were also observed in loops 9/10 and 10/11, which have also been suggested to play a role in the formation of the immature lattice, and the dimer interface (helix 9) shows distinct conformation in the immature lattice and in the mature lattice<sup>24,30</sup>. While we did not observe significant dynamic changes of the residues in the loops 9/10 and 10/11 that are associated with helices 9 and 10, chemical shift changes unequivocally indicate conformational differences, which may indeed be due to altered positioning of the adjacent helices. Therefore, the observed perturbations in the CA-SP1(T8I) assemblies suggest that the T8I mutation (phenocopying maturation-inhibitor binding) affects the structure of those CTD regions that undergo critical conformational rearrangements during maturation. Moreover, they have important implications for our understanding of maturation as well as the function of maturation inhibitors, as further discussed below.

#### **4.3.4 Investigation of Maturation by an Integrated MD and MAS NMR Approach**

##### **4.3.4.1 Conformation of SP1 is Modulated by T8I Mutation in MD Simulations**

Our prior and current MAS NMR results prove that the SP1 subdomain is highly mobile and in a predominantly random coil conformation in WT and WT-like CA-SP1 assemblies (CA-SP1 HXB2 and CA(A92E)-SP1 NL4-3)<sup>5</sup>, but adopts a stabilized helix in assemblies of CA-SP1(T8I) from the NL4-3 strain. To further understand the structure and motions of the SP1 region, our collaborator, Dr. Juan Perilla, performed simulated tempering MD calculations<sup>31</sup> over 15-24  $\mu$ s of a hexameric subunit of CTD-SP1 as well as the CTD-SP1(T8I), Fig 4.8a-b. Strikingly, in complete agreement with the MAS NMR data, these simulations reveal a highly mobile SP1 region for the CA-SP1 WT hexamer, exhibiting a dynamic equilibrium between random coil and helical conformations, with the helix a minor, transient form. Additionally, apparently frequent, transient contacts between SP1 and helix 8 of the MHR were observed along with the entire MD trajectory. In contrast, MD simulations of the CA-SP1(T8I) mutant reveal reduced dynamics and increased helical propensity.

To visualize the changing secondary structures in the MD trajectory, Dr. Huilan Zhang generated secondary structure propensity plots using STRIDE<sup>19</sup> from each single trajectory frame for CA-SP1 WT and CA-SP1(T8I) mutant. As shown in Fig 4.8c-d, CA-SP1 WT performs helix-coil transitions in five of six chains (Chain 1-3, and 5-6) in the MD simulations, and in contrast, CA-SP1(T8I) forms stable helices that persist throughout the entire length of the MD trajectory in two of the six chains (Chain 1 and 3). To quantitatively assess the magnitude of conformational changes in each MD frame, secondary structure probability plots were depicted, associated with residues in the region of the CTD tail and SP1 subdomain, Fig 4.9. The significantly

increased probability of helical structure was observed in the residues of L231 (CA) to Q6 (SP1) due to the presence of T8I mutation. Furthermore, more residues in the CTD tail, H226 to V230, exhibit highly helical propensity, which is in agreement with our MAS NMR results, Table 4.2. Our collaborator, Dr. Juan Perilla generated 7 clusters determined using the partition around medoids protocol<sup>32</sup>, classifying a group of structures that exhibit similar conformation to the individual sub-populations of the MD trajectory for CA-SP1 WT and CA-SP1(T8I) (Fig 4.10). In CA-SP1 WT structures, SP1 exhibits a partially helical conformation in two clusters (Cluster 1 and 4), and displays pre-dominantly random coil structure in the rest of the clusters. In contrast, stable helix is present in SP1 region in six of seven clusters for the CA-SP1(T8I) mutant.

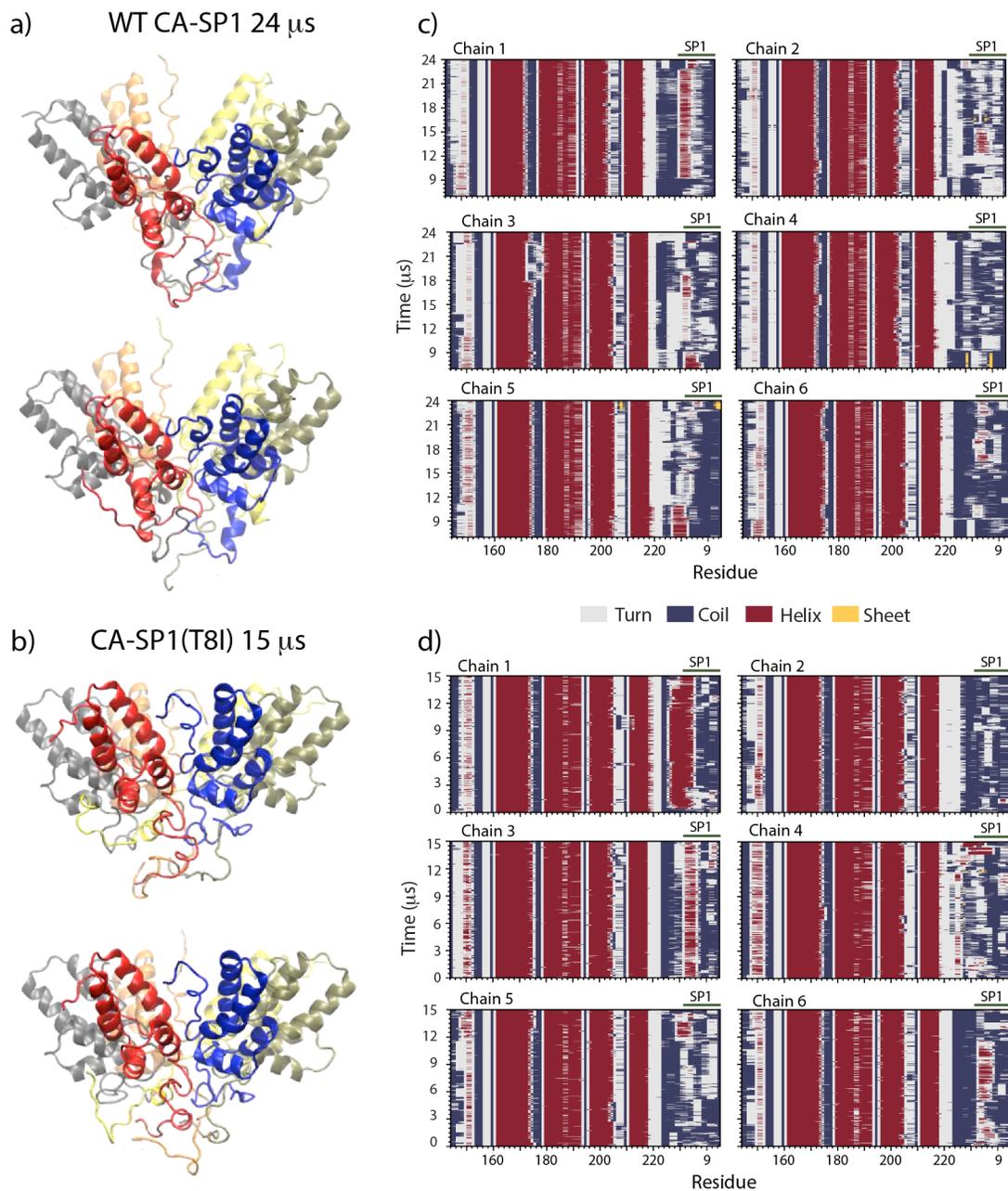


Figure 4.8 Two typical MD simulation frames showing hexameric subunits of CTD-SP1 (a) and the CTD-SP1(T8I) (b). Stride plots of secondary structure for each chain along with the MD trajectories for CA CTD-SP1 WT (c), and CA CTD-SP1(T8I) mutant (d). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

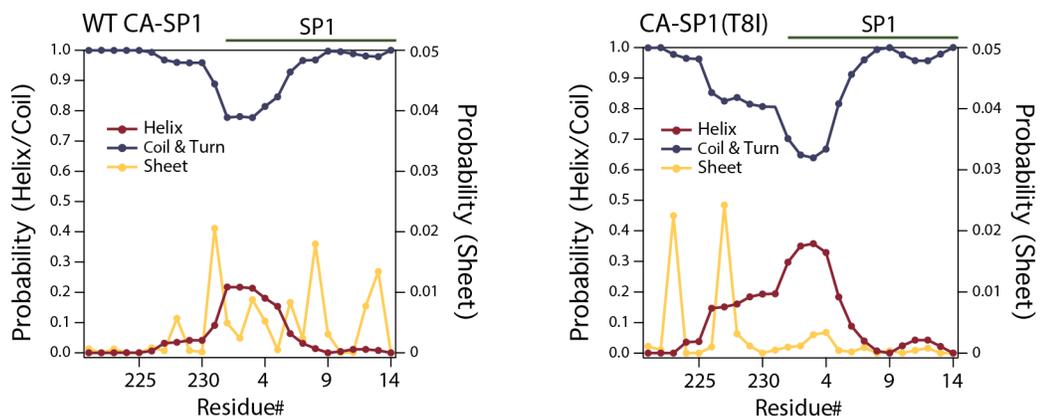


Figure 4.9 Helix/coil/sheet probability plots of the CTD (V221)-SP1 subdomain averaged over the MD trajectories: CA CTD-SP1 WT (left), and CA CTD-SP1(T8I) (right). The expanded scale (0-0.05; right hand side) is shown for the “Sheet” content. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

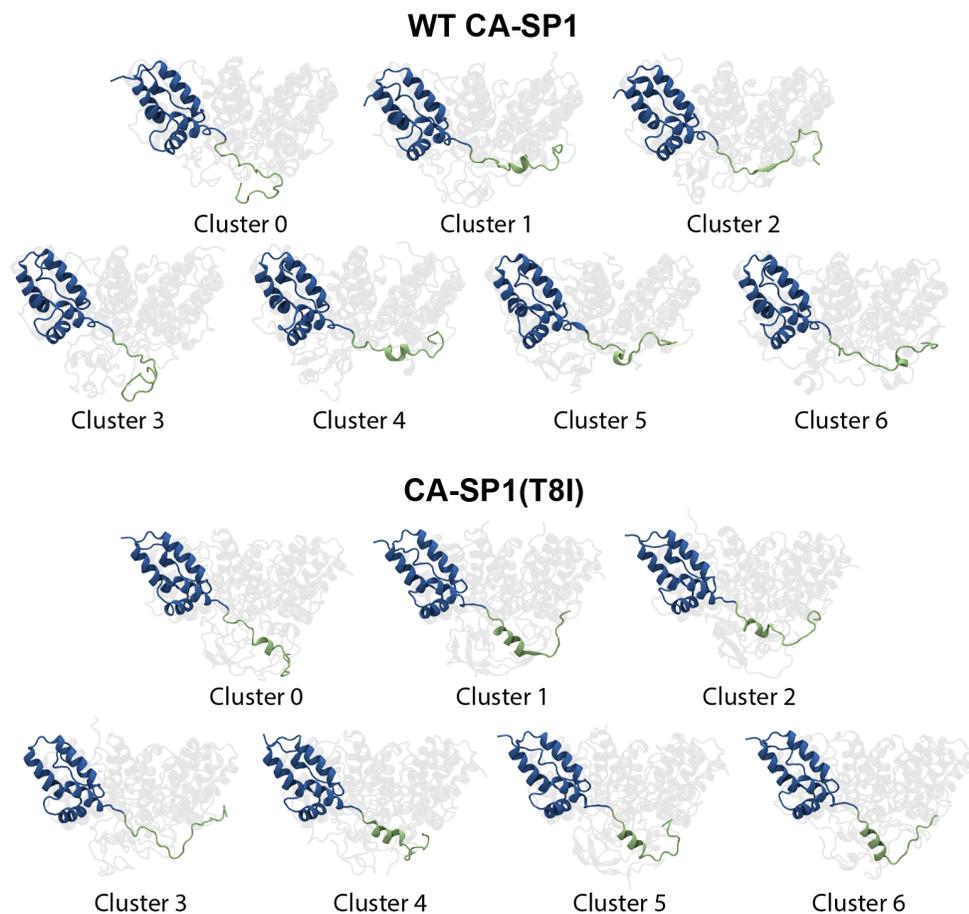


Figure 4.10 Clustering analysis of the MD trajectory of CA(CTD)-SP1 WT (top) and CA(CTD)-SP1(T8I) mutant (bottom) identified six major sub-populations. Structure clusters were extracted from the MD trajectory. One chain is highlighted in blue to illustrate the differences in secondary structure present in the SP1 region. CTD(V221)-SP1 segment is colored in green. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

#### 4.3.4.2 Quantitative Analysis of Isotropic Chemical Shift: MD vs. MAS NMR

A quantitative analysis has been carried out for experimental isotropic MAS chemical shifts and calculated ones from the MD trajectory of CA-SP1 assemblies. Using the database approach implemented in SHIFTX2<sup>33</sup>, we calculated backbone N and  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{13}\text{C}\gamma$  chemical shifts from molecular coordinates for every MD frame of each of the six WT CTD-SP1 chains. We employed two series of MD calculated shifts compared with experimental shifts: i) averaged MD trajectory over all frames and all chains of the hexamer; ii) individual seven sub-populations (clusters) (Fig 4.10) of the MD trajectory both for WT CA-SP1 and T8I mutant. The root mean square deviation (RMSD) of each atom,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{13}\text{C}\gamma$  and  $^{15}\text{N}$  are listed in Table 4.4. Overall good agreement was observed between the shifts predicted from MD trajectories and the experimental values: the typical RMSD were within the errors expected of SHIFX2 predictions, 3.0 ppm ( $^{15}\text{N}$ ) and 1.4 ppm ( $^{13}\text{C}$ )<sup>18,34</sup>. In CA-SP1 WT assemblies,  $\text{C}\alpha$ , as a representative demonstrates an excellent agreement between averaging MD calculated and experimental shifts (Fig 4.11). On the other hand, considerable scatter and poor correlations are observed when shifts are calculated for any individual sub-population of the MD trajectory. This result not only unequivocally corroborates our prior experimental MAS NMR-based conclusion that the CTD tail and SP1 subdomain adopt dynamic helix-coil conformations, but also provides quantitative understanding on the contributions of the individual cluster types to the experimental chemical shifts. In our recent dynamic nuclear polarization (DNP) experiments<sup>35</sup>, we observed weak narrow peaks corresponding to a minor helical SP1 sub-population, further supporting this notion.

Additionally, a stable SP1 six-helix bundle was found in the recent X-ray structure of CTD-SP1<sup>36</sup>, which may be due to the presence of a helix-inducing solvent

(Bis-Tris propane) in the crystallization conditions. And possibly the presence of a N-terminal His tag may force SP1 into the minor helical state, which was fortuitously captured in the crystal. In this regard, we note that the SHIFTX2 predicted chemical shifts from the X-ray structure of the CTD-SP1 crystals (PDB: 5I4T) agree well with the MAS chemical shifts for the CA-SP1(T8I) mutant assembly, and not as well with WT CA-SP1 assemblies (Fig 4.11-12). It actively supports our assertion that a highly dynamic SP1 region is present in WT CA-SP1 assemblies. Another interesting finding in CA-SP1(T8I) assemblies is that the strongest correlations are observed for structures with greater helical content, including the average over all frames of the MD trajectory, as well as cluster 4, and 5 individually. This corroborates the previous conclusion that the predominant major structure observed by MAS NMR is helical, and suggests that longer MD simulations may be needed to possibly observe the six-helix bundle.

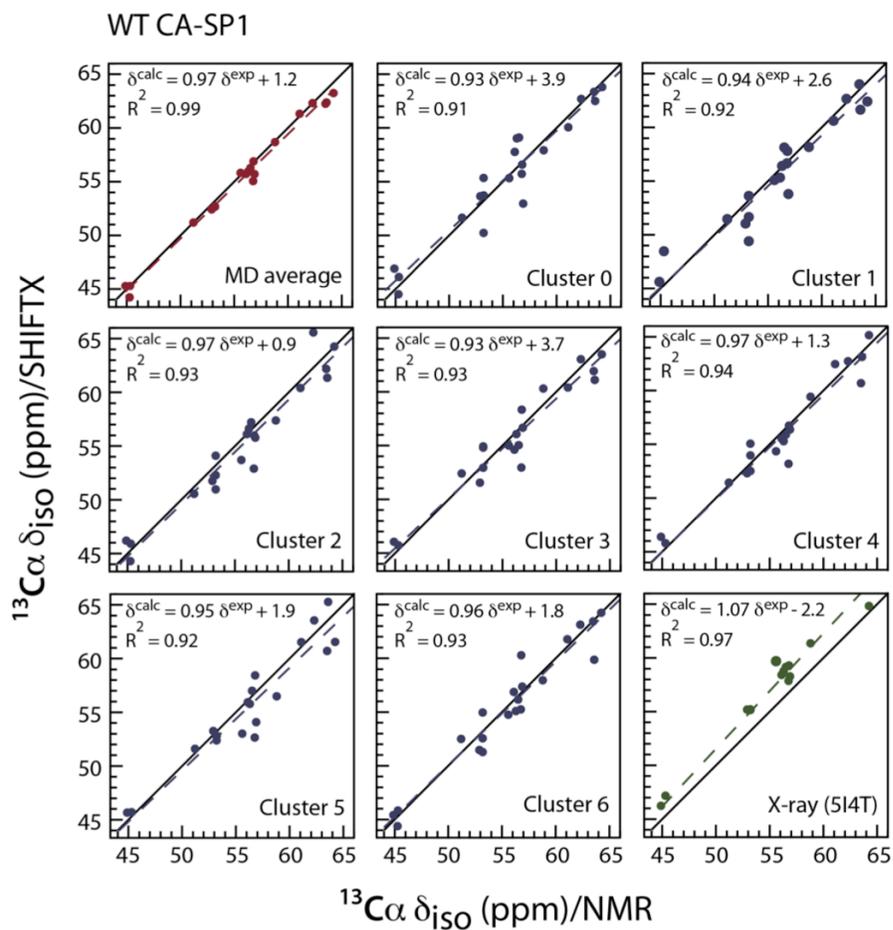


Figure 4.11 Correlation of MAS NMR chemical shifts with SHIFTX2 predicted shifts from MD trajectory and X-ray structures (PDB: 5I4T): CA CTD-SP1 WT. The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

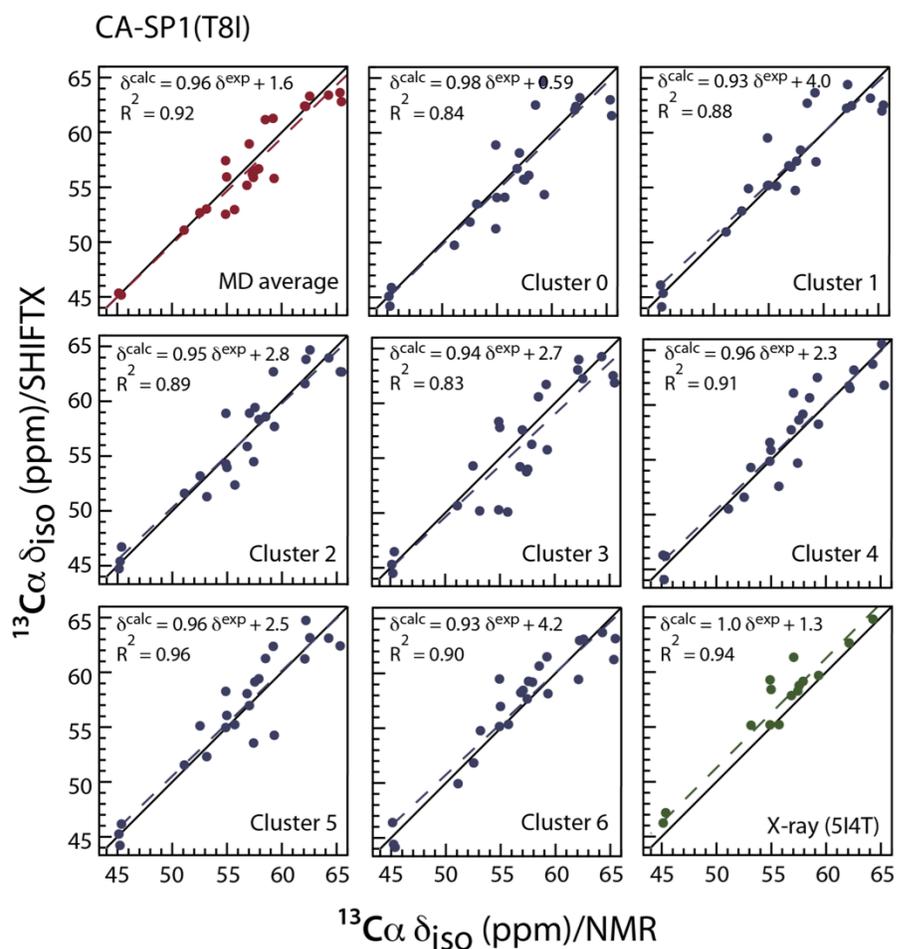


Figure 4.12 Correlation of MAS NMR chemical shifts with SHIFTX2 predicted shifts from MD trajectory and X-ray structures (PDB: 514T): CA CTD-SP1(T8I). The reuse in this dissertation is permitted by a Creative Commons Attribution 4.0 International License.

Table 4.4 An analysis of chemical shift (CS) RMSD values, comparing experimental shifts of MAS NMR with calculated shifts of MD simulations, as well as X-ray structure (PDB: 5I4T). The CS associated with the residues of CA(V221)-SP1 in CA(A92E)-SP1 and CA-SP1(T8I) assemblies were used for the comparison.

<b>WT CA-SP1</b>				
<b>Structure</b>	<b>C<math>\alpha</math> (ppm)</b>	<b>C<math>\beta</math> (ppm)</b>	<b>Co (ppm)</b>	<b>N (ppm)</b>
MD Average	0.5	0.6	1.6	0.9
Cluster 0	1.1	1.0	1.8	0.6
Cluster 1	1.1	1.0	1.6	0.8
Cluster 2	1.0	1.0	1.9	0.6
Cluster 3	0.9	0.9	1.9	0.6
Cluster 4	0.9	1.0	1.9	0.6
Cluster 5	1.1	0.9	1.7	0.7
Cluster 6	1.0	1.0	1.8	1.0
X-ray (5I4T)	0.9	0.6	1.7	1.1
<b>CA-SP1(T8I)</b>				
<b>Structure</b>	<b>C<math>\alpha</math> (ppm)</b>	<b>C<math>\beta</math> (ppm)</b>	<b>Co (ppm)</b>	<b>N (ppm)</b>
MD Average	1.0	0.6	1.1	3.1
Cluster 0	1.6	0.9	1.2	4.5
Cluster 1	1.5	0.8	1.4	5.1
Cluster 2	1.2	1.5	1.3	4.4
Cluster 3	1.5	1.0	1.4	2.9
Cluster 4	1.1	1.2	1.2	3.2
Cluster 5	1.3	0.6	1.3	3.4
Cluster 6	1.6	0.9	1.2	4.5
X-ray (5I4T)	1.3	0.6	1.3	3.0

#### **4.3.5 The Role of Dynamics in Capsid Maturation and Maturation Inhibition**

The combined results from MD and MAS NMR suggest that, notably, dynamics may play a determining role in the regulation of capsid maturation, with dynamic helix-coil equilibrium in the CTD-SP1 region of CA-SP1 being a key feature in the mechanism of the final maturation step. We hypothesize that the predominantly random coil conformation of CTD-SP1 region in CA-SP1 WT assemblies formed through this dynamic equilibrium, is important for correct and efficient processing of this Gag maturation. In the intermediate state, SP1 serves as a molecular switch that allows viral protease to recognize the cleavage site on the CTD-SP1 subdomain, leading to the formation of mature conical capsids in infectious virions, Fig 4.13. On the contrary, the T8I mutation, which phenocopies the MI-bound state, induces decreased dynamics and conformational changes not only in the CTD-SP1 region, but also in the CA lattice, and moreover, promotes a stable helical structure in the junction of CTD-SP1, Fig 4.13. The rigid helical conformation adopted by the mutant may be an impediment to proteolytic cleavage, the cleavage site remaining inaccessible for viral protease to prevent the completion of maturation. Interference with this SP1 dynamic equilibrium by the T8I mutation in the SP1 or small molecules (such as BVM) appears to be a promising strategy for developing the new class of anti-HIV therapeutics.

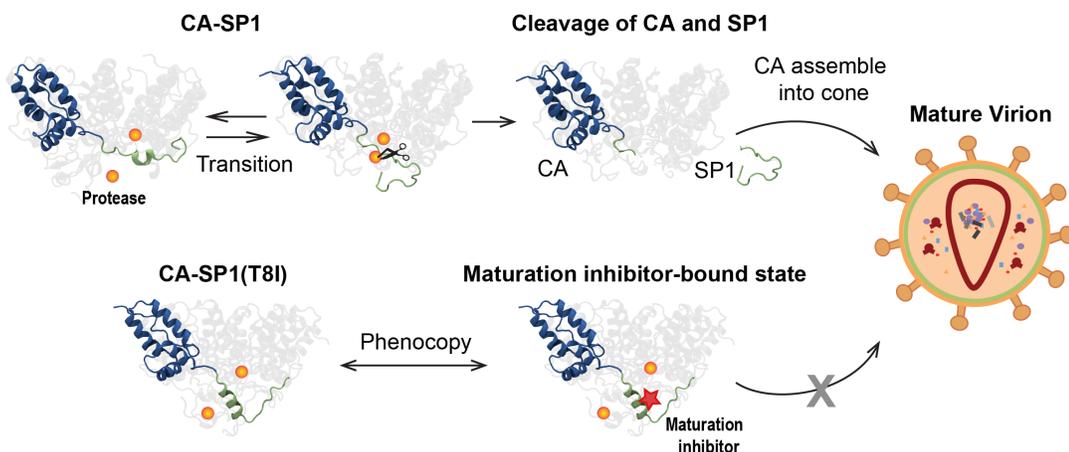


Figure 4.13 Schematic illustration of hypothesized mechanism: HIV-1 capsid maturation (top), and its inhibition by T8I mutation in SP1 or MI small-molecule. Involved elements are depicted, viral protease in orange sphere and MI in red star. One CA chain is highlighted in blue and CTD(V221)-SP1 segment is colored in green.

#### 4.4 Conclusions

Using the MAS NMR, we observed pronounced conformational and dynamics changes in the maturation-inhibiting CA-SP1(T8I) mutant, which provides key mechanistic insights into SP1 cleavage inhibition by maturation inhibitors. These residues in the CTD-SP1 junction exhibit increased helical propensity in the maturation-inhibiting CA-SP1(T8I) mutant, compared to CA-SP1 WT assemblies, and this increased helical content may be a key impediment to proteolytic cleavage. In contrast to CA-SP1(T8I), the SP1 subdomain of CA-SP1 WT exists in a highly dynamic helix-coil equilibrium, as shown by MAS NMR and MD studies. Such a flexible, dynamic conformation facilitates easy access of the PR to the cleavage site, permitting maturation to proceed. Taken together, our integrated experimental-computational approach yielded an atomistic understanding of the underlying

mechanistic details of HIV CA maturation and effects that MIs play in this important step.

## REFERENCES

- (1) Li, F.; Goila-Gaur, R.; Salzwedel, K.; Kilgore, N. R.; Reddick, M.; Matallana, C.; Castillo, A.; Zoumplis, D.; Martin, D. E.; Orenstein, J. M. et al. PA-457: A potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *P Natl Acad Sci USA* **2003**, *100* (23), 13555.
- (2) Zhou, J.; Yuan, X.; Dismuke, D.; Forshey, B. M.; Lundquist, C.; Lee, K. H.; Aiken, C.; Chen, C. H. Small-Molecule Inhibition of Human Immunodeficiency Virus Type 1 Replication by Specific Targeting of the Final Step of Virion Maturation. *Journal of Virology* **2003**, *78* (2), 922.
- (3) Adamson, C. S.; Ablan, S. D.; Boeras, I.; Goila-Gaur, R.; Soheilian, F.; Nagashima, K.; Li, F.; Salzwedel, K.; Sakalian, M.; Wild, C. T. et al. In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (Bevirimat). *J Virol* **2006**, *80* (22), 10957.
- (4) Nguyen, A. T.; Feasley, C. L.; Jackson, K. W.; Nitz, T. J.; Salzwedel, K.; Air, G. M.; Sakalian, M. The prototype HIV-1 maturation inhibitor, bevirimat, binds to the CA-SP1 cleavage site in immature Gag particles. *Retrovirology* **2011**, *8*, 101.
- (5) Han, Y.; Hou, G. J.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Lipton, A. S.; Burton, S.; Hung, I.; Gor'kov, P. L.; Gan, Z. H. et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2013**, *135* (47), 17793.
- (6) Sun, S. J.; Han, Y.; Paramasivam, S.; Yan, S.; Siglin, A. E.; Williams, J. C.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. Solid-State NMR Spectroscopy of Protein Complexes. *Methods Mol Biol* **2012**, *831*, 303.
- (7) Lu, M. M.; Hou, G. J.; Zhang, H. L.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Perilla, J. R.; Langmead, C. J.; Hung, I.; Gor'kov, P. L. et al. Dynamic allostery governs cyclophilin A-HIV capsid interplay. *P Natl Acad Sci USA* **2015**, *112* (47), 14617.

- (8) Han, Y.; Ahn, J.; Concel, J.; Byeon, I. J. L.; Gronenborn, A. M.; Yang, J.; Polenova, T. Solid-State NMR Studies of HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2010**, *132* (6), 1976.
- (9) Wang, M. Z.; Quinn, C. M.; Perilla, J. R.; Zhang, H. L.; Shirra, R.; Hou, G. J.; Byeon, I. J.; Suiter, C. L.; Ablan, S.; Urano, E. et al. Quenching protein dynamics interferes with HIV capsid maturation. *Nat Commun* **2017**, *8*.
- (10) Hou, G. J.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T. Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J Magn Reson* **2013**, *232*, 18.
- (11) Brauniger, T.; Wormald, P.; Hodgkinson, P. Improved proton decoupling in NMR spectroscopy of crystalline solids using the SPINAL-64 sequence. *Monatsh Chem* **2002**, *133* (12), 1549.
- (12) Hou, G. J.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. H-1-C-13/H-1-N-15 Heteronuclear Dipolar Recoupling by R-Symmetry Sequences Under Fast Magic Angle Spinning for Dynamics Analysis of Biological and Organic Solids. *J Am Chem Soc* **2011**, *133* (46), 18646.
- (13) Hou, G. J.; Lu, X. Y.; Vega, A. J.; Polenova, T. Accurate measurement of heteronuclear dipolar couplings by phase-alternating R-symmetry (PARS) sequences in magic angle spinning NMR spectroscopy. *J Chem Phys* **2014**, *141* (10).
- (14) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995**, *6* (3), 277.
- (15) Goddard, T. D.; Kneller, D. G. University of California, San Francisco, 2004.
- (16) Stevens, T. J.; Fogh, R. H.; Boucher, W.; Higman, V. A.; Eisenmenger, F.; Bardiaux, B.; van Rossum, B. J.; Oschkinat, H.; Laue, E. D. A software framework for analysing solid-state MAS NMR data. *Journal of biomolecular NMR* **2011**, *51* (4), 437.
- (17) Bak, M.; Rasmussen, J. T.; Nielsen, N. C. SIMPSON: A general simulation program for solid-state NMR spectroscopy. *J Magn Reson* **2000**, *147* (2), 296.

- (18) Han, B.; Liu, Y. F.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *J Biomol Nmr* **2011**, *50* (1), 43.
- (19) Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment. *Proteins* **1995**, *23* (4), 566.
- (20) Bertini, I.; Emsley, L.; Felli, I. C.; Laage, S.; Lesage, A.; Lewandowski, J. R.; Marchetti, A.; Pierattelli, R.; Pintacuda, G. High-resolution and sensitivity through-bond correlations in ultra-fast magic angle spinning (MAS) solid-state NMR. *Chem Sci* **2011**, *2* (2), 345.
- (21) Hafsa, N. E.; Wishart, D. S. CSI 2.0: a significantly improved version of the Chemical Shift Index. *J Biomol Nmr* **2014**, *60* (2-3), 131.
- (22) Shen, Y.; Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol Nmr* **2013**, *56* (3), 227.
- (23) Monroe, E. B.; Kang, S.; Kyere, S. K.; Li, R.; Prevelige, P. E. Hydrogen/Deuterium Exchange Analysis of HIV-1 Capsid Assembly and Maturation. *Structure* **2010**, *18* (11), 1483.
- (24) Schur, F. K. M.; Hagen, W. J. H.; Rumlova, M.; Ruml, T.; Muller, B.; Krausslich, H. G.; Briggs, J. A. G. Structure of the immature HIV-1 capsid in intact virus particles at 8.8 angstrom resolution. *Nature* **2015**, *517* (7535), 505.
- (25) von Schwedler, U. K.; Stemmler, T. L.; Klishko, V. Y.; Li, S.; Albertine, K. H.; Davis, D. R.; Sundquist, W. I. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *Embo J* **1998**, *17* (6), 1555.
- (26) Tang, C.; Ndassa, Y.; Summers, M. F. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nat Struct Biol* **2002**, *9* (7), 537.
- (27) Jacques, D. A.; McEwan, W. A.; Hilditch, L.; Price, A. J.; Towers, G. J.; James, L. C. HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **2016**, *536* (7616), 349.

- (28) Hafsa, N. E.; Wishart, D. S. CSI 2.0: a significantly improved version of the Chemical Shift Index. *Journal of biomolecular NMR* **2014**, *60* (2-3), 131.
- (29) Waki, K.; Durell, S. R.; Soheilian, F.; Nagashima, K.; Butler, S. L.; Freed, E. O. Structural and Functional Insights into the HIV-1 Maturation Inhibitor Binding Pocket. *Plos Pathog* **2012**, *8* (11).
- (30) Bayro, M. J.; Tycko, R. Structure of the Dimerization Interface in the Mature HIV-1 Capsid Protein Lattice from Solid State NMR of Tubular Assemblies. *J Am Chem Soc* **2016**, *138* (27), 8538.
- (31) Pan, A. C.; Weinreich, T. M.; Piana, S.; Shaw, D. E. Demonstrating an Order-of-Magnitude Sampling Enhancement in Molecular Dynamics Simulations of Complex Protein Systems. *J Chem Theory Comput* **2016**, *12* (3), 1360.
- (32) Goh, B. C.; Perilla, J. R.; England, M. R.; Heyrana, K. J.; Craven, R. C.; Schulten, K. Atomic Modeling of an Immature Retroviral Lattice Using Molecular Dynamics and Mutagenesis. *Structure* **2015**, *23* (8), 1414.
- (33) Han, B.; Liu, Y. F.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *J Biomol NMR* **2011**, *50* (1), 43.
- (34) Seidel, K.; Etkorn, M.; Schneider, R.; Ader, C.; Baldus, M. Comparative analysis of NMR chemical shift predictions for proteins in the solid phase. *Solid State Nucl Magn Reson* **2009**, *35* (4), 235.
- (35) Gupta, R.; Lu, M.; Hou, G.; Caporini, M. A.; Rosay, M.; Maas, W.; Struppe, J.; Suiter, C.; Ahn, J.; Byeon, I. J. et al. Dynamic Nuclear Polarization Enhanced MAS NMR Spectroscopy for Structural Analysis of HIV-1 Protein Assemblies. *J Phys Chem B* **2016**, *120* (2), 329.
- (36) Wagner, J. M.; Zadrozny, K. K.; Chrustowicz, J.; Purdy, M. D.; Yeager, M.; Ganser-Pornillos, B. K.; Pornillos, O. Crystal structure of an HIV assembly and maturation switch. *Elife* **2016**, *5*.

## Chapter 5

### DEVELOPMENTS AND APPLICATIONS OF $^{19}\text{F}$ MAS NMR

#### 5.1 Introduction

Fluorine is a magnetically active nucleus for NMR spectroscopy because  $^{19}\text{F}$  isotope has a nuclear spin of 1/2 with high gyromagnetic ratio (equivalent to 83% of  $^1\text{H}$ ), 100% natural abundance, and performs a large chemical shift dispersion (approx. 500 ppm).<sup>1-3</sup> Fluorine is absent from naturally occurring biological molecules, and can be biosynthetically incorporated into proteins.<sup>4</sup>  $^{19}\text{F}$  NMR has been applied to a large number of biological and biochemical studies both in solution and solid state, such as small proteins, lipids, nucleic acids, and synthetic small-molecule ligands.<sup>5-9</sup> The strong  $^{19}\text{F}$ - $^{19}\text{F}$  dipolar couplings make fluorine well suited as a long-range distance probe in the solid state, distances between fluorine atoms up to 20 Å can be detected.<sup>10,11</sup> However,  $^{19}\text{F}$  MAS NMR of biological systems remains underutilized due to the challenges associated with the lack of comprehensive insights into the relationship between  $^{19}\text{F}$  chemical shifts and local environments, as well as the generic broad lines observed in solid-state NMR spectra. It is imperative to develop applicable experimental conditions for studying biological systems by  $^{19}\text{F}$  MAS NMR.

In this chapter, I describe the work that I contributed to for the development of a robust protocol for  $^{19}\text{F}$  MAS NMR spectroscopy based characterization of structure and dynamics in fluorinated solids. I examined fluorosubstituted tryptophans and explored the application of the method to fluorinated HIV-1 capsid protein assemblies. This work was conducted in our lab by Dr. Manman Lu, Matthew Fritz, Jodi Kraus

and Sucharita Sarkar, and by our collaborators, Dr. In-Ja Byeon, Chang-Hyeock Byeon and Professor Angela Gronenborn from the University of Pittsburgh, and Dr. Jochem Struppe from Bruker Biospin. My contributions included preparation of the fluorinated HIV-1 capsid protein samples for solution and MAS NMR experiments, performing the MAS NMR experiments on fluorosubstituted tryptophans and fluorinated HIV-1 capsid protein sample, and data analysis. The study on fluorosubstituted tryptophans has been published in the recent manuscript: Lu, M.; Sarkar, S.; Wang, M.; Kraus, J.; Fritz, M.; Quinn, C. M.; Bai, S.; Holmes, S. T.; Dybowski, C.; Yap, G. P. A. et al.  $^{19}\text{F}$  Magic Angle Spinning NMR Spectroscopy and Density Functional Theory Calculations of Fluorosubstituted Tryptophans: Integrating Experiment and Theory for Accurate Determination of Chemical Shift Tensors. *The Journal of Physical Chemistry B* **2018**, DOI:10.1021/acs.jpcc.8b00377.1021/acs.jpcc.8b00377. The study on HIV-1 capsid protein assemblies will be published (manuscript in preparation).

## 5.2 $^{19}\text{F}$ MAS NMR of Fluorosubstituted Tryptophans

$^{19}\text{F}$  chemical shifts are sensitive to the structure and electronic environment in organic and biological molecules. To assess the dependence of the experimental  $^{19}\text{F}$  chemical shift tensor (CST) parameters on the fluorine position in the aromatic ring and the chirality of the tryptophan molecule, we inspected CST parameters of six crystalline  $^{19}\text{F}$ -tryptophan samples (4F-, 5F-, 6F-, 7F-DL-Trp and 5F-, 6F-L-Trp, Fig 5.1) by MAS NMR spectroscopy.

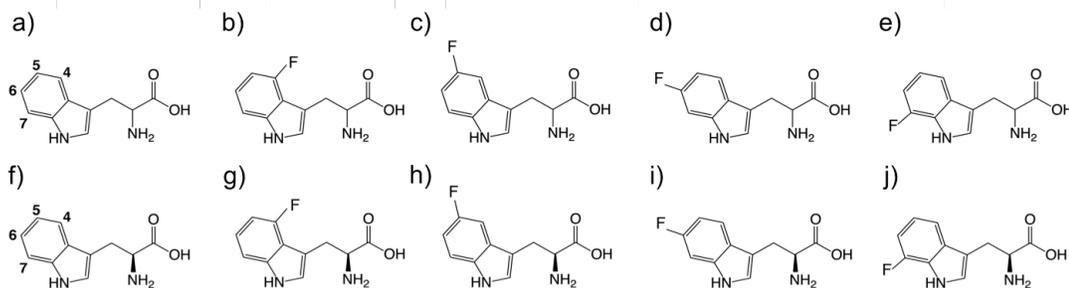


Figure 5.1 Chemical structures of tryptophan and various fluorotryptophan molecules (a) DL-Trp, (b) 4F-DL-Trp, (c) 5F-DL-Trp, (d) 6F-DL-Trp, (e) 7F-DL-Trp, (f) L-Trp, (g) 4F-L-Trp, (h) 5F-L-Trp, (i) 6F-L-Trp, and (j) 7F-L-Trp.

## 5.2.1 Methods and Experiments

### 5.2.1.1 Materials

4F-DL-tryptophan, 5F-L-tryptophan, 6F-L-tryptophan were purchased from Sigma-Aldrich. 5F-DL-tryptophan was purchased from MP Biomedicals, LLC (Santa Ana, CA). 6F-DL-tryptophan was purchased from Apollo Scientific Ltd (UK). 7F-DL-tryptophan was purchased from Amatec Chemical Co. Ltd (Hong Kong). All fluorosubstituted tryptophans were used without further purification.

### 5.2.1.2 $^{19}\text{F}$ MAS NMR Experiments

$^{19}\text{F}$  MAS NMR experiments on 4F-DL-tryptophan, 5F-DL-tryptophan, 5F-L-tryptophan, 6F-DL-tryptophan, 6F-L-tryptophan, and 7F-DL-tryptophan were carried out at room temperature (21 °C) on a 9.4 T NMR spectrometer comprised of a wide-bore Magnex Scientific magnet and a Bruker AVANCE console. The spectrometer was equipped with a 3.2 mm Varian T3 MAS HXY probe. The  $^{19}\text{F}$  Larmor frequency was 376.476 MHz. 20 mg of 4F-DL-tryptophan, 20 mg of 5F-DL-tryptophan, 10 mg

of 5F-L-tryptophan, 20 mg of 6F-DL-tryptophan, 20 mg of 6F-L-tryptophan, and 4 mg of 7F-DL-tryptophan were packed into 3.2 mm Varian rotors.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 5 kHz, 7 kHz, 9 kHz, and 11 kHz; MAS frequencies were controlled to within  $\pm 10$  Hz by a Tecmag MAS controller. The typical  $^{19}\text{F}$   $90^\circ$  pulse length was 1.4  $\mu\text{s}$ . During the odd-numbered transients, a  $180^\circ$  pulse was applied before the  $90^\circ$  excitation pulse to suppress  $^{19}\text{F}$  background signals. The recycle delay was 10 s.

$^{19}\text{F}$  MAS NMR experiments on 5F-DL-tryptophan, 5F-L-tryptophan, 6F-DL-tryptophan, and 6F-L-tryptophan were also carried out on a Varian 14.1 T narrow bore magnet equipped with a Bruker AVIII HD spectrometer, and outfitted with a 1.3 mm Bruker HCN MAS probe. The  $^{19}\text{F}$  Larmor frequency was 564.278 MHz. 3.0 mg of 5F-DL-tryptophan, 3.0 mg of 5F-L-tryptophan, 1.8 mg of 6F-DL-tryptophan, and 2.2 mg of 6F-L-tryptophan were packed into 1.3 mm Bruker rotors.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 8 kHz, 11 kHz (except for 5F-DL-tryptophan), 20 kHz, 30 kHz, 35 kHz, 40 kHz, 50 kHz, and 60 kHz; the frequencies were controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The actual sample temperature was calibrated using KBr as a temperature sensor and was maintained at  $22 \pm 1$   $^\circ\text{C}$  throughout the experiments using a Bruker temperature controller. The typical  $^{19}\text{F}$   $90^\circ$  pulse length was 1.3  $\mu\text{s}$ . The recycle delay was 3 s.

The effects of various magnetic field strengths and  $^1\text{H}$  decoupling were assessed using  $^{19}\text{F}$  MAS NMR experiments with 5F-DL-tryptophan. MAS NMR experiments were carried out on an 11.7 T wide-bore Bruker AVANCE III spectrometer outfitted with a 2.5 mm MAS HFX probe. Larmor frequencies were 500.1 MHz ( $^1\text{H}$ ) and 470.6 MHz ( $^{19}\text{F}$ ). 8 mg of 5F-DL-tryptophan were packed into a

2.5 mm thick-walled Bruker rotor.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 5 kHz, 7 kHz, 11.111 kHz, and 30 kHz, with the frequencies controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The actual sample temperature was calibrated using KBr as a temperature sensor and was maintained at  $20\pm 1^\circ\text{C}$  throughout the experiments using a Bruker temperature controller. The typical  $90^\circ$  pulse lengths were  $2.5\ \mu\text{s}$  ( $^1\text{H}$ ), and  $2.5\ \mu\text{s}$  ( $^{19}\text{F}$ ). Single pulse  $^{19}\text{F}$  spectra were acquired by using a small-angle excitation pulse with a pulse length of  $0.5\ \mu\text{s}$ .  $^1\text{H}$  TPPM decoupling<sup>12</sup> (91 kHz) was used during the acquisition period. The recycle delay was 3 s.

MAS NMR experiments were also carried out on a 19.96 T Bruker AVANCE III spectrometer, outfitted with a 1.9 mm MAS HX probe. 9 mg of 5F-DL-tryptophan were packed into a 1.9 mm thin-wall Bruker rotor. The  $^{19}\text{F}$  Larmor frequency was 800.095 MHz.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 11 kHz, 15 kHz, 30 kHz, and 40 kHz, and the frequencies were controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The actual sample temperature was calibrated using KBr as a temperature sensor and was maintained at  $25\pm 1^\circ\text{C}$  throughout the experiments using a Bruker temperature controller. The typical  $90^\circ$  pulse length for  $^{19}\text{F}$  was  $1.8\ \mu\text{s}$ . During the odd-numbered transients, a  $180^\circ$  pulse was applied before the  $90^\circ$  excitation pulse to suppress the  $^{19}\text{F}$  background signals. The recycle delay was 3 s.

MAS NMR experiments were also carried out on a 19.96 T Bruker AVANCE III spectrometer outfitted with an H-F dual mode MAS HFXYP Phoenix probe. The sample was 7.6 mg of 5F-DL-tryptophan packed into a 1.6 mm Varian rotor. Larmor frequencies were 850.4 MHz ( $^1\text{H}$ ) and 800.1 MHz ( $^{19}\text{F}$ ).  $^{19}\text{F}$  MAS NMR spectra were collected at a MAS frequency of 40 kHz, regulated to within  $\pm 5$  Hz by a Bruker MAS

controller. The actual sample temperature was calibrated using KBr as a temperature sensor and was maintained at  $25\pm 1^\circ\text{C}$  throughout the experiments using a Bruker temperature controller. The typical  $90^\circ$  pulse lengths were  $2.7\ \mu\text{s}$  ( $^1\text{H}$ ), and  $2.7\ \mu\text{s}$  ( $^{19}\text{F}$ ).  $^1\text{H}$  TPPM decoupling (10 kHz) was used during the acquisition period. The recycle delay was 80 s.

$^{19}\text{F}$  chemical shifts were REFERENCED to trifluoroacetic acid, TFA (100  $\mu\text{M}$  solution in 25 mM sodium phosphate buffer, pH 6.5), used as an external REFERENCES standard (0 ppm).

All  $^{19}\text{F}$  MAS spectra were processed using TopSpin. Baseline correction was employed using manually defined baseline points.

### **5.2.1.3 Determination of Chemical Shift Tensors**

The principal components of  $^{19}\text{F}$  chemical shift tensors were determined by fitting the spinning sideband intensities according to the Herzfeld-Berger Analysis protocol.<sup>13</sup> The peak intensities were obtained from TopSpin and were input into the HBA program, version 1.7.5.<sup>14</sup>

### **5.2.2 MAS and Magnetic Field Dependence of Sensitivity and Resolution**

$^{19}\text{F}$  MAS NMR spectra of crystalline tryptophan with fluorine substitutions in the 4-, 5-, 6-, and 7-positions on the indole ring were acquired at various magnetic field strengths, including 9.4, 14.1, and 19.96 T, and at MAS frequencies ranging from 5-60 kHz (Fig 5.2, 5.3 and 5.4). Inspecting a series of spectra of 5F, 6F-DL-tryptophan and 5F, 6F-L-tryptophan crystalline that were acquired at fast MAS frequencies, we notice that, in the absence of  $^1\text{H}$  decoupling, increasing the MAS frequency from 20 to 60 kHz results in dramatic line narrowing and increases in peak intensity because of

improved efficiency of the rotational averaging of  $^{19}\text{F}$  CSA and  $^1\text{H}$ - $^{19}\text{F}/^{19}\text{F}$ - $^{19}\text{F}$  dipolar interactions (Fig 5.2 and Table 5.1). Moreover, as shown in Fig 5.2 d-e, the spectral resolution attained for 5F-DL-tryptophan at 14.1 T and a MAS frequency of 60 kHz in the absence of decoupling is similar to that observed in experiments at 19.96 T and a MAS frequency of 40 kHz, in the presence of  $^1\text{H}$  decoupling. This is a very important finding for optimizing experimental conditions, since it demonstrates that  $^{19}\text{F}$  MAS NMR spectroscopy can be widely applied, even in the absence of specialized probes that are capable of  $^1\text{H}$  decoupling while detecting  $^{19}\text{F}$ .

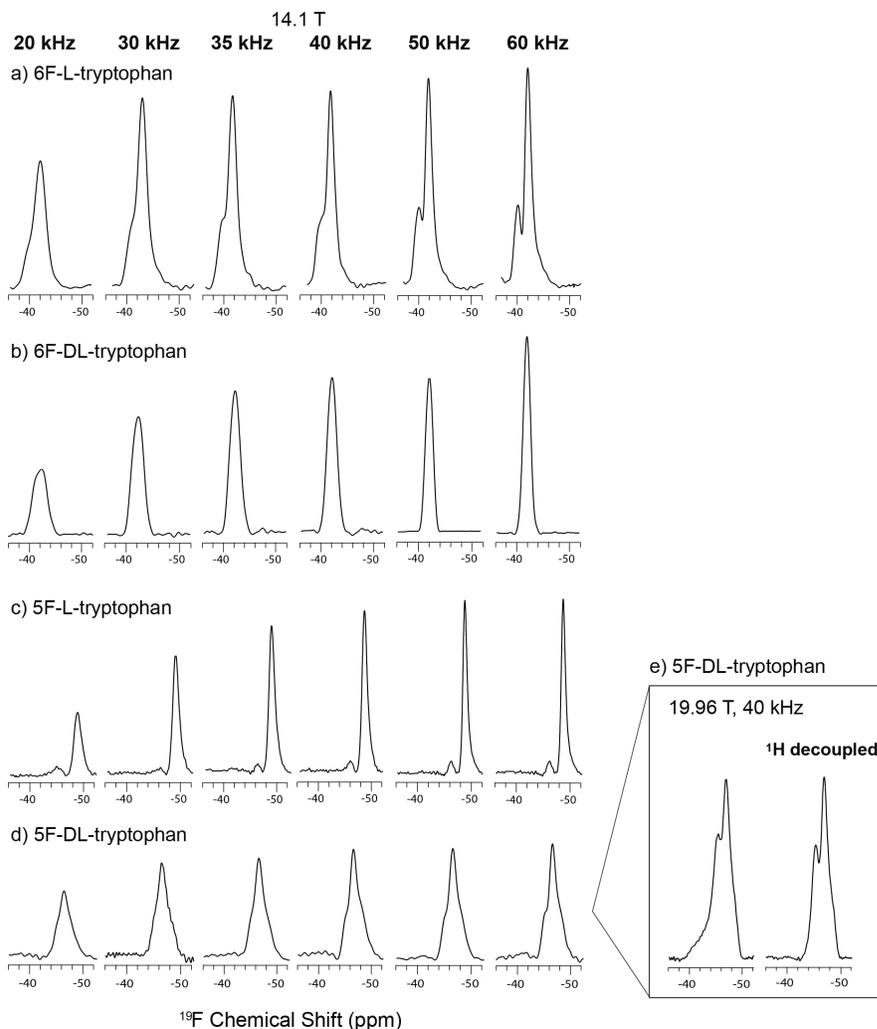


Figure 5.2 (a-d) MAS frequency dependence of the linewidth and signal intensity of fluorosubstituted tryptophan crystalline solids: (a) 6F-L-tryptophan, (b) 6F-DL-tryptophan, (c) 5F-L-tryptophan, (d) 5F-DL-tryptophan. 1D  $^{19}\text{F}$  spectra were acquired at 14.1 T (564.278 MHz  $^{19}\text{F}$  Larmor frequency) and MAS frequencies of 20 kHz, 30 kHz, 35 kHz, 40 kHz, 50 kHz, and 60 kHz. For (e) 5F-DL-tryptophan,  $^{19}\text{F}$  1D spectra were also acquired at 19.96 T (800.095 MHz  $^{19}\text{F}$  Larmor frequency) and a MAS frequency of 40 kHz, with (right spectrum) and without  $^1\text{H}$  decoupling (left spectrum). The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

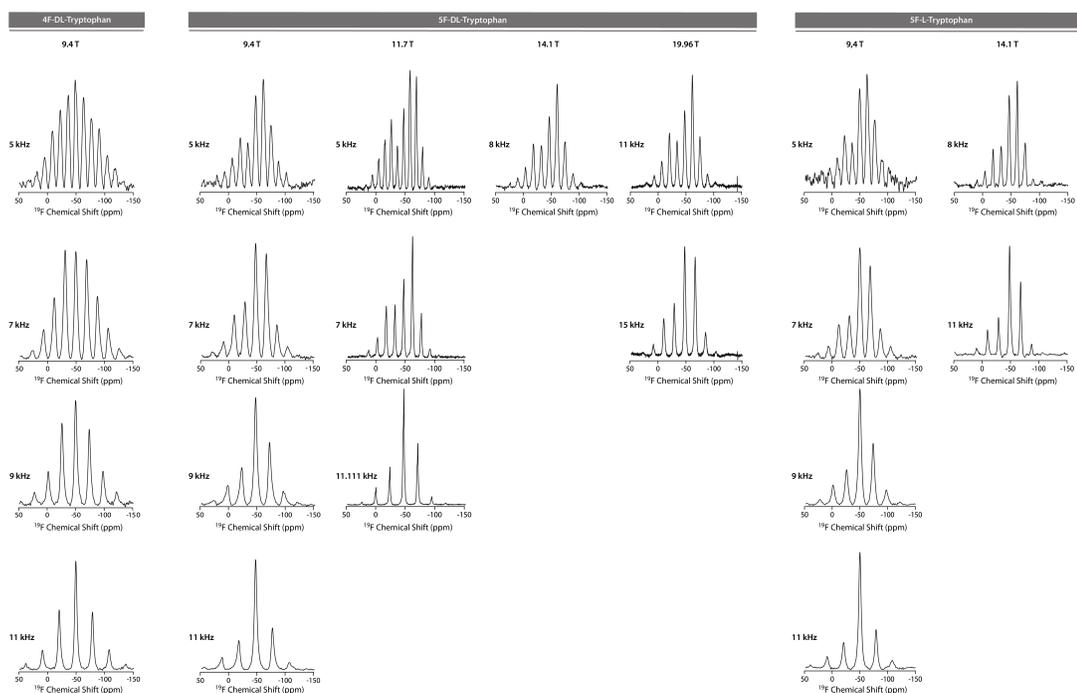


Figure 5.3 Summary of  $^{19}\text{F}$  MAS NMR spectra of fluorosubstituted tryptophans (4F-DL-tryptophan, 5F-DL-tryptophan, 5F-L-tryptophan) acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

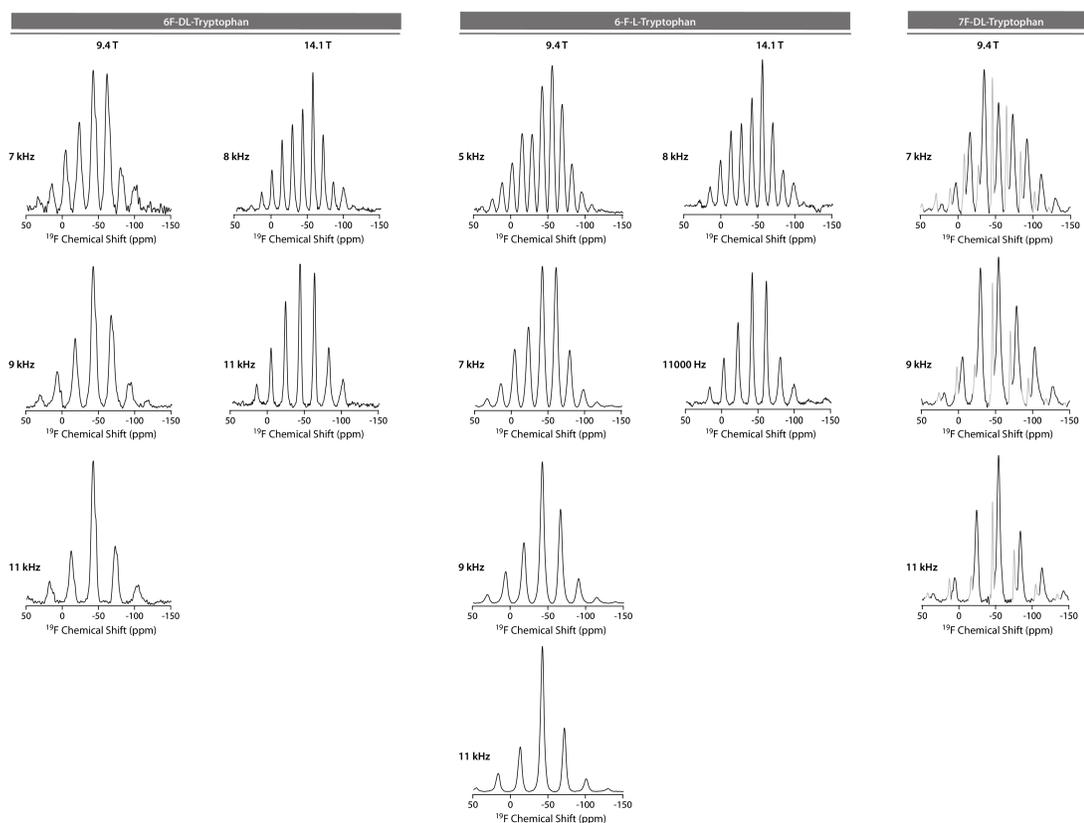


Figure 5.4 Summary of  $^{19}\text{F}$  MAS NMR spectra of fluorosubstituted tryptophans (6F-DL-tryptophan, 6F-L-tryptophan, 7F-DL-tryptophan) acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

Table 5.1 MAS frequency and magnetic field dependence of the sensitivity of  $^{19}\text{F}$  MAS NMR experiments for the 5F-L/DL-tryptophan and 6F-L/DL-tryptophan. The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

Sample	Rotor size	Sample weight	Magnetic field strength	MAS frequency	Signal-to-noise ratio (normalized, per single scan)
6F-L-tryptophan	1.3 mm	2.2 mg	14.1 T	20 kHz	1.4
				30 kHz	1.9
				35 kHz	2.0
				40 kHz	2.0
				50 kHz	2.2
				60 kHz	2.2
6F-DL-tryptophan	1.3 mm	1.8 mg	14.1 T	20 kHz	1.0
				30 kHz	1.7
				35 kHz	2.0
				40 kHz	2.2
				50 kHz	2.2
				60 kHz	3.1
5F-L-tryptophan	1.3 mm	3.0 mg	14.1 T	20 kHz	4.1
				30 kHz	7.9
				35 kHz	10.1
				40 kHz	10.5
				50 kHz	11.7
				60 kHz	11.8
5F-DL-tryptophan	1.3 mm	3.0 mg	14.1 T	20 kHz	2.3
				30 kHz	3.1
				35 kHz	3.4
				40 kHz	3.9
				50 kHz	3.9
				60 kHz	4.0
	1.6 mm	7.6 mg	19.96 T	40 kHz	21.7
				40 kHz, $^1\text{H}$ decoupled	19.1

### 5.2.3 Chemical Shift Tensor Measurements

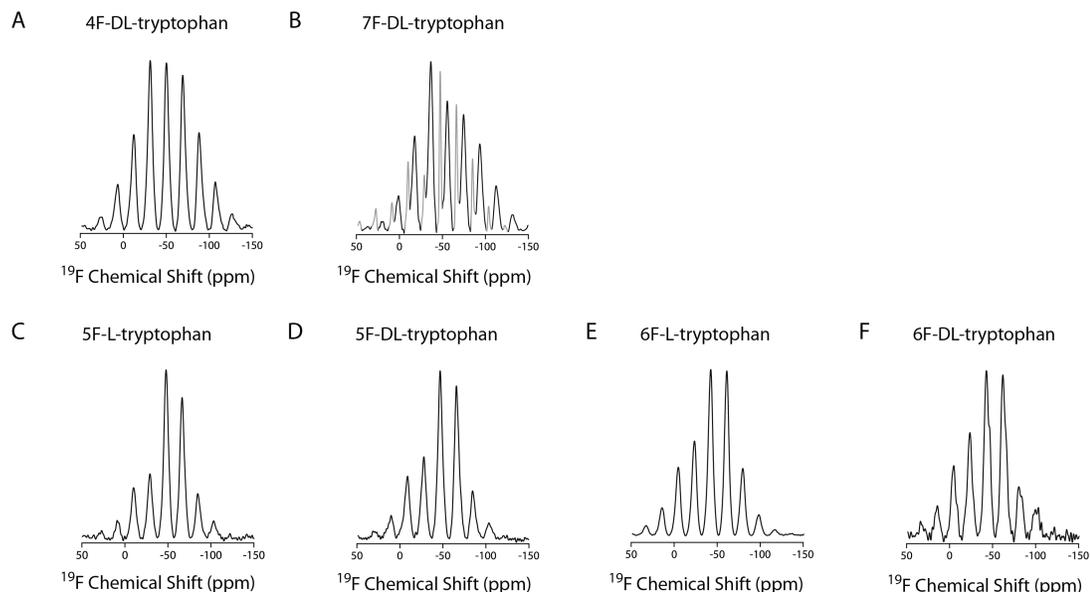
To measure accurate  $^{19}\text{F}$  chemical shift tensors of the fluorosubstituted tryptophans, we acquired  $^{19}\text{F}$  MAS spectra at different magnetic field strengths (9.4, 14.1, and 19.96 T) and at MAS frequencies ranging from 5 to 15 kHz. Representative spectra are provided in Fig. 5.5 (all spectra are shown in Fig 5.3 & 5.4).  $^{19}\text{F}$  CSTs have been previously determined for 5F-L, 5F-DL-, 6F-L- and 6F-DL-tryptophan, and the reported values for the reduced anisotropy parameter ( $\delta_{\parallel}$ ) for 5F-DL-Trp differ by 2.2 ppm for the MAS<sup>16</sup> vs. the single-crystal static<sup>17</sup> NMR study. In our work, we recorded multiple MAS spectra for each compound, which permitted us to extract accurate CST parameters and evaluate the range of experimental errors in these measurements.

The experimental CST parameters were extracted by analyzing the spinning sideband intensities using the Herzfeld-Berger protocol,<sup>18</sup> and all values are summarized in Table 5.2 (all values for each set of spectra are listed in Table 5.3). The isotropic shifts range from -43.2 to -55.7 ppm, with reduced anisotropy parameters between 48.6 and 67.6 ppm. The position of the fluorine in the indole ring yields the largest effect on the shift, and smaller but distinct differences (0.5-3.2 ppm) are observed for molecules of different chirality. All of the tensors are rhombic, with the asymmetry parameters ranging from 0.44-0.5 (for 5F-Trp) to 0.95 (for 4F-Trp). The isotropic shifts are essentially identical, to within 1 ppm, to the solution NMR values for fluorosubstituted L-Trp molecules,<sup>19</sup> except for 7F-L-Trp, for which a difference of 3.1 ppm is observed (the solution NMR shift is -58.8 ppm). Overall, pronounced differences in the isotropic and anisotropic components of CSTs are observed, indicating the remarkably high sensitivity of fluorine chemical shift tensors to local

electronic structure. Thus, fluorotryptophans are excellent  $^{19}\text{F}$  NMR probes in protein studies, as has been previously shown for solution NMR studies.<sup>19,20</sup>

CST parameters for every fluorosubstituted tryptophan and each set of spectra (Table 5.3) allowed us to evaluate the experimental errors carefully. The experimental uncertainties in the  $^{19}\text{F}$  reduced anisotropy parameter ( $\delta_{\parallel}$ ), extracted from the spinning sideband analyses of the MAS spectra, are of the order of 0.8-3.7 ppm, or 5-7% of the total magnitude of  $\delta_{\parallel}$ . In this case, the deviation is reasonable because accurate values rely on the intensities of the outer, low-intensity sidebands, whose quantification becomes more difficult for slower MAS frequencies. Overall, the CSTs for 5F-L/DL and 6F-L/DL tryptophan determined here are in good agreement with the published values,<sup>16,21</sup> within the experimental uncertainty.

9.4 T,  $\omega_r = 7$  kHz



14.1 T,  $\omega_r = 8$  kHz

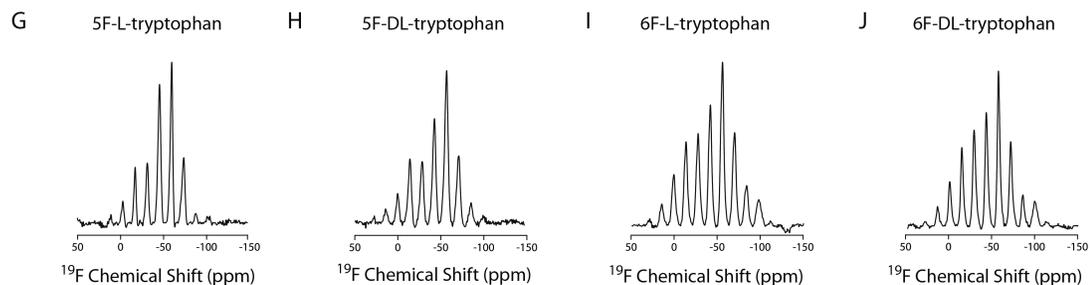


Figure 5.5  $^{19}\text{F}$  MAS NMR spectra of fluorosubstituted tryptophans acquired at 9.4 T (376.476 MHz  $^{19}\text{F}$  Larmor frequency) and a MAS frequency of 7 kHz: (a) 4F-DL-tryptophan, (b) 7F-DL-tryptophan, (c) 5F-L-tryptophan (d) 5F-DL-tryptophan, (e) 6F-L-tryptophan (f) 6F-DL-tryptophan.  $^{19}\text{F}$  MAS NMR spectra of fluorosubstituted tryptophans acquired at 14.1 T ( $^{19}\text{F}$  Larmor frequency of 564.278 MHz) and a MAS frequency of 8 kHz: (g) 5F-L-tryptophan, (h) 5F-DL-tryptophan, (i) 6F-L-tryptophan, (j) 6F-DL-tryptophan. The peaks colored in light grey in (b) arise from the  $^{19}\text{F}$  signal of the sample spacer. The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

Table 5.2 MAS NMR experimental principal components of the chemical shift tensor, isotropic chemical shifts, reduced anisotropy, and asymmetry for the F-DL-tryptophans and F-L-tryptophans. The experimental reduced anisotropy of 4F-DL-tryptophan is -62.8 ppm; for asymmetries close to 1, the sign can be inverted, indicating a symmetric tensor.

Compound	$\delta_{11}$ (ppm)	$\delta_{22}$ (ppm)	$\delta_{33}$ (ppm)	$\delta_{iso}$ (ppm)	$\delta$ (ppm)	$\eta$
4F-DL-tryptophan	11.2±1.4	-48.3±0.5	-112.8±0.8	-50.0	62.8±0.8	0.95±0.02
5F-DL-tryptophan	4.8±1.3	-60.5±1.0	-86.1±1.2	-47.2	51.8±1.3	0.50±0.03
5F-L-tryptophan	-0.8±3.7	-62.9±1.4	-84.4±4.1	-49.4	48.6±3.7	0.44±0.08
6F-DL-tryptophan	12.9±1.4	-51.2±0.5	-91.6±1.4	-43.3	56.2±1.4	0.72±0.02
6F-L-tryptophan	12.6±1.1	-50.7±2.8	-91.3±1.9	-43.2	55.7±1.2	0.73±0.09
7F-DL-tryptophan	4.6±1.9	-48.3±0.5	-123.3±1.4	-55.7	-67.6±1.4	0.78±0.02

Table 5.3 Summary of MAS NMR experimental principal components of the chemical shift tensors, isotropic chemical shifts, reduced anisotropy, and asymmetry for the F-DL-tryptophan acquired at various magnetic field strengths with various MAS frequencies. The figure was originally published in REFERENCES<sup>15</sup>. Permissions for reuse in this dissertation is granted by American Chemical Society, copyright 2018.

Compound	Field Strength	Spinning Frequency	$\delta_{iso}$ (ppm)	$\delta$ (ppm)	$\eta$	$\delta_{11}$ (ppm)	$\delta_{22}$ (ppm)	$\delta_{33}$ (ppm)	
4F-DL-tryptophan	9.4 T	*5 kHz	-50.0	-62.0	0.96	10.8	-48.9	-112.0	
		7 kHz	-50.0	-63.4	0.96	12.1	-48.7	-113.4	
		9 kHz	-50.0	-62.2	0.94	10.2	-48.0	-112.2	
		*11 kHz	-50.0	61.6	0.97	11.6	-50.8	-110.8	
	Average		-50.0	-62.8	0.95	11.2	-48.3	-112.8	
	Standard Deviation		0.0	0.8	0.02	1.4	0.5	0.8	
5F-DL-tryptophan	9.4 T	*5 kHz	-47.2	55.4	0.64	8.2	-57.2	-92.6	
		7 kHz	-47.2	51.7	0.56	4.5	-58.6	-87.4	
		9 kHz	-47.2	54.7	0.48	7.5	-61.3	-87.7	
		*11 kHz	-47.2	55.8	0.57	8.6	-59.2	-91.0	
		Average		-47.2	53.2	0.52	6.0	-60.0	-87.6
		STD		0.0	2.1	0.05	2.1	1.9	0.2
	11.7 T	5 kHz	-47.2	52.0	0.45	4.8	-61.4	-85.0	
		7 kHz	-47.2	51.7	0.52	4.5	-59.7	-86.4	
		11 kHz	-47.2	51.4	0.47	4.1	-60.9	-85.1	
		Average		-47.2	51.7	0.48	4.4	-60.7	-85.5
		STD		0.0	0.3	0.03	0.3	0.9	0.8
	14.1 T	8 kHz	-47.2	51.8	0.53	4.6	-59.4	-86.8	
	19.96 T	11 kHz	-47.2	50.6	0.49	4.8	-61.4	-85.0	
		15 kHz	-47.2	50.5	0.49	4.5	-59.7	-86.4	
		Average		-47.2	50.6	0.49	4.1	-60.9	-85.1
		STD		0.0	0.1	0.00	0.2	1.2	1.0
	Average		-47.2	51.8	0.5	4.8	-60.5	-86.1	
	Standard Deviation		0.0	1.3	0.03	1.3	1.0	1.2	
5F-L-tryptophan	9.4 T	*5 kHz	-48.4	55.3	0.59	5.9	-60.6	-93.4	
		7 kHz	-48.4	50.1	0.49	0.8	-62.2	-86.6	
		9 kHz	-48.4	52.7	0.50	3.4	-62.5	-89.0	
		*11 kHz	-48.4	53.5	0.52	4.1	-62.1	-90.1	
		Average		-48.4	51.4	0.50	2.1	-62.4	-87.8
		STD		0.0	1.8	0.01	1.8	0.2	1.7
	14.1 T	8 kHz	-47.1	44.4	0.43	-5.0	-62.1	-81.1	
		11 kHz	-47.1	47.0	0.34	-2.4	-65.0	-80.8	
		Average		-47.1	45.7	0.38	-3.7	-63.5	-80.9
		STD		0.0	1.9	0.07	1.9	2.1	0.2
Average		-47.7	48.6	0.44	-0.8	-62.9	-84.4		
Standard Deviation		0.8	3.7	0.08	3.7	1.4	4.1		

Compound	Field Strength	Spinning Frequency	$\delta_{iso}$ (ppm)	$\delta$ (ppm)	$\eta$	$\delta_{11}$ (ppm)	$\delta_{22}$ (ppm)	$\delta_{33}$ (ppm)
<b>6F-DL-tryptophan</b>	9.4 T	7 kHz	-43.3	55.2	0.70	11.9	-51.7	-90.1
		9 kHz	-43.3	57.4	0.72	14.2	-51.2	-92.8
		*11 kHz	-43.3	56.0	0.91	12.7	-45.9	-96.8
		Average	-43.3	56.3	0.71	13.1	-51.5	-91.5
		STD	0.0	1.5	0.02	1.6	0.3	1.9
	14.1 T	8 kHz	-43.3	57.4	0.73	14.1	-51.2	-92.8
		11 kHz	-43.3	54.8	0.74	11.5	-50.5	-90.8
		Average	-43.3	56.1	0.73	12.8	-50.9	-91.8
		STD	0.0	1.8	0.01	1.8	0.5	1.4
	Average		-43.3	56.2	0.72	12.9	-51.2	-91.6
Standard Deviation		0.0	1.4	0.02	1.4	0.5	1.4	
<b>F-L-tryptophan</b>	9.4 T	*5 kHz	-43.2	58.9	0.70	15.7	-52.1	-93.2
		7 kHz	-43.2	55.4	0.67	12.2	-52.3	-89.5
		9 kHz	-43.2	57.5	0.64	14.3	-53.5	-90.3
		*11 kHz	-43.2	59.3	0.64	16.1	-53.9	-91.8
		Average	-43.2	56.4	0.66	13.2	-52.9	-89.9
		STD	0.0	1.5	0.02	1.5	0.9	0.6
	14.1 T	8 kHz	-43.2	55.1	0.76	12.1	-49.7	-91.4
		11 kHz	-43.2	55.0	0.85	12.0	-47.2	-93.8
		Average	-42.0	55.0	0.80	12.0	-48.4	-92.6
		STD	0.0	0.1	0.06	0.1	1.8	1.7
Average		-43.2	55.7	0.73	12.6	-50.7	-91.3	
Standard Deviation		0.0	1.2	0.09	1.1	2.8	1.9	
<b>7F-DL-tryptophan</b>	9.4 T	7 kHz	-55.7	-68.7	0.80	5.9	-48.7	-124.4
		9 kHz	-55.7	-66.6	0.77	3.2	-48.0	-122.3
		*11 kHz	-55.7	-68.1	0.77	4.7	-48.0	-123.8
	Average		-55.7	-67.6	0.78	4.6	-48.3	-123.3
	Standard Deviation		0.0	1.4	0.02	1.9	0.5	1.4

\*At 9.4 T, the spectra acquired with 5 kHz spinning exhibit lower sensitivity, and the spectra acquired with 11 kHz spinning have fewer spinning side bands than those obtained with 7 kHz and 9 kHz spinning. Due to the lower accuracy, the values obtained with 5 kHz and 11 kHz spinning at 9.4 T are not included in the calculation of the average values of the CST parameters.

### 5.3 <sup>19</sup>F MAS NMR of HIV-1 CA Protein Assemblies

To explore the potential of <sup>19</sup>F MAS NMR for the structural analysis of protein assemblies, we conducted an investigation of HIV-1 CA protein tubular assemblies. The aromatic rings of tryptophan (Trp) residues were chosen for fluorine incorporation because of its relative sparseness in the CA sequence. HIV-1 CA protein was uniformly labeled with 5F-Trp.

#### 5.3.1 Methods and Experiments

##### 5.3.1.1 Materials

Common chemicals were purchased from Fisher Scientific or Sigma Aldrich. Chromatography columns were purchased from GE Healthcare. <sup>15</sup>NH<sub>4</sub>Cl and U-<sup>13</sup>C<sub>6</sub>-glucose were purchased from Cambridge Laboratories, Inc. The Rosetta (DE3) competent cells used for protein expression were purchased from Novagen.

##### 5.3.1.2 Sample Preparation of Fluorinated CA Protein

5-Fluoroindole (Sigma Aldrich) was used as a precursor to uniformly incorporate fluorine into all Trp residues of the HIV-1 CA protein.<sup>22</sup> 5-<sup>19</sup>F-Trp, U-<sup>13</sup>C, <sup>15</sup>N-labeled CA, 5-<sup>19</sup>F-Trp, U-<sup>15</sup>N-labeled CA, 5-<sup>19</sup>F-Trp, U-<sup>13</sup>C, <sup>15</sup>N-labeled A14C/E45C/W184A/M185A CA, and 5-<sup>19</sup>F-Trp, <sup>15</sup>N-labeled CA Trp-substituted mutant proteins were expressed and purified as reported previously with modifications.<sup>23,24</sup> In brief, 5-<sup>19</sup>F-Trp, U-<sup>13</sup>C, <sup>15</sup>N-enriched CA was expressed in a modified M9 medium, containing <sup>15</sup>NH<sub>4</sub>Cl, U-<sup>13</sup>C<sub>6</sub>-glucose. 20 mg of 5-fluoroindole for 1 L medium was added prior to inducing expression with 0.8 mM IPTG. Cells were grown at 18 °C and harvested after 16 h by centrifugation. Cell pellets were resuspended in 25 mM sodium phosphate buffer (pH 7.0), ruptured by sonication on

ice, and the cell debris was removed by centrifugation at 27,000 g at 4 °C for 1 h. The pH of the supernatant was adjusted to 5.8 with acetic acid, and the conductivity was adjusted to below 2.5 ms/cm by dilution, followed by an additional centrifugation at 27,000 g at 4 °C for 1 h. The final supernatant was loaded onto a cation exchange column (HiTrap SP HP 5 mL, GE Healthcare), and the protein was eluted with a 0-1 M NaCl gradient in 25 mM sodium phosphate buffer (pH 5.8), 1 mM DTT, and 0.02% NaN<sub>3</sub>. Fractions containing CA protein were further purified by gel filtration using a size-exclusion column (HiLoad 26/600 Superdex 75 prep grade, GE Healthcare), equilibrated with 25 mM sodium phosphate buffer (pH 5.5), 1 mM DTT, and 0.02% NaN<sub>3</sub>. Fractions containing CA protein were combined and concentrated to 20-30 mg/mL.

To prepare 5-<sup>19</sup>F-Trp,U-<sup>13</sup>C,<sup>15</sup>N-labeled CA A14C/E45C/W184A/M185A cross-linked hexamers, monomeric protein was buffer exchanged into 25 mM sodium phosphate buffer (pH 8.0) without DTT or other redox reagents overnight and purified by gel filtration on a size-exclusion column (HiLoad 16/600 Superdex 200 prep grade, GE Healthcare). Fractions containing cross-linked hexamer protein were collected and concentrated to 28 mg/mL, buffer exchanged into 50 mM Tris, 1.0 M NaCl (pH 8.0) for 2.5 days, for assembly into tubes. To prepare 5-<sup>19</sup>F-Trp CA cross-linked hexamer polyethylene glycol (PEG) precipitate, 30% PEG 4000 solution was added to an equal volume of the solution of 28 mg/mL CA cross-linked hexamer, followed by incubation at 37 °C for 1 h and then 4 °C overnight. By dialyzing the tubular assemblies (1M NaCl) against 25 mM sodium phosphate (pH 6.5), 7.5 mg (300 mM) soluble clean hexamer protein solution, containing 7 % D<sub>2</sub>O were prepared for solution NMR experiments.

For solution NMR experiments, the concentrated 5-<sup>19</sup>F-Trp,<sup>15</sup>N-labeled CA wild type (WT) and Trp-substituted mutant proteins were buffer exchanged to 25 mM sodium phosphate (pH 6.5) and diluted to 7.5 mg (300 mM) soluble protein, containing 7 % D<sub>2</sub>O. 5-<sup>19</sup>F-Trp,U-<sup>13</sup>C,<sup>15</sup>N-labeled and 5-<sup>19</sup>F-Trp,<sup>15</sup>N-labeled CA wild type (WT), 5-<sup>19</sup>F-Trp,<sup>15</sup>N-labeled CA A204, and 5-<sup>19</sup>F-Trp,<sup>15</sup>N-labeled CA W80Y assemblies were prepared from 20-26 mg/mL protein solutions in 25 mM phosphate buffer, 1.0 M NaCl (pH 6.5), followed by incubation at 37°C for 1 h and 4°C overnight. 5-<sup>19</sup>F-Trp,<sup>15</sup>N-labeled CA W23I assemblies were prepared from 22 mg/mL protein solution in 25 mM phosphate buffer, 2.4 M NaCl (pH 6.5), followed by incubation at 37°C for 1 h and 4°C overnight. The 5-<sup>19</sup>F-Trp CA WT and mutant assemblies were pelleted at 10,000 g and packed into Bruker MAS NMR rotors of appropriate sizes.

### **5.3.1.3 Transmission Electron Microscopy (TEM)**

The morphologies of 5-<sup>19</sup>F-Trp CA WT and mutants assemblies were characterized by transmission electron microscopy (TEM). TEM images were collected at the Delaware Biotechnology Institute (DBI) in the Bioimaging Center, using a Zeiss Libra 120 transmission electron microscope operating at 120 kV. Assemblies were stained with uranyl acetate (0.5-1% w/v), deposited onto 400 mesh, formvar/carbon-coated copper grids, and dried for 45 min in the air. The copper grids were glow discharged prior to staining, so that the tubular assemblies are uniformly spread on the grid surface and adhere to it.

#### 5.3.1.4 MAS NMR Experiments

$^{19}\text{F}$  MAS NMR experiments were performed on a Bruker 20.0 T Bruker AVANCE III spectrometer, outfitted with a 1.9 mm MAS HX probe. The samples consisted of 14.1 mg of 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA (1M NaCl) assemblies, 13.8 mg of 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA W80Y (1 M NaCl) assemblies, 12.1 mg of 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA W23I (2.4 M NaCl) assemblies, 15.9 mg of 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA A204C (1M NaCl) assemblies, 13.1 mg of 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA cross-linked hexamer (1M NaCl) assemblies, and 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA cross-linked hexamer (PEG 4000) precipitates were packed into 1.9 mm thin-wall Bruker rotors. The  $^{19}\text{F}$  Larmor frequency was 800.1 MHz.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 15 kHz, 35 kHz and 40 kHz, and the frequencies were controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The actual sample temperature was at  $15\pm 1$  °C throughout the experiments using a Bruker temperature controller. The typical  $90^\circ$  pulse length for  $^{19}\text{F}$  was 1.9-3.2  $\mu\text{s}$ . The recycle delay was 3 s for all experiments. For 2D  $^{19}\text{F}$ - $^{19}\text{F}$  spin diffusion spectra, the mixing time is 1 s. For 2D  $^{19}\text{F}$ - $^{19}\text{F}$  RFDR spectra, the RFDR mixing time was 8 ms. The  $^{19}\text{F}$ - $^{19}\text{F}$  RFDR buildup curves for the 5- $^{19}\text{F}$ -Trp CA A14C/E45C/W184A/M185A cross-linked hexamer tubes were prepared by using a series of different mixing time, 1.6 ms, 4 ms, 5.6 ms, 20 ms, 51.2 ms, and 100 ms.

To understand the effect of fluorine incorporation and the effect of fluorine linewidths at fast spinning frequencies, hetero-nuclear proton detected HETCOR and  $^{19}\text{F}$  MAS NMR spectra of tubular assemblies of 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA and U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA were acquired at 20.0 T, with a narrow bore Avance III spectrometers outfitted with 1.3 mm E-Free HCN probe, at MAS frequency of 60 kHz, and the frequencies were controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The

actual sample temperature was at  $15 \pm 1$  °C throughout the experiments using a Bruker temperature controller. The operating Larmor frequencies were 850.4 MHz ( $^1\text{H}$ ), 213.8 MHz ( $^{13}\text{C}$ ), 86.2 MHz ( $^{15}\text{N}$ ), and 800.1 MHz ( $^{19}\text{F}$ ). For subsequent NMR experiments, 3.7 mg of 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA (1M NaCl), and 4.4 mg of U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled CA (1M NaCl) tubular assemblies were packed into 1.3 mm Bruker rotors. The typical  $90^\circ$  pulse lengths were 2.1  $\mu\text{s}$  ( $^1\text{H}$ ), 3.1-3.5  $\mu\text{s}$  ( $^{13}\text{C}$ ), 3.25-3.35  $\mu\text{s}$  ( $^{15}\text{N}$ ), and 3.35  $\mu\text{s}$  ( $^{19}\text{F}$ ). For 2D  $^{13}\text{C}$  - $^{13}\text{C}$  RFDR spectra, the RFDR mixing time was 3.2 ms, and  $^1\text{H}$  field strength performed swfTPPM<sup>25</sup> decoupling (10 kHz) during RFDR. For 2D (H)N and (H)C HETCOR, the WALTZ-16 broadband decoupling (10 kHz) was used during the FID acquisition periods.

The effects of various magnetic field strengths and  $^1\text{H}$  decoupling were examined by using  $^{19}\text{F}$  MAS NMR experiments with 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA tubular assemblies carried out on an 11.7 T wide-bore Bruker AVANCE III spectrometer outfitted with a 2.5 mm MAS HFX probe. Larmor frequencies were 500.1 MHz ( $^1\text{H}$ ) and 470.6 MHz ( $^{19}\text{F}$ ). 19.6 mg of 5- $^{19}\text{F}$ -Trp,  $^{15}\text{N}$ -labeled CA (2.4 M NaCl) assemblies were packed into a 2.5 mm thin-wall Bruker rotor.  $^{19}\text{F}$  MAS NMR spectra were collected at MAS frequencies of 4 kHz and 30 kHz, controlled to within  $\pm 5$  Hz by a Bruker MAS controller. The actual sample temperature was calibrated using KBr as a temperature sensor and was maintained at  $15 \pm 1$  °C throughout the experiments using a Bruker temperature controller. The typical  $90^\circ$  pulse lengths were 2.5  $\mu\text{s}$  ( $^1\text{H}$ ), and 3.2  $\mu\text{s}$  ( $^{19}\text{F}$ ). The recycle delay was 3 s for all experiments. At MAS of 30 kHz and 4 kHz, 17kHz and 100 kHz of  $^1\text{H}$  swfTPPM decoupling were used during the acquisition period, respectively.

$^{19}\text{F}$  chemical shifts were indirectly REFERENCED to the adamantane-REFERENCED  $^{13}\text{C}$  chemical shifts.<sup>26</sup> A sample of 5- $^{19}\text{F}$ -DL-Trp powder was used as a secondary REFERENCES standard, -44.6 ppm at 290 K. The  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts were REFERENCED with respect to DSS, adamantane and ammonium chloride used as external referencing standards.

All  $^{19}\text{F}$  MAS spectra were processed using TopSpin. Baseline correction was employed using manually defined baseline points for 1D spectra. Two dimensional  $^{13}\text{C}$ - $^{13}\text{C}$  RFDR and HN/HC HETCOR spectra were processed in TopSpin and with NMRpipe<sup>27</sup>, and analyzed using SPARKY<sup>28</sup>. For 2D data sets,  $30^\circ$  or  $60^\circ$  shifted sine bell apodization followed by a Lorentzian-to-Gaussian transformation was applied in both dimensions. Forward linear prediction to twice the number of the original data points was used in the indirect dimension in some data sets followed by zero filling to twice the total number of points.

The principal components of  $^{19}\text{F}$  chemical shift tensors were determined by the same protocol as described in section 5.2.1.3.

### **5.3.2 Introduction of $^{19}\text{F}$ into HIV-1 CA Protein**

CA protein contains five tryptophan residues, four in the NTD (W23, W80, W117, and W133) and one in the CTD, W184, which is located at the CTD-CTD dimerization interface (Fig 5.5). Replacement of all native tryptophans by 5F-Trp has no measurable effect on the in vitro tubular assemblies, given that the morphology and the dimensions of U- $^{15}\text{N}$ -CA and 5F-Trp,U- $^{15}\text{N}$  CA tubes are indistinguishable (Fig 5.5b). Furthermore, no significant chemical shift perturbations (other than those associated with residues close to the substitution sites) were noted in the 2D  $^{13}\text{C}$ - and  $^1\text{H}$ -detected MAS RFDR, (H)NH HETCOR and (H)CH HETCOR correlation spectra

(Fig 5.6 b-d). Small perturbations (0.4 ppm or less) were observed in the vicinity of the tryptophans (Table 5.4 and Fig 5.6). The only relatively large chemical shift perturbations, 0.8 and 0.5 ppm in  $^{15}\text{N}$ , were detected for S102 and A77, both of which are within 4.6 Å and have direct electrostatic interactions with the F atom of W133. Taken together, the TEM and MAS NMR results confirm that the introduction of fluorine into CA does not interfere with its assembly and does not perturb local structure to any noticeable extent. 5F-Trp residues are thus suitable endogenous probes of structure in both the monomeric and assembled states of CA.

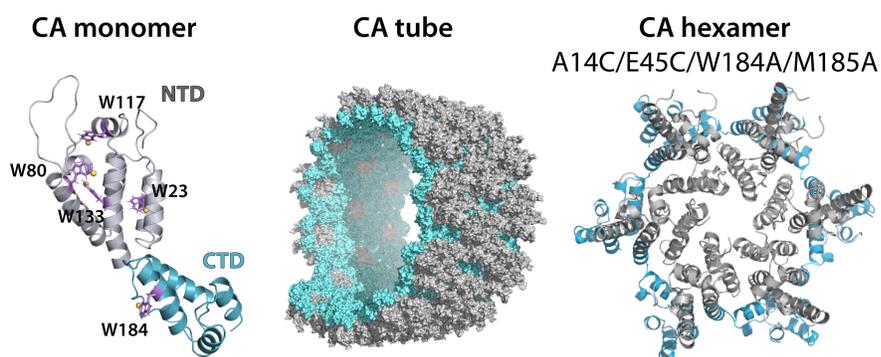


Figure 5.5 Structures of the HIV-1 CA monomer (W23, W80, W117, W133 and W184 are shown in purple stick representation with fluorine atoms as yellow spheres), a section of the CA tube and the A14C/E45C/W184A/M185A CA cross-linked hexamer (NTD; grey, CTD; teal).

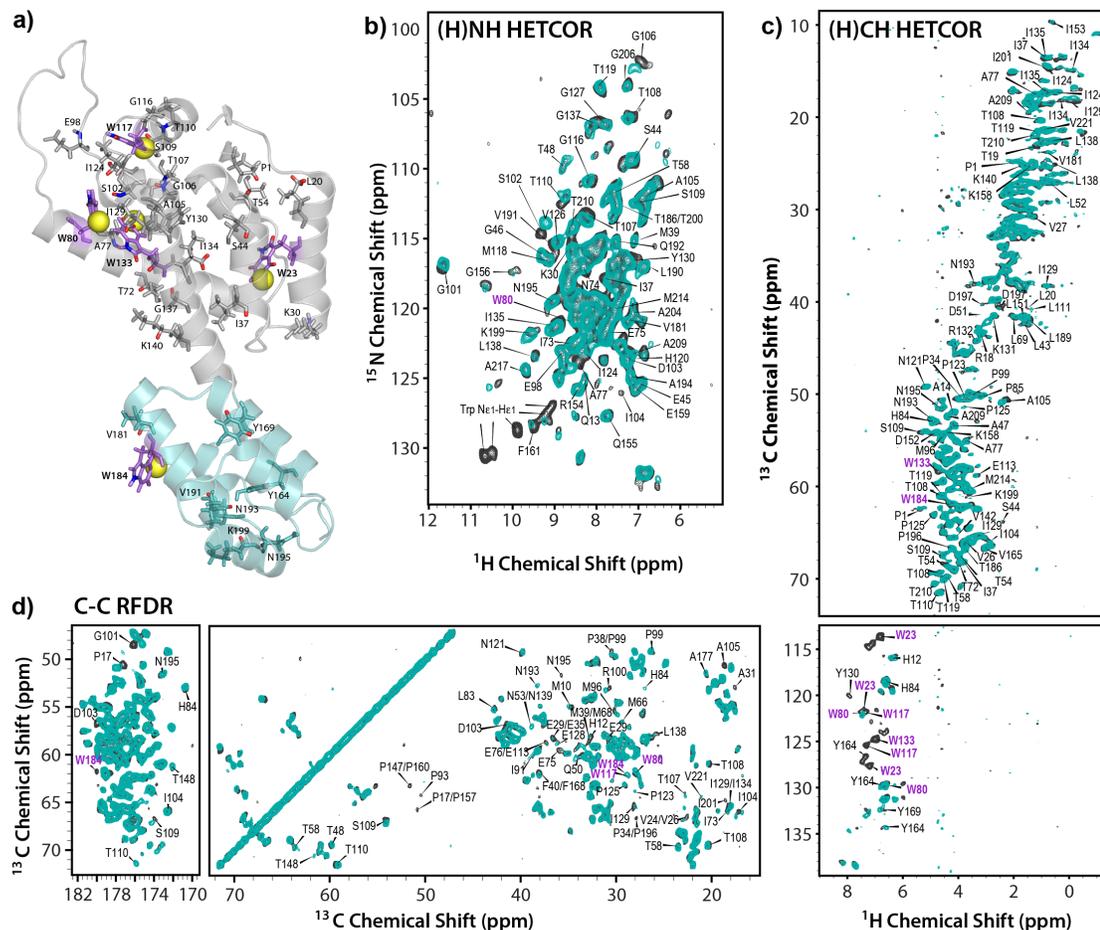


Figure 5.6 (a) 3D structure of an HIV-1 CA monomer showing the positions where 5-F-Trp was incorporated. Five 5- $^{19}\text{F}$ -Trp residues (W23, W80, W117, W133 and W184) are shown in purple sticks, and the fluorine atoms in yellow spheres. (b-d) 2D (H)NH HETCOR (b), (H)CH HETCOR (c) spectra, and  $^{13}\text{C}$ - $^{13}\text{C}$  RFDR spectra in aliphatic region (d) of tubular assemblies of U- $^{13}\text{C}$ ,  $^{15}\text{N}$  CA (grey), and 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$  CA (teal), acquired at 19.96 T; the MAS frequency of 60 kHz. Note the disappearance of resonances corresponding to the indole ring of Trp residues in the fluorinated sample. The residues for which large chemical shift perturbations were detected in the MAS NMR spectra (b-d) are shown as sticks. Nitrogen and hydrogen atoms are shown in blue and red, respectively.

Table 5.4 Chemical shift perturbations detected in (H)NH HETCOR, (H)CH HETCOR, and  $^{13}\text{C}$ - $^{13}\text{C}$  RFDR MAS NMR spectra of tubular assemblies of 5- $^{19}\text{F}$ -Trp, U- $^{13}\text{C}$ ,  $^{15}\text{N}$  CA vs. U- $^{13}\text{C}$ ,  $^{15}\text{N}$  CA.

Experiments	(H)NH / (H)CH HETCOR			$^{13}\text{C}$ - $^{13}\text{C}$ RFDR		
	Cross peak	$ \Delta \text{N/C} $ (ppm)	$ \Delta \text{H} $ (ppm)	Cross peak	$ \Delta \text{C1} $ (ppm)	$ \Delta \text{C2} $ (ppm)
P1	CG-HG2	0.1 (C)	0.14			
M10				CA-CB	0.0	0.2
P17				CD-CO	0.3	0.8
T19	CG2-HG2	0.3 (C)	0.00			
L20	CB-HB3	Missing in fluorinated CA				
W23	CE3-HE3	0.2 (C)	0.03			
	CD1-HD1	Missing in fluorinated CA				
	CZ2-HZ2	Missing in fluorinated CA				
V24/V26				CA-CG	0.3	0.4
E29				CA-CB	0.3	0.2
K30	N-H	0.4 (N)	0.00			
A31				CA-CB	0.3	0.1
P34				CB-CG	0.5	0.7
P34/P196				CA-CG	0.5	0.2
I37	N-H	0.2 (N)	0.09			
	CA-HA	0.1 (C)	0.08			
P38/P99				CD-CB	0.3	0.1
M39/M68				CA-CB	0.4	0.0
F40/F168				CA-CB	0.3	0.0
S44	N-H	0.2 (N)	0.08			
Q50				CA-CB	0.1	0.3
T54	CB-HB	0.2 (C)	0.09			
H62				CB-CO	0.3	0.3
T72	CB-HA	0.1 (C)	0.10			
E75				CA-CB	0.0	0.4
A77	N-H	0.5 (N)	0.16			
W80	CD1-HD1	0.2 (C)	0.09	CA-CB	0.3	0.1
	CZ3-HZ3	Missing in fluorinated CA				
M96				CA-CG	0.3	0.3
E98	N-H	0.3 (N)	0.03			
P38/P99				CD-CB	0.3	0.1
R100				CA-CB	0.2	0.4
G101				CA-CO	0.1	0.9

S102	N-H	0.8 (N)	0.12			
D103				CA-CG	0.1	0.6
				CA-CB	0.1	0.4
I104	CA-HA	0.2 (C)	0.00	CB-CG2	0.2	0.2
				CB-CD	0.1	0.2
A105	N-H	0.3 (N)	0.14	CA-CB	0.6	0.0
				CA-CO	0.3	0.0
G106	N-H	0.4 (N)	0.14			
T107	N-H	0.4 (N)	0.00	CA-CG2	0.3	0.0
T108	N-H	0.2 (N)	0.09			
S109	CA-HA	0.2 (C)	0.15	CB-CO	0.4	0.3
				CB-CA	0.3	0.1
				CA-CO	0.1	0.4
T110	N-H	0.4 (N)	0.04	CB-CA	0.0	0.2
Q112				CA-CB	0.0	0.2
I115				CB-CD	0.6	0.3
G116	N-H	0.1 (N)	0.11			
W117	CE3-HE3	0.2 (C)	0.08			
	CD1-HD1	Missing in fluorinated CA				
M118				CB-CG	0.0	0.2
I124	CG2-HG2	0.1 (C)	0.08	CG1-CG2	0.2	0.1
	CD1-HD1	0.2 (C)	0.03			
I129	CG2-HG2	0.3 (C)	0.08	CG1-CG2	0.0	0.3
I129/I134				CA-CG2	0.1	0.2
Y130	CE1-HE1	Missing in fluorinated CA				
W133	CA-HA	0.5 (C)	0.02			
	CD1-HD1	Missing in fluorinated CA				
I134	CG2-HG2	0.2 (C)	0.10	CG2-CD	0.3	0.0
I129/I134				CA-CG2	0.1	0.2
I135	CD1-HD1	0.3 (C)	0.01			
	CG2-HG2	0.2 (C)	0.04			
G137	N-H	0.2 (N)	0.11			
K140	CG-HG2	0.2 (C)	0.16			
T148				CA-CO	0.3	0.0
I150				CB-CG2	0.2	0.1
I153				CA-CB	0.1	0.5
Y164	CD1-HD1	0.1 (C)	0.10			
	CG-HG	Missing in fluorinated CA				
F40/F168				CA-CB	0.3	0.0
Y169	CG-HG	0.0 (C)	0.10			

V181	CG1-HG1	0.0 (C)	0.10	CG1-CG2	0.1	0.3
V191	N-H	0.0 (N)	0.12			
N193	CB-HB2	0.1 (C)	0.14			
N195	CA-HA	0.1 (C)	0.08	CA-CO	0.1	0.2
				CB-CO	0.1	0.2
P34/P196				CA-CG	0.5	0.2
K199	N-H	0.3 (N)	0.08			
	CA-HA	0.1 (C)	0.15			
I201				CA-CG2	0.3	0.1
G206				CA-CO	0.2	0.4
A209	CB-HB	0.2 (C)	0.03			

The spectra were recorded at 19.96 T ( $^1\text{H}$  Larmor frequency 850 MHz); the MAS frequency was 60 kHz. Perturbations with  $\Delta\text{N} \geq 0.3$  ppm,  $\Delta\text{C} \geq 0.2$  ppm, and  $\Delta\text{H} \geq 0.08$  ppm are considered large, using  $2*\text{STD}$  (standard deviation) as a cutoff.

### 5.3.3 Resonance Assignments of Fluorine Sites

The  $^{19}\text{F}$  solution NMR spectrum of monomeric CA (Fig. 5.7a) exhibits several resonances of varying intensity. To assign the multiple peaks, several tryptophan-substituted CA mutants were generated: W23I, W80Y, W117L, W133Y, W184A/M185A. Some mutants exhibit low stability or low yield, such as W117L and W133Y, and pronounced chemical shift perturbations have been observed in the  $^{19}\text{F}$  solution NMR spectra (Fig. 5.7a). Accordingly, five 5F-Trp sites were assigned and analyzed, where W80, W117 and W133 possess unit intensity resonances at -47.91 ppm, -49.36 ppm and -44.90 ppm, while W23 and W184 exhibit several, smaller intensity resonances at -45.85, -46.3, -46.7 ppm and -48.1, -47.2, -47.8 ppm, respectively. This observation is consistent with prior findings that W184 exists in several conformations in solution, arising from monomeric and dimeric quaternary structures.<sup>29</sup> The  $^{19}\text{F}$  chemical shifts are summarized in Table 5.5.

$5\text{-}^{19}\text{F}$ -Trp CA WT and mutants were assembled and investigated by solid-state  $^{19}\text{F}$  NMR spectra. The following assemblies were prepared: W23I, W80Y, and

A14C/E45C/W184A/M185A. The morphologies of protein assemblies were characterized by transmission electron microscopy (Fig 5.8). 5-<sup>19</sup>F-Trp CA WT and W80Y form tubes; in contrast, W23I forms non-tubular assemblies. 5-<sup>19</sup>F-Trp CA A14C/E45C/W184A/M185A was cross-linked into a hexamer (Fig 5.8d), and assembled into tubes with 1M NaCl. All solid-state <sup>19</sup>F NMR spectra of CA assemblies were shown in Fig 5.7b and Fig 5.9. The general chemical shift range (-44.7 to -49.1 ppm, MAS NMR) for all five 5F-Trp side-chain resonances is similar to that for CA in solution (-44.9 to -48.1 ppm, solution NMR), while noticeable differences exist. Assemblies of CA W23I, W80Y, and A14C/E45C/W184A/M185A samples permitted unambiguous assignments (Fig 5.7b). Interestingly, no noticeable chemical shift changes were observed due to the different conformations. CA A14C/E45C/W184A/M185A was cross-linked into a hexamer, and the spectrum of this hexamer in the tubular assembly was very useful in this regard. In the CA WT assemblies, the 5F-Trp resonances of W133 and W117 are at -44.7 and -49.1 ppm, respectively. W80 exhibits a relatively narrow resonance, superimposed on the broad resonance of W184 at -47.3 ppm. W23 resonates are at -46.3 ppm in the WT assembly and is downfield-shifted to -45.4 ppm in the A14C/E45C/W184A/M185A hexamer mutant. We also notice that there is a minor peak at -46.8 ppm that is associated with W80 in the A14C/E45C/W184A/M185A hexamer mutant. The chemical shifts are summarized in Table 5.6.

The line widths of the W23 and W184 resonances are larger than the others and of the order of 1.3 ppm, suggesting some degree of conformational heterogeneity. A much narrower resonance is observed for W23 at -46.3 ppm for the tubular

assembly of the hexamer, suggesting that a single conformer with a slightly different local environment compared to the WT assembly is present.

It is not surprising to detect the conformational heterogeneity for W184 at the CTD-CTD dimer interface, since different side-chain conformers have been observed by X-ray (PDB 3NTE, 4XFX, 4XFY)<sup>30,31</sup> and NMR<sup>29</sup>. W23 was found to be heterogeneous by solution NMR.<sup>29</sup> The side chain of W23 is buried in the NTD core and is essential for proper core assembly.<sup>32</sup> Interestingly, the W23 side chain was shown to be displaced by more than 3 Å in complexes of CA with potent antiviral capsid inhibitors, derivatives of benzodiazepines and benzoimidazoles,<sup>33</sup> supporting potential side chain conformational variability.

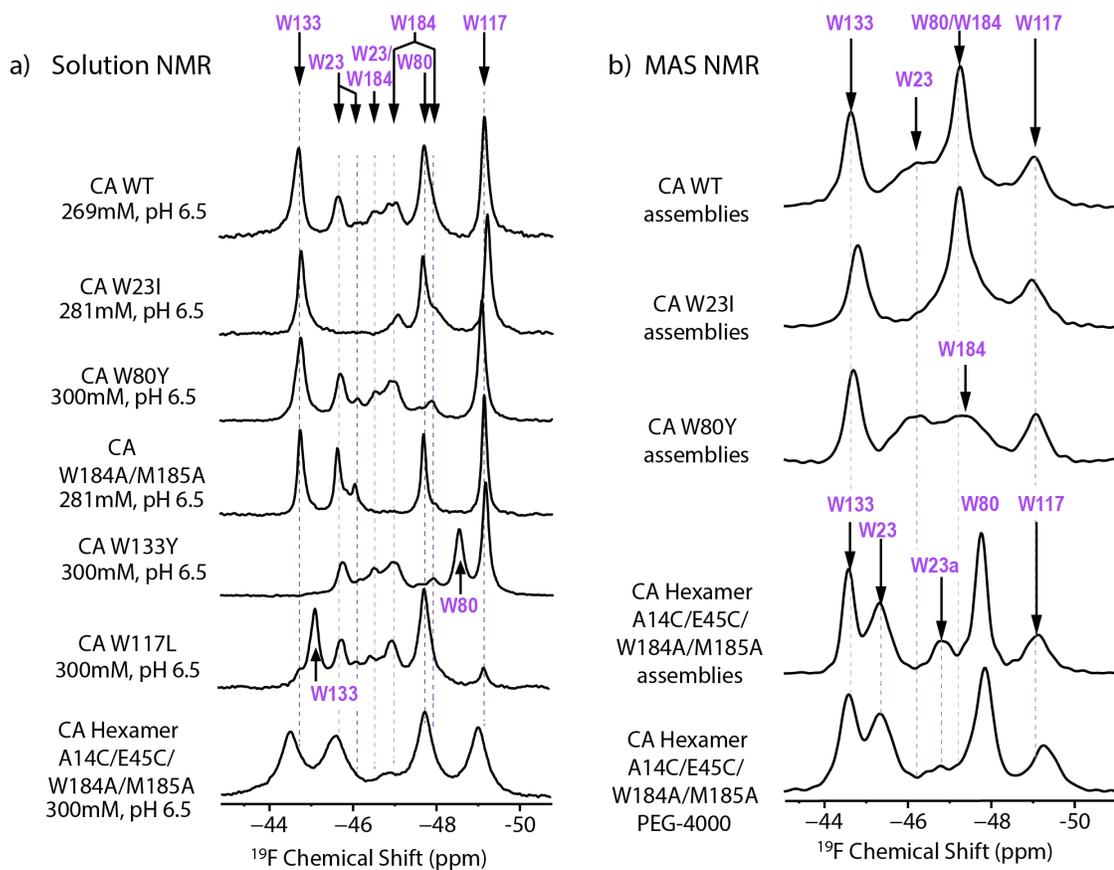


Figure 5.7  $^{19}\text{F}$  solution NMR (a) and MAS NMR (b) spectra of 5- $^{19}\text{F}$ -Trp CA WT, fluorosubstituted CA mutants and CA cross-linked hexamer (A14C/E45C/W184A/M185A). The solution spectra were recorded at 14.1 T. The MAS spectra were recorded at 19.96 T and a MAS frequency of 40 kHz. Assignments for each fluorine site are shown in purple labels.

Table 5.5  $^{19}\text{F}$  chemical shifts of HIV-1 CA protein in solution.

Residue	$^{19}\text{F}$ chemical shift (ppm)*
<b>W23</b>	-45.9, -46.3, -46.7
<b>W80</b>	-47.9
<b>W117</b>	-49.4
<b>W133</b>	-44.9
<b>W184</b>	-48.1, -47.2, -47.8

T = 298 K; 25 mM Na phosphate, pH 6.5, 5 mM DTT, 0.02%  $\text{NaN}_3$ , 7%  $\text{D}_2\text{O}$ .

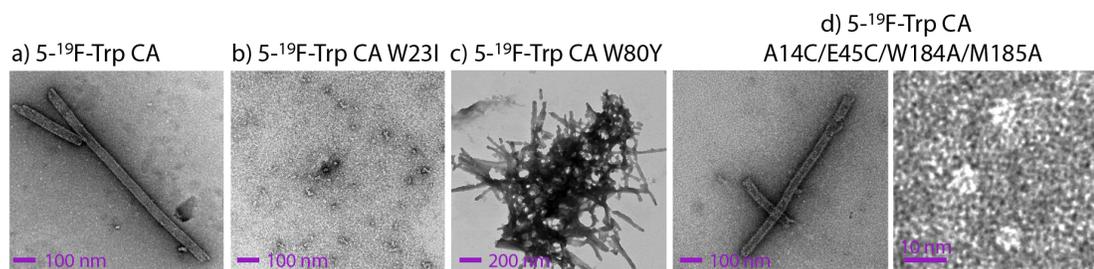


Figure 5.8 TEM images of  $5\text{-}^{19}\text{F}\text{-Trp}$  CA WT and mutant assemblies, including  $5\text{-}^{19}\text{F}\text{-Trp}$ ,  $\text{U}\text{-}^{15}\text{N}$  CA tubes (a),  $5\text{-}^{19}\text{F}\text{-Trp}$ ,  $\text{U}\text{-}^{15}\text{N}$  W80Y tubes (c),  $5\text{-}^{19}\text{F}\text{-Trp}$ ,  $\text{U}\text{-}^{15}\text{N}$  W23I non-tube assemblies (b),  $5\text{-}^{19}\text{F}\text{-Trp}$ ,  $\text{U}\text{-}^{13}\text{C}$ ,  $^{15}\text{N}$  A14C/E45C/W184A/M185A tubes and soluble cross-linked hexamers (d).

Table 5.6 MAS NMR experimental  $^{19}\text{F}$  isotropic chemical shifts for the CA assemblies of WT and 5F-Trp-substituted mutant. All spectra were recorded at 19.96 T.

Sample	Residue	$\delta_{\text{iso}}$ (ppm)
CA tubes	W23	-46.3
	W80/W184	-47.3
	W117	-49.1
	W133	-44.7
CA W23I assemblies	W80/W184	-47.2
	W117	-48.9
	W133	-44.8
CA W80Y tubes	W23	-46.3
	W117	-49.1
	W133	-44.7
	W184	-47.2
CA A14C/E45C/W184A/M185A hexamer tubes	W23	-45.3
	W23a*	-46.8
	W80	-47.7
	W117	-49.1
	W133	-44.6
CA A14C/E45C/W184A/M185A hexamer PEG precipitate	W23	-45.4
	W23a*	-46.5
	W80	-47.8
	W117	-49.2
	W133	-44.6

\* In the samples of cross-linked CA hexamer assembled with NaCl and PEG-4000, Trp 23 is assigned to a major peak (W23) and a minor peak (W23a).

Comparison of the spectra of 5F-Trp CA tubular assemblies at various MAS frequencies of 30, 40, and 60 kHz (Fig 5.9) reveals that proton decoupling is necessary for observing sharp resonances as MAS frequencies below 60 kHz. The  $^{19}\text{F}$  linewidths in the MAS NMR spectrum acquired at  $\omega_r = 30$  kHz with high-power proton decoupling (0.5-1.3 ppm) are still broader than those in the spectrum recorded at  $\omega_r = 60$  kHz without decoupling (0.3-1.0 ppm). The spectra collected at  $\omega_r = 40$  kHz, without decoupling, exhibit linewidths of 1.1 ppm. This observation is in agreement with our findings for fluorosubstituted tryptophan solids<sup>34</sup> and clearly illustrates the benefits of fast MAS conditions.

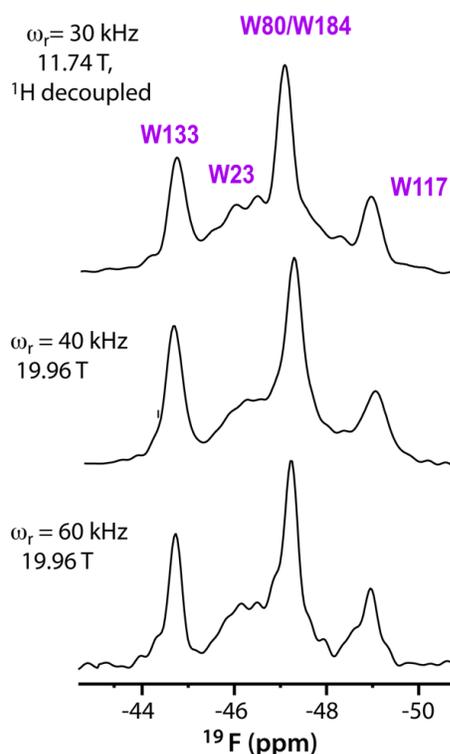


Figure 5.9 Comparison of  $^{19}\text{F}$  MAS NMR spectra of 5- $^{19}\text{F}$ -Trp, U- $^{15}\text{N}$  CA tubular assemblies that were recorded at different magnetic fields and MAS frequencies.

### 5.3.4 Chemical Shift Anisotropy and Relaxation Time Measurements

While isotropic chemical shifts are a key NMR observable, orientation-dependent chemical shift anisotropy (CSA) tensors are an even richer source for exploring local electronic and geometric structure. It also is necessary to assess CSA parameters for interfluorine distance measurements by spin exchange experiments.<sup>11</sup>  $^{19}\text{F}$  CSA tensors for the WT CA tubular assemblies were recorded, using two sets of experiments: (i) at 11.74 T with a MAS frequency of 4 kHz and high-power  $^1\text{H}$  decoupling, and (ii) at 19.96 T with a MAS frequency of 15 kHz without  $^1\text{H}$  decoupling. As shown in Fig 5.10 & 5.11 and summarized in Table 5.7, good

agreement for the  $^{19}\text{F}$  CSA tensor parameters is obtained for slow MAS frequency measurements, with uncertainties of  $\pm 0.6$  ppm and  $\pm 0.3$  for the reduced anisotropy and the asymmetry parameters, respectively. We note that even at slow MAS frequencies, accurate CSA parameters can be obtained, even in the absence of  $^1\text{H}$  decoupling, since no contribution from  $^{19}\text{F}$ - $^{19}\text{F}$  homonuclear dipolar interactions are present. Besides CA WT assemblies, CSA parameters of  $^{19}\text{F}$ -Trp CA mutant assemblies were measured at 19.96 T with a MAS frequency of 15 kHz without  $^1\text{H}$  decoupling.

We notice that CSA parameters for each 5-Trp site in different mutants are variable; the largest difference is 9.3 ppm in the W23 of CA assemblies (44.4 ppm) and that of CA cross-linked hexamer PEG precipitates (53.7 ppm), which possibly imply that high sensitivity of  $^{19}\text{F}$  to different local environments as a structural probe. Interestingly, no correlation between the magnitude of the reduced anisotropy and the isotropic chemical shift is observed.

We recorded  $^{19}\text{F}$   $T_1$  and  $T_2$  relaxation times for each 5-Trp site in CA and CA A14C/E45C/W184A/M185A mutant assemblies (Fig 5.12 and Table 5.8). The 5-Trp sites exhibit very different  $T_1$ s ranging from 1.1 to 9.2 s. The  $T_2$ s are in the range of 0.4 to 1.3 ms. The relaxation times were not found to correlate to the structural differences in these three 5-F-Trp CA assemblies.

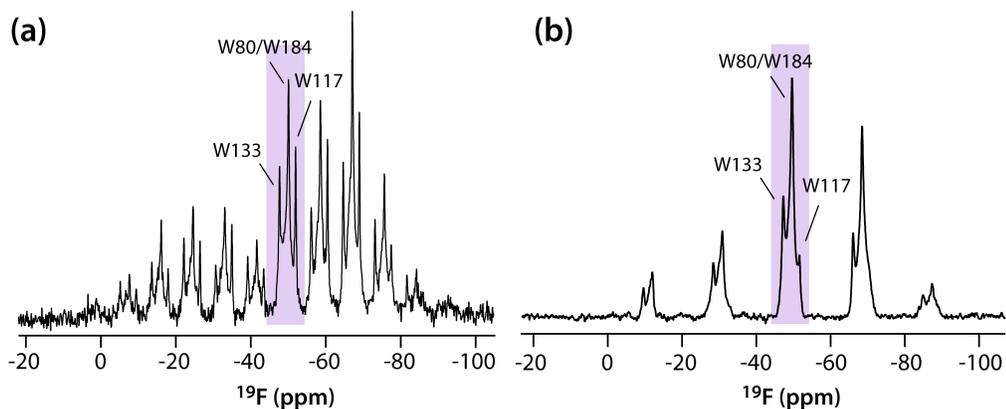


Figure 5.10  $^{19}\text{F}$  MAS NMR spectra of  $5\text{-}^{19}\text{F}\text{-Trp,U-}^{15}\text{N}$  CA tubes (11.74 T; MAS frequency 4 kHz (a), and 19.96 T; MAS frequency 15 kHz (b))

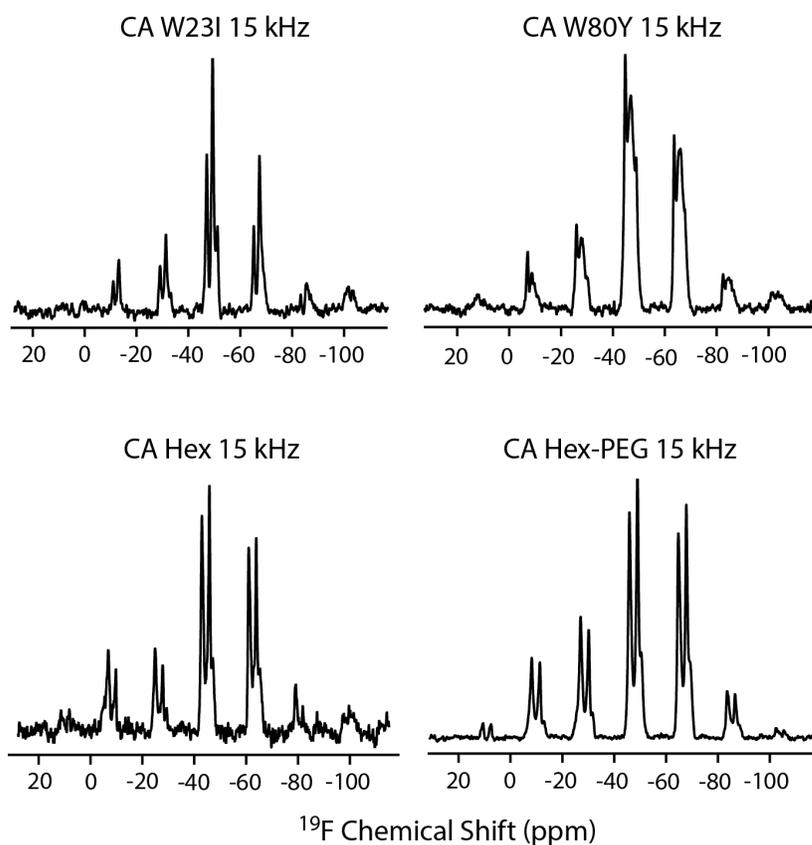


Figure 5.11  $^{19}\text{F}$  MAS NMR spectra of  $5\text{-}^{19}\text{F}\text{-Trp}$  CA mutant assemblies acquired at 19.96 T; the MAS frequency was 15 kHz without  $^1\text{H}$  decoupling.

Table 5.7 Summary of MAS NMR experimental isotropic chemical shifts, reduced anisotropy, and asymmetry for the 5-<sup>19</sup>F-Trp CA wild type and mutants acquired at 11.74 T and 19.96 T.

Sample	Field Strength	5- <sup>19</sup> F-Trp Sites	$\delta_{\text{iso}}$ (ppm)	$\delta_{\sigma}$ (ppm) ( $\pm 1.0$ )	$\eta$
CA assemblies	11.74 T 4 kHz	W23	-46.1	--	--
		W80/W184	-47.3	46.7	0.4
		W117	-49.2	44.1	0.4
		W133	-45.0	44.8	0.7
	19.96 T 15 kHz	W23	-46.3	44.4	0.4
		W80/W184	-47.3	46.1	0.0
		W117	-49.1	44.6	0.0
		W133	-44.7	45.3	0.3
CA hexamer assemblies	19.96 T 15 kHz	W23	-45.3	49.6	0.0
		W23a*	-46.8	--	--
		W80	-47.7	47.5	0.0
		W117	-49.1	45.2	0.6
		W133	-44.6	51.3	0.0
CA W80Y assemblies	19.96 T 15 kHz	W23	-46.3	46.2	0.3
		W117	-49.1	44.2	0.0
		W133	-44.7	42.8	0.4
		W184	-47.2	48.5	0.2
CA W23I assemblies	19.96 T 15 kHz	W80/W184	-47.2	42.4	0.2
		W117	-48.9	39.1	0.0
		W133	-44.8	40.7	0.5
CA hexamer assemblies with PEG	19.96 T 15 kHz	W23	-45.4	53.7	0.3
		W23a*	-46.5	--	--
		W80	-47.8	49.4	0.2
		W117	-49.2	51.0	0.0
		W133	-44.6	49.6	0.6

\* In the samples of cross-linked CA hexamer assembled with NaCl and PEG-4000, Trp 23 is assigned to a major peak (W23) and a minor peak (W23a). "--" Due to the relatively low sensitivity or resolution, some Trp sites cannot be used for reduced anisotropy and asymmetry fitting.

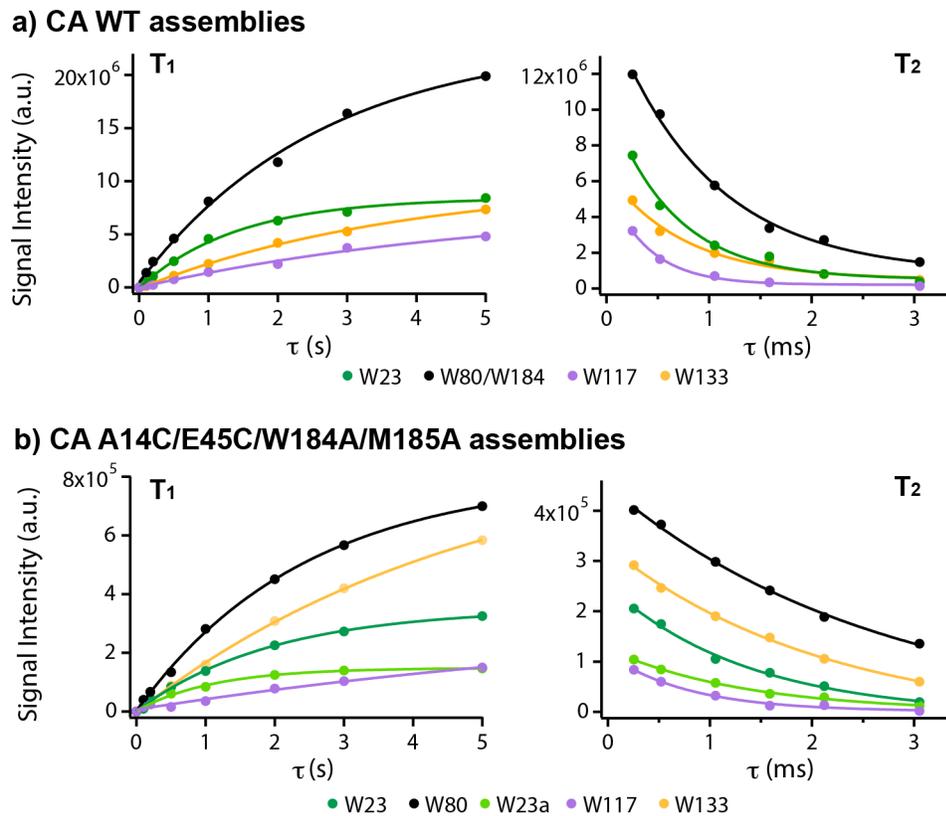


Figure 5.12  $T_1$  buildup curves and  $T_2$  decay curves of  $5\text{-}^{19}\text{F}$ -Trp CA wild type (a) and A14C/E45C/W184A/M185A mutant assemblies (b). Experiments were acquired at 19.96 T and MAS frequency of 40 kHz.

Table 5.8  $T_1$  and  $T_2$  relaxation time of 5- $^{19}\text{F}$ -Trp CA wild type (tubular assembly), and A14C/E45C/W184A/M185A mutant (cross-linked hexamer tubular assembly).

Sample	5- $^{19}\text{F}$ -Trp Sites	$T_1$ (s)	$T_2$ (ms)
CA assemblies	W23	1.4	0.6
	W80/W184	2.6	1.0
	W117	5.2	0.4
	W133	3.8	0.8
CA hexamer assemblies	W23	2.0	0.7
	W23a*	1.1	0.8
	W80	2.4	0.4
	W117	9.2	1.3
	W133	4.9	0.5

\* In cross-linked CA hexamer assemblies, Trp 23 is assigned to a major peak (W23) and a minor peak (W23a).

### 5.3.5 $^{19}\text{F}$ - $^{19}\text{F}$ Correlation Experiments

Intra- and intermolecular  $^{19}\text{F}$ - $^{19}\text{F}$  distances in 5F-Trp labeled CA assemblies are expected to range between 8.8 and 42.5 Å. Based on the X-ray structures of wild type CA and cross-linked hexamers of A14C/E45C/W184A/M185A CA (Table 5.9), four pairs of Trp residues are separated by distances that are within the range potentially accessible for the detection in dipolar-based experiments 5F-Trp interactions: W80-133 (8.8 Å), W117-W133 (9.8 Å), W80-117 (12.6 Å), and W23-W133 (19.1 Å). For fluorine distance measurements by MAS NMR, several methods are available: RFDR<sup>35</sup>, CODEX<sup>36</sup>, and, more recently, PDS<sup>37</sup> and CORD<sup>35</sup>. In a recent study at 14.1 T which employed PDS and CORD mixing at a MAS frequency of 25 kHz, correlations were observed for distances as long as 16 Å for long CORD mixing times of 306 ms.

Here, for CA assemblies, we established that ideally MAS frequencies exceeding 35 kHz are necessary to ensure sufficiently narrow lines in the absence of  $^1\text{H}$  decoupling (Fig 5.9). This is a requirement for probes without separate  $^{19}\text{F}$  and  $^1\text{H}$  channels available for decoupling. We explored MAS frequencies of 35, 40, and 60 kHz, and two mixing schemes, PDS and RFDR, to look for  $^{19}\text{F}$ - $^{19}\text{F}$  correlations in the fluorinated CA assemblies.

PDS spectra of fluorinated CA and A14C/E45C/W184A/M185A CA cross-linked hexamer assemblies, recorded at a MAS frequency of 35 kHz, using a mixing time of 1 s (Fig 5.13a, b) contain cross peaks between W80-W133 resonances and W80-W117 resonances, and W23-W133 resonances. The associated 5F positions are separated by 8.8 Å, 12.6 Å, and 19.1 Å, respectively (PDB 4XFX). In addition, a W23-W80 cross peak is also detected, which corresponds to a F-F distance of 23.1 Å. Since PDS is no longer an efficient mixing scheme for 40 and 60 kHz spinning

frequencies at any practically attainable conditions<sup>38</sup>, we employed RFDR mixing at these higher MAS frequencies. At both spinning frequencies, correlations between W80-W133 and W80-W117 resonances are observed with an RFDR mixing time of 8 ms (Fig 5.13). At 40 kHz, an additional cross peak between W80-W23 is present. At 60 kHz, the W133-W117 correlation appears, while the equivalent W117-W133 cross peak is not seen. PDS spectra of 5F-Trp-substituted CA mutants and A14C/E45C/W184A/M185A CA cross-linked hexamer PEG precipitates were recorded at a MAS frequency of 35 kHz, using a mixing time of 1 s (Figure 5.13c), which further verified the assignments of 5-Trp sites.

<sup>19</sup>F-<sup>19</sup>F RFDR buildup curves were recorded at an MAS frequency of 40 kHz for tubular assemblies of cross-linked A14C/E45C/W184A/M185A hexamer (Fig 5.13d). This assembly was used since W184 is not present and the resonance of W80 is no longer overlapped. It should be noted that, by monitoring the W80-W133 cross peak intensity at a MAS frequency of 40 kHz, the polarization buildup has not reached maximum intensity, even at a mixing time as long as 100 ms. The simulated RFDR buildup curves for the MAS frequencies of 20, 40, and 60 kHz and <sup>19</sup>F-<sup>19</sup>F distances ranging from 5 to 20 Å are shown in Fig 5.13e. We note that quantitative comparisons with the experimental buildup rates are not possible, because the SIMPSON simulations do not include relaxation and multiple protons, and the polarization transfer occurs significantly faster in these experiments. Nevertheless, it is clear that observation of correlations corresponding to 15-20 Å distances requires long mixing times, of the order of hundreds of milliseconds. The future studies will focus on establishing efficient <sup>19</sup>F-<sup>19</sup>F polarization transfer schemes at MAS frequencies of 60

kHz and above, to enable efficient recording of interfluorine distances longer than 15 Å.

Table 5.9  $^{19}\text{F}$ - $^{19}\text{F}$  distances in HIV-1 CA WT and A14C/E45C/W184A/M185A protein.

CA NL4-3, PDB 4XFX (modified with F-Trp, C-F bond is 1.372 Å). \*W184-W184 intermolecular distances are measured in PDB 3J34 HOH building block.

5-F-Trp-CA	W23	W80	W117	W133	W184
	Intramolecular Distance (Å)				
W23	20.0	23.1	23.4	19.1	28.6
W80	33.3	34.1	12.6	8.8	33.1
W117	25.1	21.7	28.5	9.8	42.5
W133	26.2	28.5	27.2	33.0	33.8
W184	33.0	35.8	32.2	27.5	2.1-9.4*
F-F	Intermolecular Distance (Å)				

CA cross-linked hexamer, PDB 3H4E (modified with F-Trp, C-F bond is 1.372 Å)

5-F-Trp-CA A14C/E45C/W184A/M185A	W23	W80	W117	W133
	Intramolecular Distance (Å)			
W23	19.9	22.9	23.8	19.3
W80	33.2	34.3	12.5	8.8
W117	25.1	22.1	29.3	9.9
W133	26.1	28.5	27.9	33.2
F-F	Intermolecular Distance (Å)			



#### 5.4 Conclusions and Future Outlook

Our investigation of  $^{19}\text{F}$  NMR parameters for a series of fluorosubstituted tryptophans (section 5.2) provides a comprehensive evaluation of the influence of the local environment on the  $^{19}\text{F}$  chemical shift tensor. Our results demonstrate that at a MAS frequency of 60 kHz, the  $^{19}\text{F}$  lines are sufficiently narrow, which obviates the need for  $^1\text{H}$  decoupling. The experimental conditions developed here are broadly applicable for studying not only other organic solids, but also large biological molecules and assemblies and is practically accessible in many laboratories.

The results presented in section 5.3, highlight the potential of  $^{19}\text{F}$  MAS NMR spectroscopy to employ 5F-Trp substituted HIV-1 capsid assemblies for analysis of its structural properties and interactions. To our knowledge, this is the first  $^{19}\text{F}$  MAS NMR application on a large protein assembly. The chemical shift dispersion is high with separation of at least three of the five resonances (W80, W117, and W133). An intriguing observation is the fact that for two tryptophans, W23 and W184, multiple resonances are present in both solution and solid-state spectra, indicating conformational heterogeneity.

One important outcome of the current investigation is that at fast MAS conditions (frequencies of 35-60 kHz), the lines are sufficiently narrow that proton decoupling is not required, consistent with our findings in fluorosubstituted tryptophan solids<sup>15</sup>. At these fast MAS conditions, nanometer-range interfluorine distance restraints may be extracted from PDS and RFDR experiments.

In conclusion, our results demonstrate exciting possibilities for widespread use of  $^{19}\text{F}$  fast MAS, including characterization of proteins and protein assemblies without the need to employ specialized probes where rf irradiation can be applied on  $^{19}\text{F}$  and  $^1\text{H}$  channels simultaneously. We envision that  $^{19}\text{F}$  probes that permit MAS frequencies

over 100 kHz, ‘ultrafast’ MAS conditions, will bring further benefits with respect to resolution and will enable  $^{19}\text{F}$  MAS NMR spectroscopy in a wide range of biological assemblies. Moreover, dynamic nuclear polarization (DNP) technique will bring assets with respect to sensitivity, and will allow more novel investigations on large protein assemblies by  $^{19}\text{F}$  MAS NMR spectroscopy.

## REFERENCES

- (1) Gakh, Y. G.; Gakh, A. A.; Gronenborn, A. M. Fluorine as an NMR probe for structural studies of chemical and biological systems. *Magn. Reson. Chem.* **2000**, *38* (7), 551.
- (2) Sharaf, N. G.; Ishima, R.; Gronenborn, A. M. Conformational Plasticity of the NNRTI-Binding Pocket in HIV-1 Reverse Transcriptase: A Fluorine Nuclear Magnetic Resonance Study. *Biochemistry* **2016**, *55* (28), 3864.
- (3) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J Am Chem Soc* **2003**, *125* (25), 7696.
- (4) Buer, B. C.; Marsh, E. N. G. Fluorine: A new element in protein design. *Protein Sci* **2012**, *21* (4), 453.
- (5) Sharaf, N. G.; Gronenborn, A. M. In *Isotope Labeling of Biomolecules - Labeling Methods*; Kelman, Z., Ed.; Elsevier Academic Press Inc: San Diego, 2015; Vol. 565.
- (6) Kitevski-LeBlanc, J. L.; Prosser, R. S. Current Applications of  $^{19}\text{F}$  NMR to Studies of Protein Structure and Dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1.
- (7) Koch, K.; Afonin, S.; Ieronimo, M.; Berditsch, M.; Ulrich, A. S. In *Solid State NMR*; Chan, J. C. C., Ed., 2012; Vol. 306.
- (8) Williams, J. K.; Tietze, D.; Lee, M.; Wang, J.; Hong, M. Solid-State NMR Investigation of the Conformation, Proton Conduction, and Hydration of the Influenza B Virus M2 Transmembrane Proton Channel. *J. Am. Chem. Soc.* **2016**, *138* (26), 8143.

- (9) Hellmich, U. A.; Pflieger, N.; Glaubitz, C.  $^{19}\text{F}$  MAS NMR on Proteorhodopsin: Enhanced Protocol for Site-Specific Labeling for General Application to Membrane Proteins. *Photochem. Photobiol.* **2009**, *85* (2), 535.
- (10) Gilchrist, M. L.; Monde, K.; Tomita, Y.; Iwashita, T.; Nakanishi, K.; McDermott, A. E. Measurement of Interfluorine Distances in Solids. *J. Magn. Reson.* **2001**, *152* (1), 1.
- (11) Roos, M.; Wang, T.; Shcherbakov, A. A.; Hong, M. Fast Magic-Angle-Spinning  $^{19}\text{F}$  Spin Exchange NMR for Determining Nanometer  $^{19}\text{F}$ - $^{19}\text{F}$  Distances in Proteins and Pharmaceutical Compounds. *J Phys Chem B* **2018**, *122* (11), 2900.
- (12) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **1995**, *103* (16), 6951.
- (13) Herzfeld, J.; Berger, A. E. Sideband Intensities in Nmr-Spectra of Samples Spinning at the Magic Angle. *J Chem Phys* **1980**, *73* (12), 6021.
- (14) Eichele, K. Universität Tübingen, 2015.
- (15) Lu, M.; Sarkar, S.; Wang, M.; Kraus, J.; Fritz, M.; Quinn, C. M.; Bai, S.; Holmes, S. T.; Dybowski, C.; Yap, G. P. A. et al.  $^{19}\text{F}$  Magic Angle Spinning NMR Spectroscopy and Density Functional Theory Calculations of Fluorosubstituted Tryptophans: Integrating Experiment and Theory for Accurate Determination of Chemical Shift Tensors. *The Journal of Physical Chemistry B* **2018**, DOI:10.1021/acs.jpcc.8b00377 10.1021/acs.jpcc.8b00377.
- (16) Dürr, U. H.; Grage, S. L.; Witter, R.; Ulrich, A. S. Solid state  $^{19}\text{F}$  NMR parameters of fluorine-labeled amino acids. Part I: aromatic substituents. *J. Magn. Reson.* **2008**, *191* (1), 7.
- (17) Zhao, X.; DeVries, J. S.; McDonald, R.; Sykes, B. D. Determination of the  $^{19}\text{F}$  NMR chemical shielding tensor and crystal structure of 5-fluoro-dl-tryptophan. *J. Magn. Reson.* **2007**, *187* (1), 88.
- (18) Herzfeld, J.; Berger, A. E. Sideband intensities in NMR spectra of samples spinning at the magic angle. *J. Chem. Phys.* **1980**, *73* (12), 6021.

- (19) Matei, E.; Gronenborn, A. M. F-19 Paramagnetic Relaxation Enhancement: A Valuable Tool for Distance Measurements in Proteins. *Angew. Chem.-Int. Edit.* **2016**, *55* (1), 150.
- (20) Sharaf, N. G.; Gronenborn, A. M. In *Isotope Labeling of Biomolecules - Labeling Methods*; Kelman, Z., Ed.; Elsevier Academic Press Inc: San Diego, 2015; Vol. 565.
- (21) Zhao, G. P.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J. Y.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **2013**, *497* (7451), 643.
- (22) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and Inexpensive Incorporation of <sup>19</sup>F-Tryptophan for Protein NMR Spectroscopy. *Chem. Commun.* **2012**, *48* (86), 10681.
- (23) Han, Y.; Hou, G. J.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Lipton, A. S.; Burton, S.; Hung, I.; Gor'kov, P. L.; Gan, Z. H. et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. *J Am Chem Soc* **2013**, *135* (47), 17793.
- (24) Sun, S. J.; Han, Y.; Paramasivam, S.; Yan, S.; Siglin, A. E.; Williams, J. C.; Byeon, I. J. L.; Ahn, J.; Gronenborn, A. M.; Polenova, T. Solid-State NMR Spectroscopy of Protein Complexes. *Methods. Mol. Biol.* **2012**, *831*, 303.
- (25) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. Heteronuclear Decoupling in Rotating Solids. *J Chem Phys* **1995**, *103* (16), 6951.
- (26) Maurer, T.; Kalbitzer, H. R. Indirect Referencing of P-31 and F-19 NMR Spectra. *J. Magn. Reson. Ser. B* **1996**, *113* (2), 177.
- (27) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995**, *6* (3), 277.
- (28) Goddard, T. D.; Kneller, D. G. University of California, San Francisco, 2004.

- (29) Byeon, I. J. L.; Hou, G. J.; Han, Y.; Suiter, C. L.; Ahn, J.; Jung, J.; Byeon, C. H.; Gronenborn, A. M.; Polenova, T. Motions on the Millisecond Time Scale and Multiple Conformations of HIV-1 Capsid Protein: Implications for Structural Polymorphism of CA Assemblies. *J Am Chem Soc* **2012**, *134* (14), 6455.
- (30) Nguyen, A. T.; Feasley, C. L.; Jackson, K. W.; Nitz, T. J.; Salzwedel, K.; Air, G. M.; Sakalian, M. The prototype HIV-1 maturation inhibitor, bevirimat, binds to the CA-SP1 cleavage site in immature Gag particles. *Retrovirology* **2011**, *8*, 101.
- (31) Gres, A. T.; Kirby, K. A.; KewalRamani, V. N.; Tanner, J. J.; Pornillos, O.; Sarafianos, S. G. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* **2015**, *349* (6243), 99.
- (32) Tang, S.; Murakami, T.; Agresta, B. E.; Campbell, S.; Freed, E. O.; Levin, J. G. Human Immunodeficiency Virus Type 1 N-Terminal Capsid Mutants that Exhibit Aberrant Core Morphology and Are Blocked in Initiation of Reverse Transcription in Infected Cells. *J. Virol.* **2001**, *75* (19), 9357.
- (33) Lemke, C. T.; Titolo, S.; von Schwedler, U.; Goudreau, N.; Mercier, J. F.; Wardrop, E.; Faucher, A. M.; Coulombe, R.; Banik, S. S.; Fader, L. et al. Distinct Effects of Two HIV-1 Capsid Assembly Inhibitor Families that Bind the Same Site Within the N-Terminal Domain of the Viral CA Protein. *J. Virol.* **2012**, *86* (12), 6643.
- (34) Fritz, M.; Quinn, C. M.; Wang, M. Z.; Hou, G. J.; Lu, X. Y.; Koharudin, L. M. I.; Struppe, J.; Case, D. A.; Polenova, T.; Gronenborn, A. M. Determination of accurate backbone chemical shift tensors in microcrystalline proteins by integrating MAS NMR and QM/MM. *Phys Chem Chem Phys* **2018**, *20* (14), 9543.
- (35) Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W.; Lansbury, P. T. J.; Griffin, R. G. Homonuclear Radio Frequency-Driven Recoupling in Rotating Solids. *J. Chem. Phys.* **1998**, *108* (22), 9463.
- (36) deAzevedo, E. R.; Hu, W. G.; Bonagamba, T. J.; Schmidt-Rohr, K. Centerband-Only Detection of Exchange: Efficient Analysis of Dynamics in Solids by NMR. *J. Am. Chem. Soc.* **1999**, *121* (36), 8411.

- (37) Bloembergen, N. On the Interaction of Nuclear Spins in a Crystalline Lattice. *Physica* **1949**, *15* (3), 386.
- (38) Hou, G.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T. Broadband Homonuclear Correlation Spectroscopy Driven by Combined  $R2_n^v$  Sequences Under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* **2013**, *232*, 18.

## Chapter 6

### CONCLUSIONS

In this dissertation, I employed MAS solid-state NMR spectroscopy and integrated biophysical methods to characterize the structural rearrangement and dynamic changes accompanying HIV-1 capsid maturation and capsid assembly. Additionally, for method development of  $^{19}\text{F}$  MAS NMR at fast MAS frequencies of 35-60 kHz, I explored  $^{19}\text{F}$  as structural probe to large biological molecules and assemblies, and obtained the longer interfluorine distance (15-20 Å), which benefits the structural determination of capsid assemblies by MAS NMR spectroscopy. The fast and ultrafast MAS frequencies allow  $^{19}\text{F}$  NMR to be broadly applied even without the need of employing specialized  $^{19}\text{F}$ - $^1\text{H}$  dual mode probes.

In chapter 3, I examined two covalently cross-linked CA constructs A14C/E45C/W184A/M185A and N21C/A22C/W184A/M185A, which form the cross-linked hexamer and pentamer, respectively. Using the MAS NMR approach, 89% of the backbone assignments of CA cross-linked hexamer tubular assemblies have been completed. Comparing WT and cross-linked tubular assemblies, I observed that the effect of cross-linking was pronounced on the inter-hexamer interfaces, but not on the protein secondary structure and tertiary structure. The uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  pentamer/natural abundance hexamer co-assembly sample was first prepared to investigate the model pentamer building blocks in the HIV-1 capsid core, exhibiting sphere-like morphology. By analyzing 2D  $^{13}\text{C}$ - $^{13}\text{C}$  correlations of pentamer MAS NMR spectra in the context of co-assembly, the multiple conformational changes

indicated the distinct assembly in hexamer and pentamer. Further investigation will be performed on dynamic properties of cross-linked pentamers as well as cross-linked hexamers to understand the HIV-1 capsid core structure and its stability in different phases.

In chapter 4, I inspected the maturation intermediate, CA-SP1 and its maturation-inhibiting CA-SP1(T8I) mutant using the integrated method of MAS NMR and MD simulations. The data reveal that the CTD tail and SP1 subdomain exists in a highly dynamic helix-coil equilibrium. Such a flexible, dynamic conformation elucidates that SP1 plays the molecular switch in the last step of capsid maturation. T8I mutation in SP1 that phenocopies the MI bound state, introduces dramatic dynamics and structure changes: i) molecular motions being suppressed, and ii) a stable helical structure formed in CTD-SP1 junction, which indicates key mechanistic insights into SP1 cleavage inhibition by maturation inhibitors. In this case, our data suggest that modulation of protein dynamics appears to be a determining factor in capsid maturation, and small-molecule ligands can influence these motions, with profound effects for viral maturation. Overall, this integrated experimental-computational approach has demonstrated the power to facilitate atomistic understanding of the mechanisms in such highly dynamic systems. It offers unique transient structural features that are significantly important to elucidate the details of capsid maturation and the influences of MIs on this step.

In the last chapter, I investigated a series of fluorosubstituted tryptophans by  $^{19}\text{F}$  MAS NMR, which provides a comprehensive evaluation of the influence of the local environment on the  $^{19}\text{F}$  chemical shift tensor. The results demonstrate that at a MAS frequency of 60 kHz, the  $^{19}\text{F}$  lines are sufficiently narrow, which obviates the

need for  $^1\text{H}$  decoupling. Furthermore, we explored  $^{19}\text{F}$  MAS NMR to 5F-Trp substituted HIV-1 capsid assemblies for analysis of its structural properties and interactions. The chemical shift dispersion is high with separation of at least three of the five resonances (W80, W117, and W133). Multiple resonances are present for two tryptophans, W23 and W184 in  $^{19}\text{F}$  spectra, indicating conformational heterogeneity. One important finding is that at fast MAS conditions (frequencies of 35-60 kHz), the lines are sufficiently narrow that proton decoupling is not required. At these fast MAS conditions, the  $^{19}\text{F}$ - $^{19}\text{F}$  correlations that yield interatomic distances as long as 23 Å were observed in the capsid assemblies, which suggests that nanometer-range interfluorine distance restraints may be extracted from PDS and RFDR experiments. Our results demonstrate exciting possibilities for widespread use of  $^{19}\text{F}$  fast MAS, including characterization of proteins and protein assemblies without the need to employ specialized probes where rf irradiation can be applied on  $^{19}\text{F}$  and  $^1\text{H}$  channels simultaneously.

**Appendix A**  
**COPYRIGHT LETTER**

## SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jul 05, 2018

This Agreement between Mingzhang Wang ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4382840643480
License date	Jul 05, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Microbiology
Licensed Content Title	HIV-1 assembly, release and maturation
Licensed Content Author	Eric O. Freed
Licensed Content Date	Jun 29, 2015
Licensed Content Volume	13
Licensed Content Issue	8
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Structural and Dynamic Investigation of HIV-1 Capsid Assemblies and Capsid Maturation by Magic Angle Spinning NMR spectroscopy
Instructor name	Mingzhang Wang
Institution name	University of Delaware
Expected presentation date	Jul 2018
Portions	Figure 5 a and b
Requestor Location	Mingzhang Wang University of Delaware  NEWARK, DE 19716 United States Attn: Mingzhang Wang
Billing Type	Invoice
Billing Address	Mingzhang Wang University of Delaware  NEWARK, DE 19716 United States Attn: Mingzhang Wang
Total	<b>0.00 USD</b>

<https://s100.copyright.com/MyAccount/web/jsp/viewprintablelicensefrommyorders.jsp?ref=bc2b181f-240d-4700-a5fe-f535e9d8dd6f&email=>

1/3

[Terms and Conditions](#)**Springer Nature Terms and Conditions for RightsLink Permissions**

**Springer Customer Service Centre GmbH (the Licensor)** hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.
3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.
4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.
5. Where '**reuse in a dissertation/thesis**' has been selected the following terms apply: Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline ([www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/)).
6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines <http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/>), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.
7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to [Journalpermissions@springernature.com](mailto:Journalpermissions@springernature.com)/[bookpermissions@springernature.com](mailto:bookpermissions@springernature.com) for these rights.
8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.
9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.
10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

**Appendix — Acknowledgements:****For Journal Content:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME]  
[REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

**For Advance Online Publication papers:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME]  
[REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

**For Adaptations/Translations:**

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL

**NAME** [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

**Note: For any republication from the British Journal of Cancer, the following credit line style applies:**

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

For **Advance Online Publication** papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For **Book content**:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc) [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)

**Other Conditions:**

Version 1.0

Questions? [customer-care@copyright.com](mailto:customer-care@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

---

---



RightsLink®

Home

Account Info

Help



AMERICAN SOCIETY FOR MICROBIOLOGY

**Title:** Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation  
**Author:** Juan Fontana, Kellie A. Jurado, Naqian Cheng et al.

Logged in as:  
Mingzhang Wang  
Account #:  
3001099791

LOGOUT

**Publication:** Journal of Virology  
**Publisher:** American Society for Microbiology

**Date:** Oct 1, 2015

Copyright © 2015, American Society for Microbiology

**Permissions Request**

ASM authorizes an advanced degree candidate to republish the requested material in his/her doctoral thesis or dissertation. If your thesis, or dissertation, is to be published commercially, then you must reply for permission.

BACK

CLOSE WINDOW

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#), [Terms and Conditions](#).  
Comments? We would like to hear from you. E-mail us at [customer.care@copyright.com](mailto:customer.care@copyright.com)

## SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jul 05, 2018

This Agreement between Mingzhang Wang ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4382840452210
License date	Jul 05, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Microbiology
Licensed Content Title	The structural biology of HIV-1: mechanistic and therapeutic insights
Licensed Content Author	Alan Engelman, Peter Cherepanov
Licensed Content Date	Mar 16, 2012
Licensed Content Volume	10
Licensed Content Issue	4
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Structural and Dynamic Investigation of HIV-1 Capsid Assemblies and Capsid Maturation by Magic Angle Spinning NMR spectroscopy
Instructor name	Mingzhang Wang
Institution name	University of Delaware
Expected presentation date	Jul 2018
Order reference number	1
Portions	Figure 1
Requestor Location	Mingzhang Wang University of Delaware
	NEWARK, DE 19716 United States Attn: Mingzhang Wang
Billing Type	Invoice
Billing Address	Mingzhang Wang University of Delaware
	NEWARK, DE 19716 United States Attn: Mingzhang Wang

<https://s100.copyright.com/MyAccount/web/jsp/viewprintablelicensefrommyorders.jsp?ref=c0e1617d-e02c-4015-a9f3-f6da23d60c65&email=>

1/3

Total **0.00 USD**

[Terms and Conditions](#)

#### Springer Nature Terms and Conditions for RightsLink Permissions

**Springer Customer Service Centre GmbH (the Licensor)** hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where '**reuse in a dissertation/thesis**' has been selected the following terms apply: Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline ([www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/)).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines <http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/>), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to [Journalpermissions@springernature.com](mailto:Journalpermissions@springernature.com)/[bookpermissions@springernature.com](mailto:bookpermissions@springernature.com) for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

#### Appendix — Acknowledgements:

##### **For Journal Content:**

Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**]  
[**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)]

##### **For Advance Online Publication papers:**

Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**]  
[**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[**JOURNAL ACRONYM**].)]

**For Adaptations/Translations:**

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

**Note: For any republication from the British Journal of Cancer, the following credit line style applies:**

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

**For Advance Online Publication papers:**

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

**For Book content:**

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc) [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)

**Other Conditions:**

Version 1.0

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.



RightsLink®

[Home](#)[Account Info](#)[Help](#)ACS Publications  
Most Trusted. Most Cited. Most Read.**Title:**

Magic Angle Spinning NMR Spectroscopy: A Versatile Technique for Structural and Dynamic Analysis of Solid-Phase Systems

Logged in as:

Mingzhang Wang

Account #:

3001099791

[LOGOUT](#)**Author:**

Tatyana Polenova, Rupal Gupta, Amir Goldbourt

**Publication:** Analytical Chemistry**Publisher:** American Chemical Society**Date:** Jun 1, 2015

Copyright © 2015, American Chemical Society

**PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE**

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#), [Terms and Conditions](#).  
Comments? We would like to hear from you. E-mail us at [customer.care@copyright.com](mailto:customer.care@copyright.com)

## SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jul 05, 2018

This Agreement between Mingzhang Wang ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4382840727154
License date	Jul 05, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	Dynamic personalities of proteins
Licensed Content Author	Katherine Henzler-Wildman, Dorothee Kern
Licensed Content Date	Dec 12, 2007
Licensed Content Volume	450
Licensed Content Issue	7172
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Structural and Dynamic Investigation of HIV-1 Capsid Assemblies and Capsid Maturation by Magic Angle Spinning NMR spectroscopy
Instructor name	Mingzhang Wang
Institution name	University of Delaware
Expected presentation date	Jul 2018
Portions	Figure 1b
Requestor Location	Mingzhang Wang University of Delaware
	NEWARK, DE 19716 United States Attn: Mingzhang Wang
Billing Type	Invoice
Billing Address	Mingzhang Wang University of Delaware
	NEWARK, DE 19716 United States Attn: Mingzhang Wang
Total	<b>0.00 USD</b>

<https://s100.copyright.com/MyAccount/web/jsp/viewprintablelicensefrommyorders.jsp?ref=759e1e4d-cba2-44a4-8a80-3be34c44404d&email=>

1/3

[Terms and Conditions](#)**Springer Nature Terms and Conditions for RightsLink Permissions**

**Springer Customer Service Centre GmbH (the Licensor)** hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.
3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.
4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.
5. Where '**reuse in a dissertation/thesis**' has been selected the following terms apply: Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline ([www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/)).
6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines <http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/>), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.
7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to [Journalpermissions@springernature.com](mailto:Journalpermissions@springernature.com)/[bookpermissions@springernature.com](mailto:bookpermissions@springernature.com) for these rights.
8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.
9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.
10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

**Appendix — Acknowledgements:****For Journal Content:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME]  
[REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

**For Advance Online Publication papers:**

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME]  
[REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

**For Adaptations/Translations:**

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL

**NAME** [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

**Note: For any republication from the British Journal of Cancer, the following credit line style applies:**

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)

For **Advance Online Publication** papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For **Book content**:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc) [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)

**Other Conditions:**

Version 1.0

Questions? [customer-care@copyright.com](mailto:customer-care@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.



## ELSEVIER LICENSE TERMS AND CONDITIONS

Jul 05, 2018

This Agreement between Mingzhang Wang ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4382840951277
License date	Jul 05, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Progress in Nuclear Magnetic Resonance Spectroscopy
Licensed Content Title	Solid-state NMR and protein dynamics
Licensed Content Author	Alexey Krushelnitsky, Detlef Reichert
Licensed Content Date	Oct 31, 2005
Licensed Content Volume	47
Licensed Content Issue	1-2
Licensed Content Pages	25
Start Page	1
End Page	25
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation	Structural and Dynamic Investigation of HIV-1 Capsid Assemblies and Capsid Maturation by Magic Angle Spinning NMR spectroscopy
Publisher of new work	University of Delaware
Author of new work	Mingzhang Wang
Expected completion date	Jul 2018
Estimated size (number of pages)	1
Requestor Location	Mingzhang Wang University of Delaware  NEWARK, DE 19716 United States Attn: Mingzhang Wang
Publisher Tax ID	98-0397604
Total	<b>0.00 USD</b>
Terms and Conditions	

### INTRODUCTION

<https://s100.copyright.com/MyAccount/web/jsp/viewprintablelicensefrommyorders.jsp?ref=64d69545-2fd0-4c5d-91ce-d6239be0feff&email=>

1/4

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

#### GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:  
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

#### LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage:

This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.  
Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com> . All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

**Preprints:**

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

#### **Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

#### **Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

#### **Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at

<http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

#### **20. Other Conditions:**

v1.9

**Questions? [customer@copyright.com](mailto:customer@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.**



RightsLink®

[Home](#)[Account  
Info](#)[Help](#)ACS Publications Title:  
Most Trusted. Most Cited. Most Read.19F Magic Angle Spinning NMR  
Spectroscopy and Density  
Functional Theory Calculations  
of Fluorosubstituted  
Tryptophans: Integrating  
Experiment and Theory for  
Accurate Determination of  
Chemical Shift TensorsLogged in as:  
Mingzhang Wang  
Account #:  
3001099791[LOGOUT](#)**Author:** Manman Lu, Sucharita Sarkar,  
Mingzhang Wang, et al**Publication:** The Journal of Physical  
Chemistry B**Publisher:** American Chemical Society**Date:** May 1, 2018

Copyright © 2018, American Chemical Society

**PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE**

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#), [Terms and Conditions](#).  
Comments? We would like to hear from you. E-mail us at [customer.care@copyright.com](mailto:customer.care@copyright.com)