UNIVERSITY OF CALIFORNIA
RIVERSIDE

Investigation of HIV-1 Gag Binding With RNAs and Lipids

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Shaolong Chen

June 2018

Dissertation Committee:
Dr. Umar Mohideen, Chairperson
Dr. Roya Zandi
Dr. Sarjeet Gill

The Dissertation of Shaolong Chen is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

The pursuit of a PhD degree in physics is a long journey to uncovered land. It is by no means easy and needs one to put tremendous time and efforts to face scientific challenges. Joy and pain intertwined with each other throughout this journey. Fortunately, I was not alone when exploring this journey. There have been a lot of people who have been supportive and helpful.

First of all, I would like to express my gratitude to my advisor Professor Umar Mohideen for his patient guidance during my PhD program. He has sophisticated insight in problems I have encountered during my research. I have learned tremendously from him regarding data analysis and problem solving.

I would also like to sincerely acknowledge my committee members, Professor Roya Zandi and Professor Sarjeet Gill for their kind willingness to sacrifice their valuable time to serve on my dissertation committee board. In addition, they are our collaborators and they provided a great deal of precious suggestions to my research. I also want to give some credit for Professor Sarjeet Gill because he provided his fund for my research over a period of time.

There are other collaborators I want to thank. Professor Ayala Rao and his graduate students Areeje Almasary and Antara Chakravarty helped me with the gel electrophoresis experiments. Professor Karin Musier-Forsyth and her graduate student Erik Olson provided us with ΨRNA and TARpolyA RNA. Dr. Alan Rein supplied us with HIV-1 GagΔP6. Professor Sarjeet Gill's postdoc Jianwu Chen provided me with DEPC water.

Professor Yongtao Cui generously allowed me to use his dimension icon AFM. I really appreciate all of our collaborators for their kind support and valuable suggestions.

Furthermore, I would like to think my labmates Jun Xu, Mingyue Liu, Robert Schafer, and Dong Gui for their enormous assistance during my research. I also want to thank my colleagues at UCR, such as Yadong Xu, Yanmeng Shi, Weimin Zhou, Chi Tang, Changtao Hou, Zhisheng Lin, Supeng Ge, Yi Wu, Tianbai Li, Bowen Yang, Liankun Zhou, Kara Kurt, Elizabeth Kennedy, William Freeman, Kevin Myhro, Clement Wang, Ryan Foltz, Evan Sosenko, and so many others that I can not list all of them, for their precious friendship in my life. It is a precious and unforgettable experience living and studying at UCR.

Last but not least, I would like to deliver my deepest love and genuine gratitude to my family, especially my parents, Mr. Jianzhong Yao and Mrs. Cuiyun Chen, for their unconditional support and understanding for my PhD study in physics. I also really appreciate my young brother Shaowei Chen for his enormous encouragement during my difficulties. There have been so many hard times and stressful experiences during my PhD study which I can barely go through without my family.

ABSTRACT OF THE DISSERTATION

Investigation of HIV-1 Gag Binding With RNAs and Lipids

by

Shaolong Chen

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, June 2018
Dr. Umar Mohideen, Chairperson

Human immunodeficiency virus (HIV) is a member of a subgroup of retroviruses that lead to HIV infection and ultimately result in acquired immunodeficiency syndrome (AIDS), one of the most fatal viral diseases in human history. There are two types of HIV: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV-1, which was discovered first, is much more contagious and virulent than HIV-2. In addition, HIV-1 is the cause of the majority of HIV infections all over the world. HIV-2 is less severe due to its relatively low infectivity. HIV viruses infect pivotal cells of the human immune system to lead to the failure of patient's immune system. Dormancy of HIV viruses is comparatively long and the average survival time after HIV infection is 10 years without treatment. So far, millions of people around the world are suffering from HIV infection while there is no good way to treat it completely. Therefore, it is extremely important to study the process of HIV assembly to better understand the mechanism of HIV infection.

HIV assembly is an extremely complicated multi-step process. It takes place at the plasma membrane and involves the polyprotein Gag, genomic viral RNAs and lipids

phospha-tidylinositol-(4,5)-bisphosphate (PI(4,5)P2). During the early stage of HIV assembly, Gag is all that is required to assemble into a spherical, membrane-enveloped immature virion with a diameter ranging from 100-150nm. The Gag has an approximate mass of 55kDa and consists of six structural domains. From N-terminus to C-terminus, they are matrix (MA), capsid (CA), spacer peptides 1 (SP1), nucleocapsid (NC), spacer peptides 2 (SP2), and p6, respectively. The three major functional domains are MA, CA and NC. MA domain prefers to interact with PI(4,5)P2 on the lipid plasma membrane during HIV assembly. CA-CA interactions form hexagonal lattices with an approximate spacing of 8nm in the immature virion. NC domain has a high affinity to bind with the RNA packaging signal (Ψ) located in the 5'-end of a dimeric viral RNA that subsequently forms the conical shape virus core.

In this dissertation Atomic Force Microscopy (AFM) was utilized to study the morphology of Gag, ΨRNA, lipid PI(4,5)P2 as well as their different mixed complexes. They were imaged on either positively charged or negatively charged micas depending on the net charges carried by the respective materials. After achieving high quality AFM images in a liquid environment, programming scripts coded in MATLAB and Mathematica were used to analyze the data statistically. The sizes of the corresponding structures and the statistical distributions of Gag among monomer, dimer and tetramer complexes before and after mixing with either ΨRNA or PI(4,5)P2, or both were found. Then the binding principle of HIV Gag with ΨRNA and PI(4,5)P2 was investigated.

Table of Contents

Chapter 1 Biological introduction…….....1

1.1 Overview of human immunodeficiency virus Gag…….1

1.2 Introduction to RNAs……..6

1.3 Introduction to lipid phospha-tidylinositol-(4,5)-bisphosphate ……10

1.4 References……..12

Chapter 2 Interaction of Gag with ΨRNA and giant unilamellar vesicles containing lipid PI(4,5)P2 measured by confocal microscopy……16

2.1 Introduction to confocal microscopy ……16

2.2 Materials and methods……..18

2.2.1 Materials……...18

2.2.2 Preparation of giant unilamellar vesicles……..19

2.2.3 Confocal microscopy measurement……..22

2.3 Conclusion…….27

2.4 References…….….29

Chapter 3 Morphology and statistics of GagΔP6 studied by atomic force microscopy....32

3.1 Introduction to atomic force microscopy……...32

3.2 Experimental procedures……...38

3.2.1 Sample preparation……...38

3.2.2 Calibration of AFM…….…….41

3.2.3 Setup of AFM…….……..43

3.2.4 AFM measurement of structure and statistical analysis……...45

3.2.4.1 ΨRNA……..……...……...45

3.2.4.2 TARpolyA RNA……..……...……...48

3.2.4.3 GagΔP6……...……...48

3.2.4.4 PI(4,5)P2-DPhPC-POPC.……...……...52

3.2.4.5 GagΔP6-ΨRNA……...……...…...53

3.2.4.6 GagΔP6-TARpolyA RNA……...……..56

3.2.4.7 GagΔP6-PI(4,5)P2……...…….....57

3.2.4.8 PI(4,5)P2-ΨRNA-GagΔP6……...…….....60

3.3 Gel electrophoresis.……..……..64

3.3.1 Introduction to gel electrophoresis……64

3.3.2 Protocol of gel electrophoresis …….65

3.3.3 Analysis of gel electrophoresis outcomes…….....68

3.4 Summary……70

3.5 References……..71

Appendix A AFM tip calibration……...……..……..79

Appendix B Algorithm of computation of the sample size..……...…..81

Appendix C MATLAB script to analyze raw AFM image data……..……..……...82

Appendix D Mathematica script to plot histogram and fit distribution….……..….....90

List of Figures

Figure 1-1 Schematic diagram of the HIV-1 replication cycle. (Reprinted from Ganser-Pornillos et al. (2008), with permission from Elsevier [2].)...……..……...1

Figure 1-2 HIV-1 immature and mature virions. (a) Schematic model of the immature virion. (b) Schematic model of the mature virion. (c) Electron cryotomography image of the immature virion. (d) Electron cryotomography image of the mature virion. (Reprinted from Ganser-Pornillos et al. (2008), with permission from Elsevier [2].)…….……..2

Figure 1-3 Schematic representation of Gag. From N-terminus to C-terminus, six functional domains of Gag are matrix (MA) with amino-terminal myristylation (Myr), Capsid (CA), Spacer peptide 1 (SP1), Nucleocapsid (NC) Spacer peptide 2 (SP2), and P6, respectively……..……4

Figure 1-4 HIV-1 genomic RNA partial sequence structure: TARpolyA RNA……..…8

Figure 1-5 HIV-1 genomic RNA partial sequence structure: ΨRNA with stem loops SL1, SL2 and SL3...…...……..…….....10

Figure 2-1 Schematic representation of the principle of Confocal Microscopy (CM)…..16

Figure 2-2 Mini extruder and its parts. (a) Extruder stand. (b) Extruder outer casing. (c) Internal membrane support. (d) O-ring. (e) 2 filter supports. (f) Polycarbonate membrane. (g) O-ring wedged into the O-ring channel on the internal membrane support. (h) Teflon bearing. (i) Retainer nut. (j) Two glass syringes.....20

Figure 2-3 10μL 1mM GUVs containing 50μM PI(4,5)P2 in 190μL 0.2μM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. The scale bar is 20.0μm.....24

Figure 2-4 10μL 2μM labeled GagΔP6 in 190μL 0.2μM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. Blue arrows are added to help eyes to locate labeled GagΔP6. The scale bar is 20.0μm……..….25

Figure 2-5 20μL mixture of 1μM labeled GagΔP6 and 0.5mM GUVs containing 25μM PI(4,5)P2 in 180μL 0.2μM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. Blue arrows are added to help eyes to locate labeled GagΔP6. The scale bar is 20.0μm.…….…….……..26

Figure 3-1 Schematic representation of Atomic Force Microscopy (AFM)……..…...33

Figure 3-2 Image of the cantilever E and the tip of AFM probe MSNL. (a) The optical microscope image of AFM probe MSNL with cantilever E highlighted. The scale is

200μm. (b) The optical microscope image of the cantilever E with a scale of 100μm. (c) The SEM image of the tip on the cantilever E with a scale of 2μm....……...……...35

Figure 3-3 AFM tip calibration with 2nm Au Sphere. (a) AFM image of 2nm Au sphere on mica in tapping mode in liquid, the scan size is 250nm×250nm. (b) Histogram of 2nm Au sphere for the measured size (top) and the measured height (bottom). Shown in red are normal distribution fits to the peaks. The total number of samples was 419.……..42

Figure 3-4 0.5μM ΨRNA on positively charged mica(+). (a) AFM image with a scan size of 500nm×500nm. A few ΨRNAs are boxed: monomer (red) and dimer (green). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer and dimer are indicated by red arrows. The total number of samples was 551..……...45

Figure 3-5 0.5μM TARpolyA RNA on positively charged mica(+). (a) AFM image with a scan size of 500nm×500nm. A few TARpolyA RNAs are boxed: monomer (red). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks for length and height. Width is fit to a gamma distribution due to its non-negativity and skewness to right. (c) Three dimensional smooth histogram, where monomer is indicated by a red arrow. The total number of samples was 504.…...47

Figure 3-6 0.5μM GagΔP6 on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6s are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 858……49

Figure 3-7 Models of GagΔP6. (a) Monomer, (b) Dimer, (c) Tetramer. MA domain is in red, CA domain is in yellow, and NC domain is in green…….51

Figure 3-8 AFM image of lipid mixture of PI(4,5)P2-DPhPC-POPC (0.5μM : 5μM : 5μM) complex. (a) AFM image with a scan size of 1μm×1μm. (b) The line delineated in (a)...52

Figure 3-9 Mixture of GagΔP6-ΨRNA (0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6-ΨRNA complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 895.…….....54

Figure 3-10 Models of GagΔP6-ΨRNA complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, and ΨRNA is in cyan.....……..…...55

Figure 3-11 Mixture of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6-PI(4,5)P2 complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 903. ……....58

Figure 3-12 Models of GagΔP6-PI(4,5)P2 complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, and PI(4,5)P2 is in purple…….59

Figure 3-13 Mixture of PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few PI(4,5)P2-ΨRNA-GagΔP6 complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 616.……...……61

Figure 3-14 Models of PI(4,5)P2-ΨRNA-GagΔP6 complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, ΨRNA is in cyan, and PI(4,5)P2 is in purple..……63

Figure 3-15 6% Tris-Glycine SDS-PAGE. (a) GagΔP6 (2μM). (b) GagΔP6 (0.5μM). (c) GagΔP6-ΨRNA (0.5μM : 0.5μM). (d) GagΔP6-ΨRNA (0.5μM : 2μM). (e) GagΔP6-ΨRNA (2μM : 0.5μM). (f) PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM). (g) PI(4,5)P2-ΨRNA-GagΔP6 (2μM : 0.5μM : 0.5μM). (h) 50kDa Protein marker. (i) Protein standards. The rightmost is the criterion of molecular weight of protein standards. Bands corresponding to monomer, dimer, and tetramer are indicated on the left……...….68

Figure A-1 AFM tip calibration. (a) Schematic representation of the tip and a spherical sample. The red curve is the trajectory of the tip when it scans from left to right as indicated by the arrow on left. (b) The configuration of the tip after it is placed into the AFM probe holder (fluid cell if imaging in liquid)……79

Figure B-1 Schematic diagram of a sample with defined length and width.……..……...81

List of Tables

Table 1-1 Proteins encoded by the HIV-1 genome…….7

Table 2-1 Recipes for preparation of lipid vesicles……...19

Table 3-1 Recipes of buffer……...38

Table 3-2 Statistics of AFM measurement of 0.5μM ΨRNA on mica(+)…….....46

Table 3-3 Statistics of AFM measurement of 0.5μM GagΔP6 on mica(-)…….…...50

Table 3-4 Statistics of AFM measurement of GagΔP6-ΨRNA (0.5μM : 0.5μM) on mica(-)……..55

Table 3-5 Statistics of AFM measurement of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) on mica(-).…….……..59

Table 3-6 Statistics of AFM measurement of PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM) on mica(-)……..62

Table 3-7 Tris-Glycine SDS-PAGE recipes……..65

Chapter 1

Biological introduction

1.1 Overview of human immunodeficiency virus Gag

Human immunodeficiency virus (HIV) is a type of lentivirus, a member of a subgroup of retroviruses, which are single-stranded RNA viruses and contain two copies of their viral RNA genomes. HIV is a spherical membrane-enveloped virus with a reported diameter ranging from 100nm to 150nm. It initially buds from the plasma membrane of infected host cells as a noninfectious immature virion [1-5]. The formation

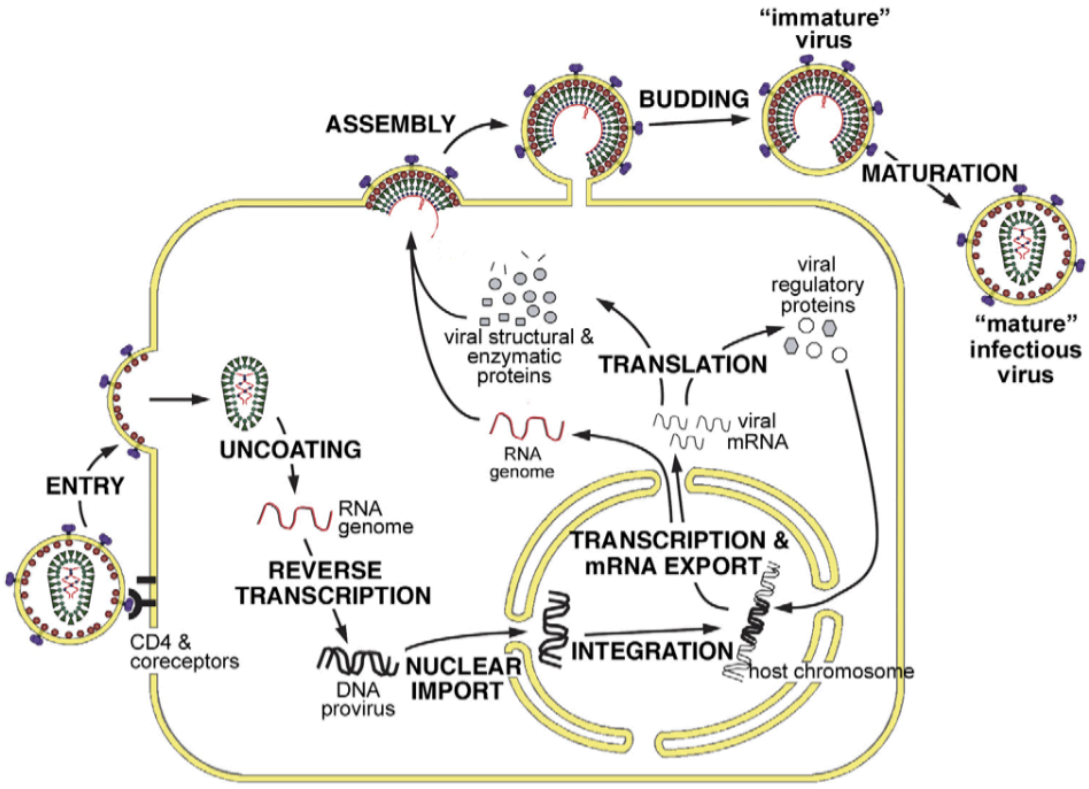

Figure 1-1 Schematic diagram of the HIV-1 replication cycle. (Reprinted from Ganser-Pornillos et al. (2008), with permission from Elsevier [2].)

process of infectious HIV has multiple steps and is usually considered to have three stages: (1) assembly, (2) budding and release, (3) and maturation [1-2, 6-8]. The schematic diagram of the HIV-1 replication cycle is shown in Figure 1-1 [2]. The first stage is the assembly where critical components are encapsulated to create an immature virion at the lipid plasma membrane. The second stage is the budding and release wherein

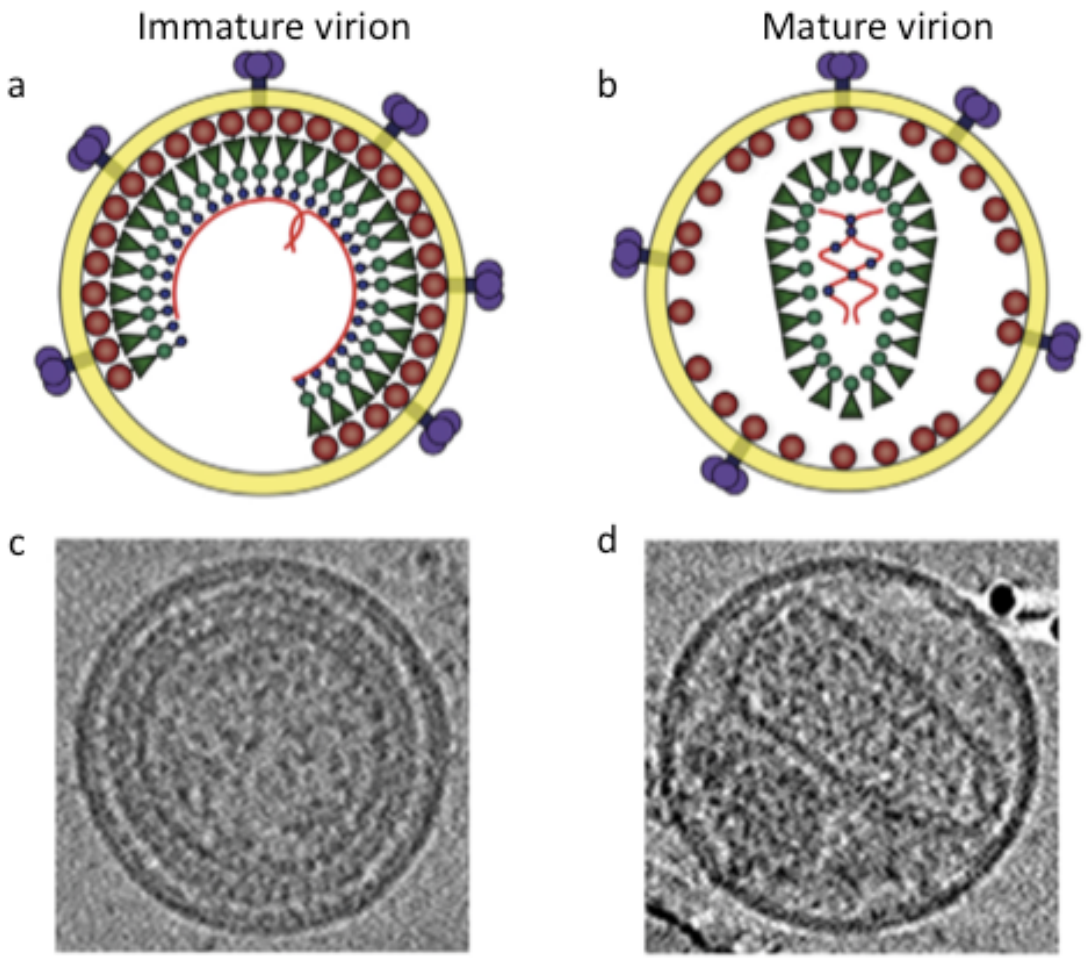

Figure 1-2 HIV-1 immature and mature virions. (a) Schematic model of the immature virion. (b) Schematic model of the mature virion. (c) Electron cryotomography image of the immature virion. (d) Electron cryotomography image of the mature virion. (Reprinted from Ganser-Pornillos et al. (2008), with permission from Elsevier [2].)

the immature virion acquires the necessary lipids to form its envelope and buds from the plasma membrane. The last stage is the maturation where the immature virion adapts conformational changes and turns into a maturely infectious virus. The schematic models of the immature virion and mature virion as well as their electron cryotomography images are shown in Figure 1-2 [2]. The most important structural component of HIV is the genetic polyprotein precursor Gag, the group specific antigen, which is involved in all of the three stages of HIV formation. It has been well studied that the HIV immature virion has approximately 2500-5000 copies of Gag polyprotein [1, 4]. The approximate mass of Gag is 55kDa [8]. Gag has rod-like shape with a length of 20-30nm and a diameter of 2-3nm [3]. From N-terminus to C-terminus, Gag consists of six structural domains: matrix (MA) with amino-terminal myristylation (Myr), capsid (CA), spacer peptides 1 (SP1), nucleocapsid (NC), spacer peptides 2 (SP2), and p6, respectively [1-2, 6-7, 9]. The schematic representation of full length Gag polypeptide is shown in Figure 1-3. The three major functional domains of Gag are MA, CA and NC. The HIV Gag precursor in the immature virion is radically oriented. The N-terminal MA domain is bound to the inner leaflet of lipid membrane. The C-terminal NC domain binds with two copies of genomic viral RNA. The central CA domain interacts with each other forming a hexagonal lattice. Gag serves as the bridge to connect the lipid membrane and dimeric viral RNA and forms a scaffold with other Gags to support the structure of the immature HIV virion. In later stages, the immature HIV virion becomes a mature virion via proteolysis with the help of a viral protease through which the HIV Gag is cleaved into pieces with specific orders [10]. The first cleavage is expected to occur between SP1 and

NC. Next the MA-CA-SP1 segment will have cleavage at the linkage between MA and CA. The NC-SP2-P6 segment will split at linkage between SP2 and P6 followed by cleavage between NC and SP2. The cleavage between CA-SP1 will occur at last. The CA cleavage at both ends by the protease seems to prompt large conformational changes.

Figure 1-3 Schematic representation of Gag. From N-terminus to C-terminus, six functional domains of Gag are matrix (MA) with amino-terminal myristylation (Myr), Capsid (CA), Spacer peptide 1 (SP1), Nucleocapsid (NC), Spacer peptide 2 (SP2), and P6, respectively.

The N-terminal MA domain of HIV-1 Gag consisting of 104 amino acids contains five alpha helices and a triple-stranded beta sheet [11-13]. MA has a myristoylated fatty acid group at its N-terminus which is responsible for Gag assembly and targeting with the phospha-tidylinositol-(4,5)-bisphosphate (PI(4,5)P2) on the lipid membrane [14-15]. The 14-carbon myristoylated group is initially sequestered in a hydrophobic cleft in the MA domain that later will be exposed to facilitate Gag binding with the lipid membrane. The exposure of the myristyl acid group triggered by PI(4,5)P2 which is abundant in the lipid plasma membrane. PI(4,5)P2 consists of a glycerol backbone with position 1' attached to a saturated fatty acid chain, position 2' attached to a unsaturated fatty acid chain, and position 3' attached to a phosphoinositol headgroup. Myristoylated group exposure and PI(4,5)P2 binding are coupled. The MA domain binds with negatively charged

phosphoinositol headgroup through nonspecific electrostatic interaction and 2' unsaturated fatty acid chain via specific interaction. The 1' saturated fatty acid chain intrudes into the hydrophobic myristyl group and triggers the myristyl group to expose and insert into lipid membrane [16]. The MA domain can also bind with nucleic acids but with less affinity in contrast to the lipid PI(4,5)P2 [17]. It is reported cholesterol can increase the selectivity of MA domain binding with lipid plasma membrane containing PI(4,5)P2. MA domain can assemble into trimers which organize into hexamers on plasma membrane that interconnect with CA hexamers in the immature virion [18].

The central CA domain of HIV Gag is responsible for adjacent Gag-Gag interaction. CA domain is separated into two parts, the N-terminal domain (CA^{NTD}) and the C-terminal domain (CA^{CTD}) connected by a flexible linker. The arrowhead-like shaped CA^{NTD} containing seven alpha helices is essential for the formation of conical outer shell of the capsid core [2, 19-20]. The CA^{NTD} forms hexameric rings with an approximate spacing of 8nm [21-22]. The CA^{NTD} also forms pentameric rings which are not necessary for the immature virions but are indispensable for formation of a closed shell of the core that adopts a fullerene cone structure in the mature virus. The globular CA^{CTD} comprises a short 3_{10} helix followed by a major homology region and four alpha helices [23-24]. The CA^{CTD} forms symmetric homodimers that link hexameric and pentameric rings into another layer of hexamers that is slightly smaller than the CA^{NTD} hexamers [22-23, 25]. Some studies showed that there is a third layer of hexamers formed by SP1. The spacing of each layer of hexamers decreases gradually across the sequential layers CA^{NTD}, CA^{CTD} and SP1. The CA domain plays a vital role in the formation of both immature and mature

virions. In the mature virus, the mature capsid core consists of 1000-1500 copies of the mature CA protein assembled into a hexameric lattice with a spacing of 10nm rather than 8nm which is the spacing of CA hexamers in the immature virions [4]. The mature CA hexamers are different from the immature compartments due to these different spacing distances and therefore must be rearranged during the stage of maturation.

Although the NC domain is the smallest component of HIV Gag compared to MA and CA domains, it is critical for nucleic acid recognition and interaction with genomic viral RNA dimerization. The NC domain tethers Gag to RNA genome through nonspecific interaction as well as specific binding to the stem-loop 3 (SL3) in the packaging signal Ψ (ΨRNA) [26]. In the inner core of the mature virus, the genomic viral RNA is wrapped by about 1500-2000 copies of NC proteins [27]. Within the NC domain, there are two CCHC type zinc fingers which are crucial for specific ΨRNA binging and genomic viral RNA packaging [28-29]. The zinc finger has a pattern Cys- Cys- X_2-Cys- X_4-His-X_4-Cys, where X's represent variable amino acid residues [8, 26]. The spacing distance between two zinc fingers is highly conserved which is important for NC functions [30]. Genomic RNA dimerization and packaging are coupled and require the presence of NC domain.

1.2 Introduction to RNAs

HIV-1 genome is principally a sequence of RNA that consists of 9173 nucleotides (~9kb) [9, 31]. HIV-1 genomic RNA involves in a lot of activities throughout virus

replication cycle, such as expressing transcription, facilitating genomic dimerization and transportation, signaling polyadenylation, mediating HIV genome packaging, initiating reverse transcription etc. The HIV-1 genome comprises 9 genes encoding 18 functionally distinct proteins, three of which encode the viral structural polyproteins Gag, Pol and Env that are common to all retroviruses [32]. All the 18 proteins encoded by the HIV-1 genome and their corresponding gene names are listed in Table 1-1. The Gag polyprotein

Table 1-1 Proteins encoded by the HIV-1 genome

Protein class	Gene name	Primary proteins	Proteolyzed proteins
Viral structural proteins	gag	Gag	MA, CA, SP1, NC, SP2,P6
	pol	Pol	RT, RNase H, IN, PR
	env	gp160	gp120, gp41
Esssential regulatory proteins	tat	Tat	
	rev	Rev	
Accessory regulatory proteins	nef	Nef	
	vif	Vif	
	vpr	Vpr	
	vpu	Vpu	

precursor is subsequently proteolyzed, in the presence of the enzymatic protein protease (PR) from Pol polyprotein, into individual proteins MA, CA, SP1, NC, SP2 and P6. The Pol polyprotein is later divided into reverse transcriptase (RT), RNase H, integrase (IN) and PR. The Env polyprotein will split into surface (SU, or gp120) and transmembrane (TM, or gp41). There are another 2 essential regulatory proteins: Tat and Rev, and 4 accessory regulatory proteins: Nef, Vif, Vpr, and Vpu encoded by HIV-1 genome. HIV-1 genome is always packed into the virion as a RNA dimer, i.e. two copies of full length ~9kb RNA, which is 5' capped and 3' polyadenylated. At the 5' untranslated region

(UTR) of HIV-1 genomic RNA, there are a variety of critical regions which are considered important for genome dimerization and binding with Gag: the transactivation response stem loop (TAR), the polyadenylation stem loop (polyA), the prime binding site (PBS) and the packaging signal domain Ψ.

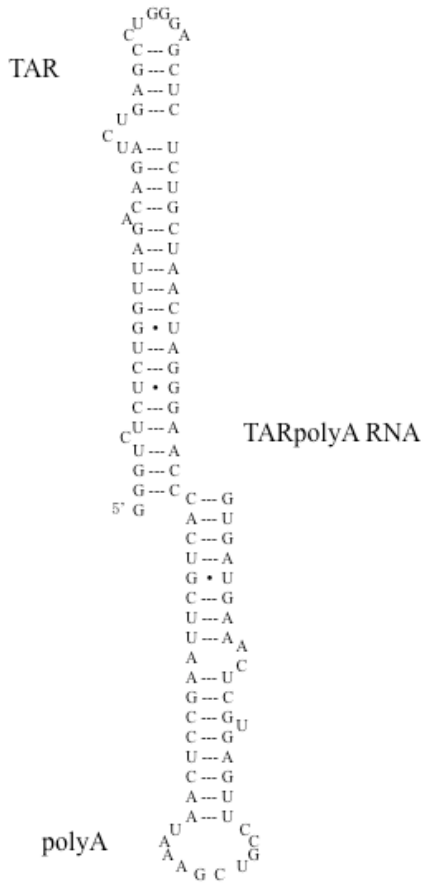

Figure 1-4 HIV-1 genomic RNA partial sequence structure: TARpolyA RNA.

The first 104 nucleotides of HIV viral RNA consists of TAR from nucleotides 1-57 and polyA from nucleotides 58-104 [33-34]. The mass of the TARpolyA RNA is around 34kDa. The TARpolyA RNA sequence used in the experiments is shown in Figure 1-4

[28]. It plays an essential role in HIV genome packaging and reverse transcription. The TAR hairpin is responsible for Tat protein binding while the polyA hairpin contains the polyadenylation signal. Small-angle X-ray scattering (SAXS) study showed TAR and polyA hairpins extend into a stable coaxially stacked helices [33]. In addition to the dimerization initiation site (DIS) in Ψ, the TAR may also facilitate HIV genomic RNA dimerization only when the NC protein is present. This is because the exposure of the hidden palindromic sequence in the TAR hairpin requires the NC protein to form TAR-TAR dimers [35].

The packaging signal Ψ contains approximately 109 nucleotides, some reported nucleotides 241-350 while other study showed nucleotides 228-334 [32-33]. The mass of the ΨRNA is about 36kDa. The ΨRNA sequence used in the experiments is shown in Figure 1-5 where most of nucleotides are paired with each other [33]. It contains 4 stem loops SL1, SL2, SL3 and SL4. SAXS revealed ΨRNA adopts an unfolded conformation where all stem loops are open for later interaction with both viral and host elements. The ΨRNA binds with the NC protein and is crucial for packaging of HIV genomic RNA into the virus. Within the ΨRNA, SL1 contains a palindromic sequence DIS that is responsible for HIV genomic RNA dimerization and Gag binding at the early stage of HIV virus replication cycle [36-38]. SL2 includes the splice donor site (SD) that is used to produce all subsequent spliced message RNAs (mRNA) for production and translation of viral accessory proteins. SL3 is required for both viral RNA dimerization and packaging [39]. SL4 contains Gag start codon. Besides SL1, SL3 and SL4 are another two independent high affinity Gag binding sites [40].

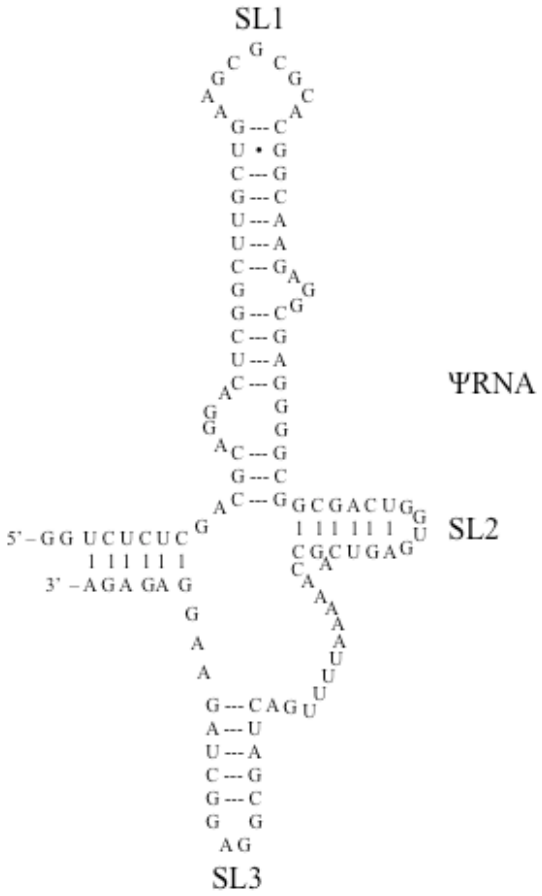

Figure 1-5 HIV-1 genomic RNA partial sequence structure: ΨRNA with stem loops SL1, SL2 and SL3.

1.3 Introduction to lipid phosphatidylinositol-(4,5)-bisphosphate

Phosphatidylinositol-(4,5)-bisphosphate (PI(4,5)P2) belongs to the negatively charged lipid family known as phosphoinositide. PI(4,5)P2 has chemical formula $C_{47}H_{80}O_{19}P_3$ and molecular weight 1042g/mol. PI(4,5)P2 consists of a glycerol backbone with one saturated fatty acid chain at position 1' and one unsaturated fatty acid chain at position 2' and a phosphoinositol headgroup at position 3'. PI(4,5)P2 is an important component of

host cell membrane and is enriched at the cytoplasmic leaflet of the plasma membrane. PI(4,5)P2 serves as a raft for HIV-1 Gag targeting to the plasma membrane and regulates the HIV virus assembly [41-42]. More specifically, PI(4,5)P2 directly binds with the myristylated MA domain of Gag through highly specific interaction. First, the phosphoinositol headgroup and the 2' unsaturated fatty acid chain of PI(4,5)P2 insert into a hydrophobic pocket in the MA domain. Then this triggers the exposure of the myristylated group to intrude into the lipid membrane bilayer [16, 43-45]. The 1' saturated fatty acid chain serves as an anchor for Gag targeting to plasma membrane although it does not interact with the MA domain. Therefore, there are two saturated fatty acid chain anchoring Gag-membrane binding, one is the myristylated group from the MA domain and the other is the 1' saturated fatty acid chain from PI(4,5)P2 [1]. In summary, PI(4,5)P2 contributes to the Gag targeting to the plasma membrane via the interaction of the MA domain and promotes the virus assembly by enhancing the Gag multimerization.

1.4 References

[1] Ganser-Pornillos B.K., Yeager M., and Pornillos O. (2012). Assembly and architecture of HIV. Adv. Exp. Med. Biol., 726, 441-465.

[2] Ganser-Pornillos B.K., Yeager M., and Sundquist W.I. (2008). The structural biology of HIV assembly. Curr. Opin. Struct. Biol., 18, 2, 203-217.

[3] Rein A., Datta S.A.K., Jones C.P., and Musier-Forsyth K. (2011). Diverse interactions of retroviral Gag proteins with RNAs. Trends. Biochem. Sci., 36, 373-380.

[4] Briggs J.A.G., Simon M.N., Gross I., Kräusslich H.G., Fuller S.D., Vogt V.M., and Johnson M.C. (2004). The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol., 11, 672–675.

[5] Datta S.A.K., Curtis J.E., Ratcliff W., Clark P.K., Crist R.M., Lebowitz J., Krueger S., Rein A. (2007). Conformation of the HIV-1 Gag protein in solution. J. Mol. Biol., 365, 812-824.

[6] Sundquist, W.I. and Kräusslich, H.G. (2012). HIV-1 Assembly, budding, and maturation. CSH Perspect. Med., 2, 7, a006924.

[7] Freed E.O. (2015). HIV-1 assembly, release and maturation. Nat. Rev. Microbiol., 13, 484–496.

[8] Bell N.M., and Lever A.M.L. (2013). HIV Gag polyprotein: processing and early viral particle assembly. Trends Microbiol., 21, 136-144.

[9] Briggs J.A.G., and Kräusslich H.G. (2011). The Molecular Architecture of HIV. J. Mol. Biol., 410, 491-500.

[10] Pettit S.C., Sheng N., Tritch R., Erickson-Viitanen S., Swanstrom R. (1998). The regulation of sequential processing of HIV-1 Gag by the viral protease. Adv. Exp. Med. Biol., 436, 15-25.

[11] Massiah M.A., Starich M.R., Paschall C., Summers M.F., Christensen A.M., and Sundquist W.I. (1994). Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein, J. Mol. Biol., 244, 198-223.

[12] Chukkapalli V., and Ono A. (2011). Molecular determinants that regulate plasma membrane association of HIV-1 Gag. J. Mol. Biol., 410, 512-524.

[13] Hill C.P., Worthylake D., Bancroft D.P., Christensen A.M., and Sundquist W.I.. (1996). Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. PNAS, 93, 7, 3099-3104.

[14] Göttlinger H.G., Sodroski J.G., and Haseltine W.A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. PNAS, 86, 15, 5781-5785.

[15] Spearman P., Wang J.J., Heyden N.V., and Ratner L. (1994). Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. J. Virol., 68, 5, 3232-3242.

[16] Saad J.S., Miller J., Tai J., Kim A., Ghanam R.H., and Summers M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. PNAS, 103, 30, 11364-11369.

[17] Alfadhli A., Still A., and Barklis E. (2009). Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids. J. Virol., 82, 23, 12196-12203.

[18] Alfadhli A., Barklis R.L., and Barklis E. (2009). HIV-1 matrix organizes as a hexamer of trimers on membranes containing phosphatidylinositol-(4,5)-bisphosphate. Virology, 387, 2, 466-472.

[19] Bharat T.A.M., Davey N.E., Ulbrich P., Riches J.D., de Marco A., Rumlova M., Sachse C., Ruml T., and Briggs J.A. G. (2012). Structure of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy. Nature, 487, 385–389.

[20] Kelly B.N., Howard B.R., Wang H., Robinson H., Sundquist W.I., Hill C.P. (2006). Implications for viral capsid assembly from crystal structures of HIV-1 Gag(1-278) and CA(N)(133-278). Biochemistry, 45, 38, 11257–11266.

[21] Wright E.R., Schooler J.B., Ding H.J., Kieffer C., Fillmore C., Sundquist W.I., and Jensen G.J. (2007). Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J., 26, 2218-2226.

[22] Briggs J.A., Riches J.D., Glass B., Bartonova V., Zanetti G., and Kräusslich H.G. (2009). Structure and assembly of immature HIV. PNAS, 106, 27, 11090-11095.

[23] Gamble T.R., Yoo S., Vajdos F.F., von Schwedler U.K., Worthylake D.K., Wang H., McCutcheon J.P., Sundquist W.I., and Hill C.P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. Science, 278, 5339, 849-853.

[24] Worthylake D.K., Wang H., Yoo S., Sundquist W.I., and Hill C.P. (1999). Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. Acta Crystallogr. D Biol. Crystallogr., 55, 85-92.

[25] Pornillos O., Ganser-Pornillos B.K., and Yeager M. (2011). Atomic level modeling of the HIV capsid. Nature, 469, 7330, 424-427.

[26] De Guzman R.N., Wu Z.R., Stalling C.C., Pappalardo L., Borer P.N., and Summers M.F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science, 16, 279,5349, 384-388.

[27] Godet J., Boudier C., Humbert N., Ivanyi-Nagy R., Darlix J.L., and Mély Y. (2012). Comparative nucleic acid chaperone properties of the nucleocapsid protein NCp7 and Tat protein of HIV-1. Virus Res., 169, 2, 349-360.

[28] Webb J.A., Jones C.P., Parent L.J., Rouzina I., Musier-Forsyth K. (2013). Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging. RNA, 19, 8, 1078-1088.

[29] Kafaie J., Song R., Abrahamyan L., Mouland A.J., and Laughrea M. (2008). Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging. Virology, 375, 2, 592-610.

[30] Thomas J.A., and Gorelick R.J. (2008). Nucleocapsid protein function in early infection processes. Virus Res., 134, 1-2, 39-63.

[31] Watts J.M., Dang K.K., Gorelick R.J., Leonard C.W., Bess Jr J.W., Swanstrom R., Burch C.L., and Weeks K.M. (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. Nature, 460, 711–716.

[32] Frankel A.D., and Young J.A. (1998). HIV-1: fifteen proteins and an RNA. Annu. Rev. Biochem., 67, 1-25.

[33] Jones C.P., Cantara W.A., Olson E.D., and Musier-Forsyth K. (2014). Small-angle X-ray scattering-derived structure of the HIV-1 5' UTR reveals 3D tRNA mimicry. PNAS, 111, 9, 3395-3400.

[34] Wilkinson K.A., Gorelick R.J., Vasa S.M., Guex N., Rein A., Mathews D.H., Giddings M.C., and Weeks K.M. (2008). High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol. 6, 4, e96.

[35] Andersen E.S., Contera S.A., Knudsen B., Damgaard C.K., Besenbacher F., Kjems J. (2004). Role of the trans-activation response element in dimerization of HIV-1 RNA. J. Biol. Chem., 279, 21, 22243-22249.

[36] Moore M.D., and Hu W.S. (2009). HIV-1 RNA dimerization: It takes two to tango. AIDS Rev., 11, 2, 91-102.

[37] Skripkin E., Paillart J.C., Marquet R., Ehresmann B., and Ehresmann C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. PNAS, 91, 11, 4945-4949.

[38] Lu K., Heng X., and Summers M.F. (2011). Structural determinants and mechanism of HIV-1 genome packaging. J. Mol. Biol., 410, 2, 609-633.

[39] Russell R.S., Hu J., Bériault V., Mouland A.J., Laughrea M., Kleiman L., Wainberg M.A., Liang C. (2003). Sequences downstream of the 5' splice donor site are required for both packaging and dimerization of human immunodeficiency virus type 1 RNA. J. Virol., 77, 1, 84-96.

[40] Clever J.L., and Parslow T.G. (1997). Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J. Virol., 71, 5, 3407-3414.

[41] Ono A., Ablan S.D., Lockett S.J., Nagashima K., Freed E.O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. PNAS, 101, 41, 14889-14894.

[42] Zhou W., Parent L.J., Wills J.W., and Resh M.D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J. Virol., 68, 4 2556-2569.

[43] Zhou W., and Resh M.D. (1996). Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. J. Virol., 70, 12, 8540-8548.

[44] Spearman P., Horton R., Ratner L., Kuli-Zade I. (1997). Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. J. Virol., 71, 9, 6582-6592.

[45] Ono A., and Freed E.O. (1999). Binding of Human Immunodeficiency Virus Type 1 Gag to Membrane: Role of the Matrix Amino Terminus. J. Virol., 73, 5, 4136-4144.

Chapter 2

Interaction of Gag with ΨRNA and giant unilamellar vesicles containing lipid PI(4,5)P2 measured by confocal microscopy

2.1 Introduction to confocal microscopy

Confocal microscopy (CM), also known as confocal laser scanning microscopy, is an optical imaging technique with highly enhanced resolution. The principle of CM, invented in 1957 by Marvin Minsky, is to place an aperture at the confocal plane of the lens to eliminate out-of-focus light [1-4]. The schematic representation of the principle of

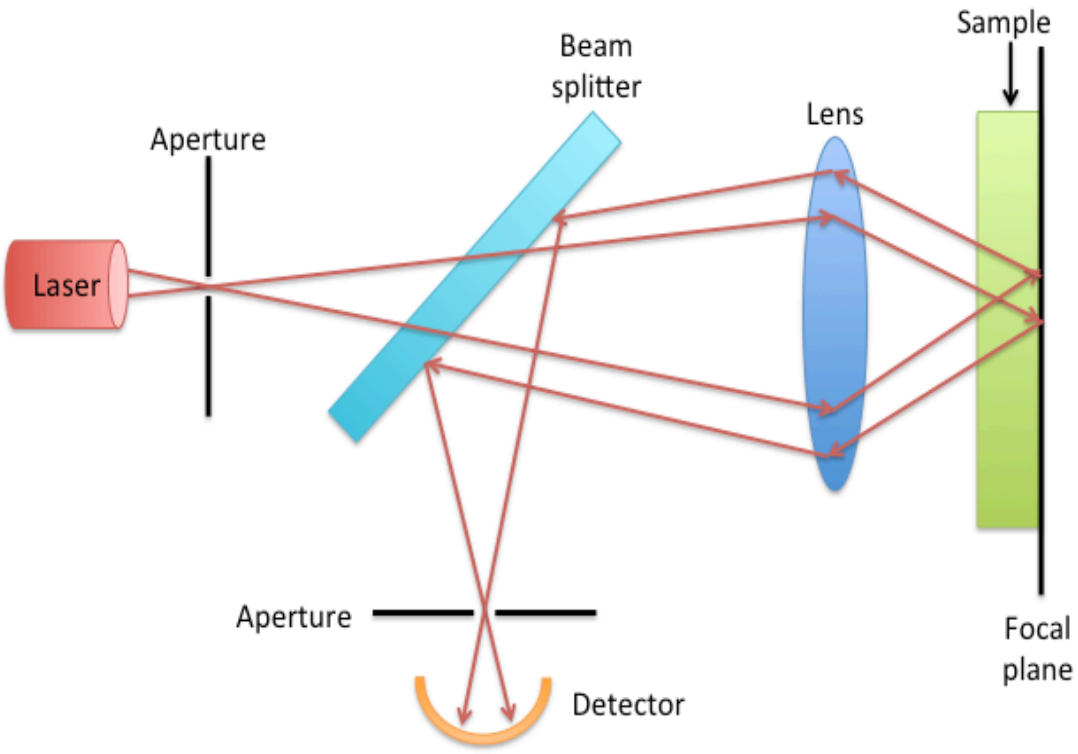

Figure 2-1 Schematic representation of the principle of Confocal Microscopy (CM).

CM is shown in Fig 2-1. CM increases the contrast of microscope images for thick samples at different depths in the sample through the process of optical sectioning [5]. Optical sectioning enables researchers to reconstruct the three dimensional structures even for a deep imaged point in a thick sample. CM has a variety of advantages over conventional microscopes, including resolution enhancement, no artifacts, three dimensional image reconstruction, etc. On the other hand, CM also has a limitation in the depth of imaging which is essentially restricted by light penetration as well as signal to noise ratio [6]. CM is a scientific technique that is widely used in biological science, chemistry, quantum optics as well as materials science.

A beam of laser light passes through an optical apparatus called an aperture and a beam splitter. The beam split by the beam splitter passes through a microscope convex lens and then is focused on a really small spot deep inside a thick sample. The imaged spot can be as small as about 500nm in diameter, which is limited by the wavelength of light used [6]. The same microscope convex lens collects the light reflected back from the sample and projects on to the beam splitter. The beam splitter reflects the light again to pass through a second aperture before being captured by a light detector. By using the aperture that has adjustable size and has to be placed properly, any out-of-focus light will be blocked. Only the light reflected back from the desirable focal spot on the thick sample can pass through so as to enhance the contrast of the microscope image and achieve higher resolution than conventional optical microscopes. On the other hand, CM can only scan one point at a time due to the exclusion of light by the confocal aperture. To restructure a three dimensional image of the thick sample, the focal spot is quickly

scanned in the horizontal plane and then moved in the vertical direction to virtually section the thick sample at different depths. All the points of the image in the same horizontal plane are collected and stacked with those in different horizontal planes to create a virtual three dimensional image of a thick sample [6].

The CM used in our experiments (Leica inverted SP5 confocal microscopy, Leica Microsystems Inc., Buffalo Grove, IL, US) is a fluorescence CM which uses dyes to fluoresce when excited by light. We used a CM 40× water immersion microscope objective in the experiments described below. The two exciting laser sources are HeNe and Argon to generate red and green lights with wavelength of 633nm and 488nm, respectively. Therefore, the spectrums of the collection light wavelength filters are 600-700nm for the red light and 450-550nm for the green light.

2.2 Materials and methods

2.2.1 Materials

The materials used to make the giant unilamellar vesicles (GUV) are cholesterol, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC), and L-α-phosphatidylinositol-4,5-bisphosphate (PI(4,5)P2) (all from Avanti Polar Lipids, Alabaster, AL, US) [7]. The fluorescent stain used to label the lipid membrane is 1,1'-dioctadecyl-3,3,3',3'-tetramethylin dodicarbocyanine perchlorate ($DilC_{18}$). The dye used to label the Gag is DyLight 488 NHS-Ester (Thermo Fisher Scientific, Carlsbad, CA, US). Other chemicals used in our experiments are methanol, chloroform, urea, sucrose and glucose.

2.2.2 Preparation of giant unilamellar vesicles

There are a variety of methods to make GUVs, including large unilamellar vesicles (LUV) fusion, electroformation and dehydration-rehydration [8-11]. The principle behind the different methods of making GUVs is similar. They all require dehydrating lipid films to swell and then rehydrating to produce GUVs. The typical size of GUV made from these methods ranges from 1μm to 100μm. In our experiments we tried both LUV fusion and dehydration-rehydration methods. LUV fusion is a very simple method to make GUVs. However, it is usually not able to produce sufficient GUVs as our experiments required. Although dehydration-rehydration method can generates vesicles as large as 50μm in diameters, it has a disadvantage over the LUV fusion method in that vesicles made by dehydration-rehydration are not always unilamellar.

Table 2-1 Recipes for preparation of lipid vesicles

Lipid	Molecular weight(g/mol)	Concentration(mg/mL)	Molar ratio	Volume(μL)
DPHDC	846.25	10	11	46.5
POPC	760.08	10	4	15.7
PI(4,5)P2	1096.39	1	1	54.8
cholesterol	386.65	5	4	15.4

To produce GUVs via the LUV fusion method, first one needs to make LUVs through a standard technique called extrusion [12-16]. The procedure of LUV fusion is straightforward. Next, use glass syringes to make a mixed lipid solution in a glass vial with molar ratio of DPhDP : POPC : PI(4,5)P2 : cholesterol = 11 : 4 : 1 : 4. The detailed

recipe is given as in Table 2-1. To avoid contamination glass syringes rather than plastic pipettes have to be used to transfer lipids that are dissolved in chloroform. Then, a glass syringe is used to add 10μL 0.4mg/mL $DilC_{18}$ to the mixed lipid solution. $DilC_{18}$ is used to dye the GUVs. Use glass syringes to add 150μL methanol and 312μL chloroform to the mixed solution. Dry it using nitrogen gas and then put in vacuum for at least two hours to remove the chloroform. Add 1mL 0.2M sucrose to hydrate the lipid film formed on a Teflon disk to achieve the lipid concentration of 1mM. Sucrose is used to make GUVs heavier and does not participate in any reactions. Next we incubate the lipid

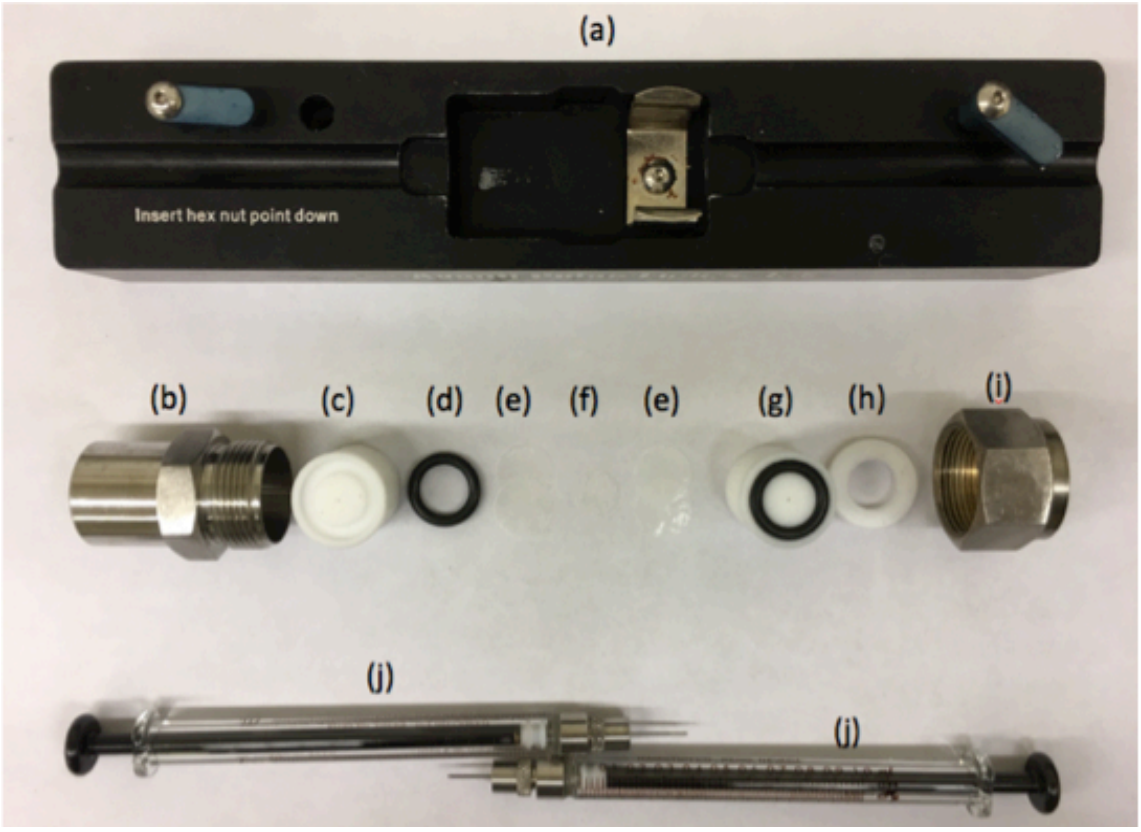

Figure 2-2 Mini extruder and its parts. (a) Extruder stand. (b) Extruder outer casing. (c) Internal membrane support. (d) O-ring. (e) 2 filter supports. (f) Polycarbonate membrane. (g) O-ring wedged into the O-ring channel on the internal membrane support. (h) Teflon bearing. (i) Retainer nut. (j) Two glass syringes.

solution in water bath at 55 ^{0}C for one hour. The lipid solution mixture is repeatedly frozen in liquid nitrogen and then thawed in hot water for at least 5 cycles. The next step is to use a mini extruder (Avanti Polar Lipids, Alabaster, AL, US) to generate the LUVs. The image of the mini extruder and all of its parts are shown in Figure 2-2. To use the mini extruder we first assemble it and then pre-warm the extruder assembly on a hot plate at 70 ^{0}C for 5 minutes. Load the fully hydrated lipid solution into one 1mL gas-tight glass syringe and gently insert into one end of the mini extruder. An empty 1mL gas-tight glass syringe is placed at the other end of the mini extruder. After 5 minutes to allow the temperature of the lipid solution to reach the equilibrium with the extruder assembly. The plunger of the syringe filled with lipid solution is slowly pushed until the lipid solution is fully transferred into the empty syringe. On repeating such transfers for about 10 times ensures that the lipid solution passes through the polycarbonate membrane. After which we inject the LUV solution into a clean glass vial. The LUV solution is kept at room temperature for 48 hours to allow LUVs to fuse into GUVs. The GUV solution is used immediately for experiments or stored in a refrigerator at -4 ^{0}C for later use within one week.

The procedure of production of GUVs through dehydration-rehydration method is as follows. The first a few steps are similar to that of production of LUVs. First, use glass syringes to make a desirable lipid solution in a glass vial with molar ratio of DPhDP : POPC : PI(4,5)P2 : cholesterol = 11 : 4 : 1 : 4. The recipe is the same as in Table 2-1. Second, use a glass syringe to add 10μL 0.4mg/mL DilC_{18} to the mixed lipid solution. Then, use plastic pipettes to add 150μL methanol and 312μL chloroform to the mixed

solution. Drop 70μL mixed solution on a clean Teflon disk and spread out the solution by the tip of the pipette. Wait for 5minutes and then dry it gently under slow nitrogen gas flow to remove all air on its surface. Place the Teflon disk with the dry lipid film on top of it into a glass petri dish. Cover it with aluminum foil to block light and then keep it in vacuum for at least 6 hours. Remove the aluminum foil and transfer the glass petri dish into a glass container filled with distilled water. Add at least 1mL 0.2M sucrose to fully cover the lipid film on the Teflon disk. The addition of sucrose is just to make GUVs heavier so that they are less mobile when measured under CM. Seal the glass container with parafilms and blow argon gas carefully through the edge to eliminate air before completely sealing the container. Put the sealed glass container upon a hot plate, set the temperature at 75 ^{0}C. Incubate it in the warm water bath over night. Make sure the glass container has enough distilled water and that the parafilms have sealed the glass container completely. GUVs should form substantially after at least 12 hours incubation. Use a transfer pipette to collect GUVs into a clean vial. Clean up Teflon disk and other parts with ethanol and dry with nitrogen gas for next use. Use GUV solution for experiments or store it in a refrigerator at -4 ^{0}C for later use within one week.

2.2.3 Confocal microscopy measurement

We need to dye the HIV Gag before the CM experiments. HIV Gag used in our experiments is unmyristoylated and lacks the P6 segment (GagΔP6, a generous gift from Dr. Alan Rein). We use GagΔP6 at 30μM concentration in 20 mM Tris- HCl pH 7.5 buffer with 0.5 M NaCl, 10 % (v/v) Glycerol, 5 mM DTT. The procedure of labeling

GagΔP6 with DyLight 488 is as follows. First, add 70μL 0.67M borate buffer contained in the DyLight 488 microscale antibody labeling kit is added into a vial of DyLight 488. Next, 5μL 30μM GagΔP6 is added into the DyLight 488 vial to get 2μM dye labeled GagΔP6. Then, the vial is put on a vortex mixer and run for 5 minutes to make sure it is completely mixed. The vial is covered with aluminum foil and then incubated at room temperature for 1 hour. Next, 100μL purification resin is added in a spin column nested with a microcentrifuge collection tube. This is placed in a centrifuge with a balance and run at 1000g for 60 seconds to remove the storage solution. Discard the used microcentrifuge collection tube and put the spin column into a new collection tube. Then, 75μL 2μM labeled GagΔP6 is added into the spin column and vortexed to allow the labeled GagΔP6 to fully mix with the resin. Centrifuge the spin column nested with the collection tube at 1000g for another 60 seconds to collect the purified labeled GagΔP6 at a concentration of 2μM. The labeled GagΔP6 can be stored at 4 ^{0}C in the dark for up to one month.

Following the labeling 50μL of the 2μM labeled GagΔP6 is mixed with 50μL 1mM GUVs containing 50μM PI(4,5)P2. The mixture is incubated at room temperature for 3 hours to get 100μL mixture of 1μM labeled GagΔP6 and 0.5mM GUVs containing 25μM PI(4,5)P2. Then, use CM to measure (I) GUVs itself as control, (II) labeled GagΔP6 as control, (III) incubated GUVs with labeled GagΔP6 to confirm if GagΔP6 reacts with PI(4,5)P2, and (IV) GagΔP6 interaction with PI(4,5)P2 in real time by first locating one GUV containing PI(4,5)P2 followed by adding labeled GagΔP6 while keeping tracking the same GUV. The procedure for the operation of CM is as follows.

Turn on the Leica inverted SP5 CM. Add one drop of distilled water on the chosen 40.0x water immersion objective lens and adjust its height. Adjust the holder and place a glass bottom culture dish (35 mm Dish, No. 0 Coverslip, 10 mm Glass Diameter, MatTek Corporation, Ashland, MA, US). For (I) add 190μL 0.2μM glucose and 10μL 1mM GUVs containing 50μM PI(4,5)P2. For (II) add 190μL 0.2μM glucose and 10μL 10μL 2μM labeled GagΔP6. For (III) add 180μL 0.2μM glucose and 20μL mixture of 1μM labeled GagΔP6 and 0.5mM GUVs containing 25μM PI(4,5)P2. For (IV) add 180μL 0.2μM glucose, 10μL 1mM GUVs containing 50μM PI(4,5)P2, and 10μL 2μM labeled GagΔP6 after locating one GUV. Set all parameters properly and try to locate a perfect GUV containing PI(4,5)P2 or GagΔP6 combined GUV containing PI(4,5)P2.

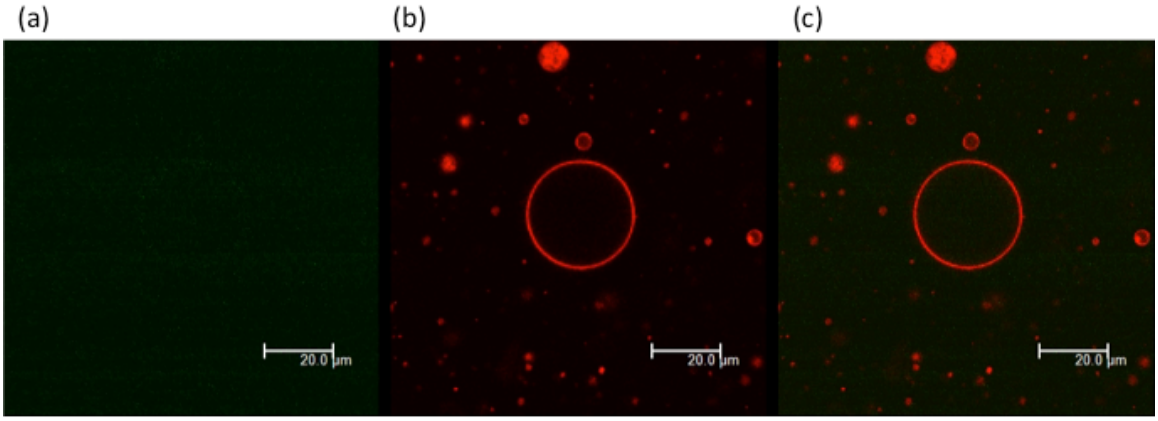

Figure 2-3 10μL 1mM GUVs containing 50μM PI(4,5)P2 in 190μL 0.2μM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. The scale bar is 20.0μm.

GUVs containing PI(4,5)P2 labeled by $DilC_{18}$ can be observed only in the red channel that corresponds to the detection spectrum of the light ranging from 600-700nm. Stable GUVs obtained through dehydration-rehydration method can be as large as 30μm

in diameter as shown in Figure 2-3. We have tried to add both NaCl and urea into GUVs. GUVs are stable and not sensitive to addition of both NaCl and urea when their concentrations are relatively low, 50mM for NaCl and 60mM for urea. However, GUVs are inclined to shrink and burst when the concentration of urea is 1M or higher.

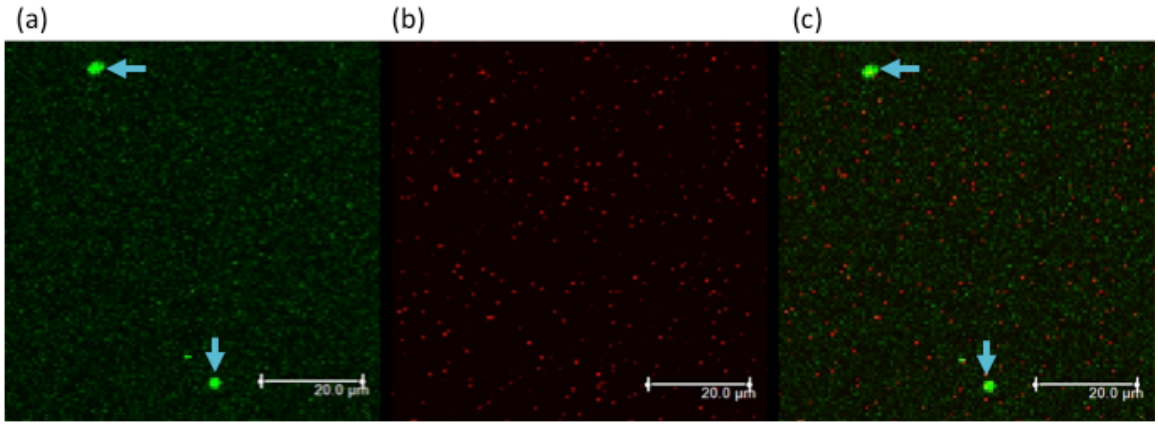

Figure 2-4 10µL 2µM labeled GagΔP6 in 190µL 0.2µM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. Blue arrows are added to help eyes to locate labeled GagΔP6. The scale bar is 20.0µm.

The GagΔP6 labeled with the dye DyLight 488 can be observed only through the green channel that corresponds to the detection spectrum of the light ranging from 450-550nm as for scenario (II). The labeled GagΔP6 is smaller than 1µm in diameter under CM so blue arrows have been added to help with the identification as shown in Figure 2-4. It turns out this size of GagΔP6 is too large to be real and the precise size and shape of the GagΔP6 cannot be obtained by CM due to its optical resolution limit. In other reports, HIV Gag in an immature virion has a shape of a rod with length of 20-30nm and diameter of 2-3nm. For virus like particles formed by Gags they are only 25-30nm in diameter with shell thickness of 8nm [17]. Even for a complete immature or mature HIV virus, it is

only as large as up to 150nm in diameter [17-21]. The measured thickness of GUVs by CM is around 700nm on average while the actual thickness of the lipid membrane of GUVs is only 5nm. This is reasonable given that the resolution limit of CM is the wavelength of visible light, i.e. 500nm. Therefore, the actual size of GagΔP6 should be on the scale of tens of nm.

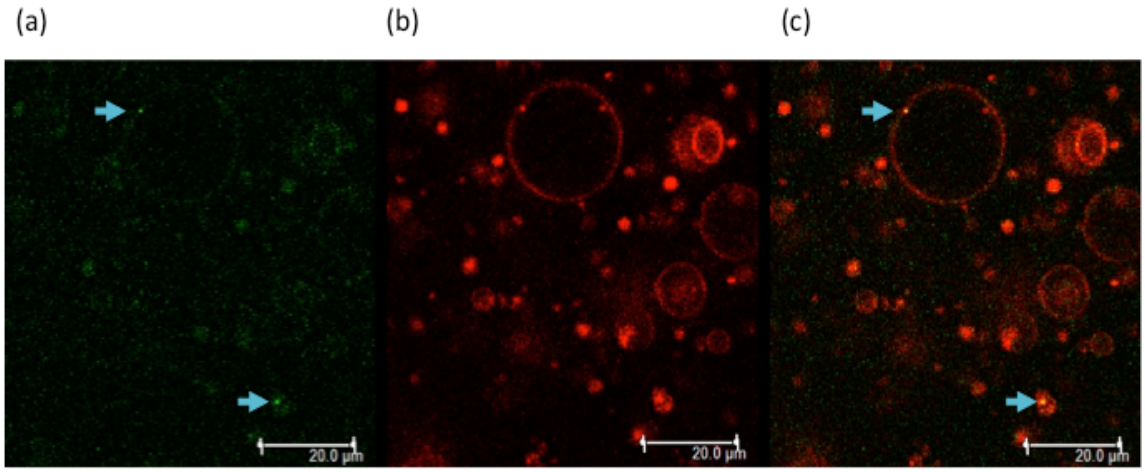

Figure 2-5 20µL mixture of 1µM labeled GagΔP6 and 0.5mM GUVs containing 25µM PI(4,5)P2 in 180µL 0.2µM glucose. (a) Green channel. (b) Red channel. (c) Overlap of green and red channels. Blue arrows are added to help eyes to locate labeled GagΔP6. The scale bar is 20.0µm.

For both scenarios (III) incubated GUVs containing PI(4,5)P2 with labeled GagΔP6 before CM measurement, and (IV) addition of labeled GagΔP6 after locating one GUV containing PI(4,5)P2 under CM, green punctae can be observed on the outer leaflet of the lipid membrane of GUVs containing PI(4,5)P2 as shown in Figure 2-5. These green punctae are considered to be GagΔP6 as reported in literature [7, 22-24]. As in the case of labeled GagΔP6, the size of these green punctae is less than 1µm in diameter as the resolution limitation of CM is about 500nm. Another proposed reason

could be that GagΔP6 molecules aggregate only upon the domain of lipid membrane of GUVs where PI(4,5)P2 is abundant so that the size of PI(,45)P2 domain is the upper limit of aggregated GagΔP6 size. For scenario (IV), it is exceptionally difficult to track the GagΔP6 when it approaches a GUV and then binds to the lipid membrane of the same GUV considering that the real size of GagΔP6 is far smaller than the resolution limit of the CM. The second reason is that keeping track of the same GUV before and after addition of GagΔP6 solution to the GUVs in glucose solution without perturbation is extremely difficult. The purpose of addition of sucrose inside the GUVs is to make them heavy and reduce the effects of perturbation. To offset the osmotic pressure resulting from sucrose inside GUVs, the same concentrated glucose was used as measuring solvent in which GUVs were added. It turns out heavier sucrose can increase the immobility of GUVs in solution but cannot completely eliminate perturbation when the GagΔP6 is added to the solution. As a control, GUVs without PI(4,5)P2 were used in these experiments in parallel and no green punctae can be observed on these GUVs.

2.3 Conclusion

GagΔP6 can react with GUVs only when GUVs contain PI(4,5)P2. Therefore, the binding sites of GagΔP6 upon the outer leaflet of lipid membrane of GUVs are considered to be the regions which are abundant in PI(4,5)P2. Both MA and NC domains of HIV Gag are positively charged and can bind with negatively charged lipids PI(4,5)P2. However, the MA domain has a much higher affinity to binding with lipids PI(4,5)P2 and NC is more likely to bind with RNAs when both lipids and RNAs are present [17,25]. As

reported by other researchers, it is acidic phospholipid PI(4,5)P2 that regulates HIV Gag localization to the plasma membrane and assembly [26-29]. PI(4,5)P2 can trigger myristate exposure of MA domain to enhance the binding when it is myristoylated. It can bind with the MA domain of HIV Gag even if it is unmyristoylated. Therefore, it is reasonable to conclude that the mechanism by which HIV Gag targets the lipid membrane is through the specific interaction between the MA domain of HIV Gag and the protein recruiting factor PI(4,5)P2 contained on the GUVs.

2.4 References

[1] Minsky M. (1988). Memoir on Inventing the Confocal Scanning Microscope. J. Scanning Microsc., 10, 4, 128-138.

[2] Pawley J. (2006). Handbook of biological bonfocal microscopy, 3rd edition. Springer US.

[3] Fine A. (2007). Confocal Microscopy: Principles and Practice. CSH Protoc., 2007, 22.

[4] Peterson D.A. (2010). Confocal Microscopy. Ency. Mov. Disord., 1, 250-252.

[5] Webb R.H. (1996) Confocal optical microscopy. Rep. Prog. Phys., 59, 427-471.

[6] Nwaneshiudu A., Kuschal C., Sakamoto F.H., Anderson R.R., Schwarzenberger K., and Young R.C. (2012). Introduction to Confocal Microscopy. J. Investig. Dermatol., 132, 12, e3.

[7] Gui D., Gupta S., Xu J., Zandi R., Gill S., Huang I.C., Rao A.L., and Mohideen U. (2015). J. Biol. Phys., 41, 2135-149.

[8] Walde P., Cosentino K., Engel H., Stano P. (2010) Giant vesicles: preparations and applications. Chembiochem., 11, 7, 848-865.

[9] Pereno V., Carugo D., Bau L., Sezgin E., de la Serna J.B., EggelingC., and Stride E. (2017). Electroformation of Giant Unilamellar Vesicles on Stainless Steel Electrodes. ACS Omega, 2, 3, 994–1002.

[10] Moscho A., Orwar O., Chiu D.T., Modi B.P., and Zare R.N. (1996). Rapid preparation of giant unilamellar vesicles. PNAS, 93, 21, 11443-11447.

[11] Wesołowska O., Michalak K., Maniewska J., and Hendrich A.B. (2009). Giant unilamellar vesicles - a perfect tool to visualize phase separation and lipid rafts in model systems. Acta Biochim Pol., 56, 33-39.

[12] Hub H.H., Zimmermann U., Ringsdorf H. (1982). Preparation of large unilamellar vesicles. FEBS Letters, 140, 2, 254-256.

[13] Hopea M.J., Ballya M.B., Webbb G., and Cullis P.R. (1985). Production of large unilamellar vesicles by a rapid extrusion procedure. Characterization of size distribution, trapped volume and ability to maintain a membrane potential. Biochim. Biophys. Acta., 812, 1, 55-65.

[14] Nayar R., Hope, M.J., and Cullis P.R. (1989). Generation of large unilamellar vesicles from long-chain saturated phosphatidylcholines by extrusion technique. Biochim. Biophys. Acta., 986, 2, 200-206.

[15] Mayer L.D., Hope M.J., and Cullis P.R. (1985). Vesicles of variable sizes produced by a rapid extrusion procedure. Biochim. Biophys. Acta., 858, 1, 161-168.

[16] Frisken B.J., Asman C., and Patty P. J. (2000). Studies of Vesicle Extrusion. Langmuir, 16, 3, 928-933.

[17] Rein A, Datta S.A., Jones C.P., and Musier-Forsyth K. (2011). Diverse interactions of retroviral Gag proteins with RNAs.Trends Biochem. Sci., 36, 7, 373-380.

[18] Ganser-Pornillos B.K., Yeager M., and Pornillos O. (2012). Assembly and architecture of HIV. Viral Molecular Machines, Adv. Exp. Med. Biol., 726, 441-465.

[19] Ganser-Pornillos B.K., Yeager M., and Sundquist W.I. (2008). The structural biology of HIV assembly. Curr. Opin. Struct. Biol., 18, 2, 203-217.

[20] Briggs J.A.G., Simon M.N., Gross I., Kräusslich H.G., Fuller S.D., Vogt V.M., and Johnson M.C. (2004). The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol., 11, 672–675.

[21] Datta S.A.K., Curtis J.E., Ratcliff W., Clark P.K., Crist R.M., Lebowitz J., Krueger S., Rein A. (2007). Conformation of the HIV-1 Gag protein in solution. J. Mol. Biol., 365, 812-824.

[22] Carlson L.A., and Hurley J.H.. (2012). In vitro reconstitution of the ordered assembly of the endosomal sorting complex required for transport at membrane-bound HIV-1 Gag clusters. PNAS, 36, 7, 373-380.

[23] Carlson L.A., Bai Y., Keane S.C., Doudna J.A., and Hurley J.H.. (2016). Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. eLIFE, 5, e14663.

[24] Keller H., Kräusslich H.G., and Schwille P. (2012). Multimerizable HIV Gag derivative binds to the liquid-disordered phase in model membranes. Cell. Microbiol., 15, 2, 237-247.

[25] Shkriabai N., Datta S.A., Zhao Z., Hess S., Rein A., and Kvaratskhelia M. (2006) Interactions of HIV-1 Gag with assembly cofactors. Biochemistry, 45, 13, 4077-4083.

[26] Ono A., Ablan S.D., Lockett S.J., Nagashima K., and Freed E.O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. PNAS, 101, 41, 14889-14894.

[27] Saad J.S., Miller J., Tai J., Kim A., Ghanam R.H., and Summers M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. PNAS, 103, 30, 11364-11369.

[28] Chukkapalli V., Hogue I.B., Boyko V., Hu W.S., and Ono A. (2008). Interaction between the human immunodeficiency virus type 1 Gag matrix domain and phosphatidylinositol-(4,5)-bisphosphate is essential for efficient gag membrane binding. J. Virol., 82, 5, 2405-2417.

[29] Chukkapalli V., and Ono A.. Molecular determinants that regulate plasma membrane association of HIV-1 Gag. J. Mol. Biol., 410, 512-524 (2011).

Chapter 3

Morphology and statistics of GagΔP6 studied by atomic force microscopy

3.1 Introduction to atomic force microscopy

Atomic force microscopy (AFM) was first invented in 1986 by Binning, Quate and Gerber with a lateral resolution of 3nm and a vertical resolution less than 1Å in air [1-2]. AFM has emerged as a powerful tool to investigate topographic, elastic, electrical, and magnetic properties of different samples in a number of scientific disciplines, including physics, chemistry, biology, and materials science [3]. The force between the tip of the AFM probe and the sample can be as small as 10pN [4]. AFM's atomic scale resolution and sensitivity of very small forces demonstrate that it is a perfect tool to investigate the morphology of soft biological samples on the scale of nanometers [5-9]. AFMs have three primary modes of operation:(1) Contact Mode, also called static mode, opposing to the other dynamic modes, (2) Tapping Mode, or intermittent contact mode, (3) and Non-contact Mode. The three major applications of AFM are measurements of forces, imaging, and matter manipulation. For force measurements, there are a lot of applications, including Coulomb force, Van der Waals force, bridge force, capillary force, depletion force, double-layer force, hydration force, hydrophobic force, solvation force, and steric force etc. The interaction between the tip of the AFM probe and the sample can be measured as a function of their mutual distance, the so-called force-distance curve [10-11]. This force distance curve can be used to explore characteristics of enormous properties of a variety of materials through the measurement of stiffness, conductivity, and adhesion strength [12-17]. For imaging, the interaction of the tip of AFM probe and

the sample surface can be used to restructure the three-dimensional topographic profile of a sample surface. It is based on the motion of the tip with respect to the sample surface. For manipulation of matter, the force between the tip of AFM probe and the sample can be utilized to achieve specific purposes, such as moving molecules into a specific pattern and stimulating cells. AFM has the flexibility to either operate in air or in liquid. To achieve high resolution, AFM can also be used in vacuum and at low temperature [18-20].

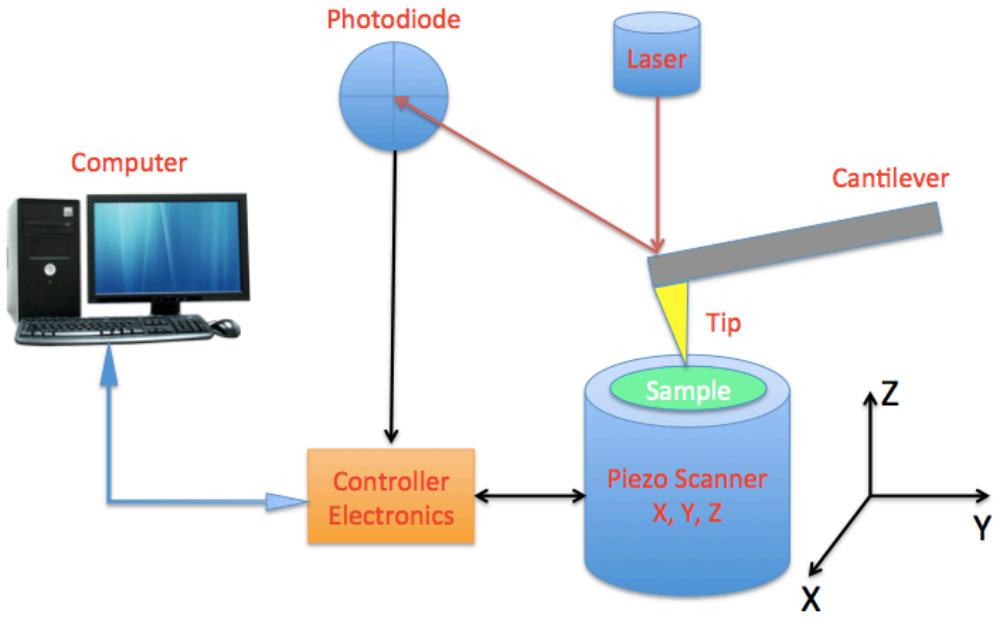

Figure 3-1: Schematic representation of Atomic Force Microscopy (AFM)

The AFM used in the experiments here is a multimode AFM with Nanoscope IIIa controller (Veeco Metrology, Santa Barbara, CA, US). We also used the peak force mode on dimension icon AFM (Bruker, Camarillo, CA, US). The schematic representation of

the AFM operation principle is shown in Figure 3-1 [21]. The sample is deposited on a substrate with atomically flat surface. The substrate is attached to an AFM metal specimen disc (TED PELLA, Redding, CA, US) mounted on a piezoelectric stage through magnetic force. The piezoelectric scanner can move the sample relative to the stationary tip vertically with high precision. AFM probes used were silicon nitride cantilevers with a sharp silicon tip at the end of the cantilever. The geometry of the tip is crucial to the AFM resolution. In particular the AFM probe used is MSNL (Bruker, Camarillo, CA, USA). MSNL has six cantilevers, with cantilever "A" on one side of the probe, and cantilevers "B", "C", "D", "E", and "F" on the other side of the probe. In the experiments, cantilever E was used. It is 140μm long and 18μm wide with a spring constant of 0.1N/m and a nominal tip radius of 2nm. It turns out the measured tip size for the cantilever E of the MSNL is around 6nm still too large to yield the correct size of GagΔP6. We also used a much sharper tip HI'RES-C19/CR-AU (MikroMasch USA, Watsonville, CA, USA). The probe HI'RES-C19/CR-AU is 125μm long and 22.5μm wide with a spring constant of 0.5N/m and a nominal tip radius less than 1nm. The image of the cantilever E and the tip of the AFM probe MSNL is shown in Figure 3-2. A thin laser beam is incident on the back of the cantilever and reflected to a mirror and then reflected upon a photodiode. The photodiode transfers the signal to the controller electronics controlled by a computer for feedback and analysis. During imaging, the tip scans horizontally and the piezo scanner moves vertically to maintain a relatively constant distance between the tip and the sample surface in contact mode, constant oscillating amplitude in tapping mode, or constant force threshold in peak force mode.

The vertical displacement of the piezo scanner is recorded by a computer to reconstruct a three dimensional topographic image of the surface of the sample.

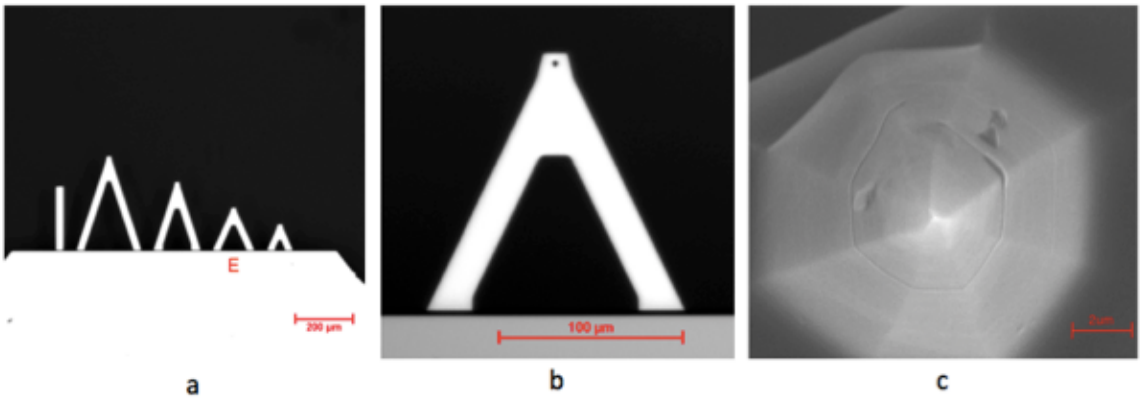

Figure 3-2 Image of the cantilever E and the tip of AFM probe MSNL. (a) The optical microscope image of AFM probe MSNL with cantilever E highlighted. The scale is 200µm. (b) The optical microscope image of the cantilever E with a scale of 100µm. (c) The SEM image of the tip on the cantilever E with a scale of 2µm.

In general, biological samples are really soft and delicate so either tapping mode or peak force mode is preferred. In contrast to contact mode, tapping mode has advantages including high lateral resolution, smaller forces and less damage to soft samples. Peak force mode has similar advantages as tapping mode with precise control of the maximum force exerted on the sample. For AFM operation in tapping mode, a tip attached to the end of an oscillating cantilever scans across the sample surface. The cantilever is driven by a voltage to oscillate around its resonant frequency with the amplitude on nanometer scale. Operation can take place both in air and in fluid. To have more profound insight in understanding the binding mechanisms of HIV Gag with

various RNAs and membrane lipids, experiments were designed by operating AFM in tapping mode in fluid. In fluid, the cantilever may not oscillate at its resonant frequency, which is the case in our experiments. The tip E of the probe MSNL has a resonant frequency of ~36.41kHz in air while it resonates at ~9.36kHz in solution. For sharper tip HI'RES-C19/CR-AU, it has a nominal resonant frequency of 65kHz in air while it resonates at ~32.36kHz in liquid.

The piezoelectric scanner responds to an applied voltage to either expand or contract. The ratio of piezo movement to piezo voltage is used to measure this response that is called sensitivity. Sensitivity usually is influenced by scan size due to piezo nonlinearities. This nonlinearity leads to the forward and reverse scanning directions behaving differently and results in hysteresis. Hysteresis has to be properly calibrated to eliminate feature distortion in AFM images. Another issue that is commonly associated with scan size is bow. Bow occurs because the piezo scanner is attached at one end while moving the sample at the free end. This leads the sample end to deviate from a horizontal plane. The bow effect increases with the scan size and can be eliminated from captured images through a software technique. Both scan size and scan rate affect imaging quality. Lowering scan rate can improve the quality of images while drifting of images plays a more and more important role. Therefore, we in general set the initial scan size to 0.5~1um and scan rate to 1~2Hz for our experiments. Besides the scan size and scan rate, there are three other primary feedback parameters during real time operation that need to be adjusted to capture high quality images in tapping mode: setpoint, integral gain and proportional gain. Setpoint is a parameter to tell the feedback loop what amplitude it

needs to maintain during scanning in tapping mode. It is automatically acquired after measurement of the resonant frequency. To get good scanned images, the setpoint is adjusted such that the minimum force is applied during imaging while still maintaining a stable engagement of the tip on the sample surface. The amount of integrated error signal controlled by the integral gain and the amount of proportional error signal controlled by the proportional gain are used in feedback calculation. Initially, integral gain and proportional gain are set to 0.4 and 0.6 in tapping mode, respectively. The higher the integral gain is set, the better the tip will image the sample topography, and the higher the quality of the image. However, noise will be introduced if the integral gain is set too high. Proportional gain is typically set to 30% to 100% higher than the integral gain. All in all, setpoint, integral gain and proportional gain have to be adjusted collectively to make sure the captured AFM images are of high quality. In peak force mode, proper scan size and scan rate also have to be chosen. However, instead of adjusting setpoint, integral gain and proportional gain, feedback gain and peak force setpoint need to be adjusted. Before engaging, the initial values for feedback gain and peak force setpoint can be 10 and 10pN, respectively. Similar to tapping mode, simultaneous adjustment of feedback gain and peak force setpoint is required to obtain high quality images in peak force mode. As reported in other studies, the limiting factors to gaining high resolution images of soft biological samples in a fluid environment are the force between the tip and the sample as well as the sharpness of the tip. It is independent of whether the imaging mode is tapping mode or peak force mode [4]. If the height of the sample is not too high, we can also

adjust the z limit to a smaller value rather than the default value to achieve a better resolution in vertical direction.

3.2 Experimental procedures

3.2.1 Sample preparation

The original HIV GagΔP6 is 30μM in the buffer containing 20 mM Tris- HCl, pH 7.5, 0.5 M NaCl, 10 % (v/v) Glycerol, 5 mM DTT. The sample concentration is too high to observe individual GagΔP6 molecules. To dilute HIV GagΔP6 to 0.5μM properly for AFM imaging in liquid, we used the buffer whose recipe is given in Table 3-1. HEPES is 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, TCEP is Tris(2-carboxyethyl) phosphine hydrochloride (both HEPES and TCEP from Thermo Fisher Scientific, Carlsbad, CA, USA) and βME is β-Mercaptoethanol (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany).

Table 3-1 Recipes of buffer

Material	Concentration	Volume added	Final concentration
HEPES	1M (PH 7.5)	20μL	20mM
$MgCl_2$	0.1M	10μL	1mM
TCEP	1mM	10μL	10μM
βME	1M	5μL	5mM
NaCl	1M	50μL	50mM
DEPC-H_2O		905μL	
Total		1mL	

The original ΨRNA and TARpolyA RNA are in the same buffer as in Table 3-1 at the concentration of 74.18μM and 119μM, respectively. Both RNAs need to refold before using. After applying refolding protocol, 30μM RNAs are diluted to 0.5μM using the same buffer as in the Table 3-1 before AFM imaging as a control experiment. The protocol of refolding RNA is as follows. First, 22.2μL 74.18μM ΨRNA (or 13.8uL 119μM TARpolyA RNA) is added into a clean vial. Next, 2.5μL 1M(PH 7.5) HEPES is added into the vial. Then, 20.3μL (or 28.7μL for TARpolyA RNA) DEPC-H_2O is added. The temperature of the mixture is raised to 80 ^{0}C for 2 minutes and then at 60 ^{0}C for another 2 minutes using a water bath. Finally, 5μL 0.1M $MgCl_2$ is added into the vial. Next, the mixture is kept at 37 ^{0}C for 5 minutes followed by ice for 30 minutes. This can be used immediately or stored at 4 ^{0}C for use up to a week.

The substrate used throughout all AFM imaging experiments is mica which is atomically smooth with roughness of 0.1~0.2nm. We also tried other substrates, such as gold coated mica, platinum coated mica, and silicon wafer. The alternative substrates we tried have an advantage of electric neutrality over mica. However, they either have a bad adsorption of samples or the metal coating will become loose after a few hours of AFM imaging in liquid. Clean mica surface is negatively charged with charge density $\sigma = -0.33C/m^2$ in air and $\sigma = -2.5mC/m^2$ in liquid [22-24]. Therefore, freshly cleaved raw mica is a perfect substrate for AFM imaging of HIV GagΔP6 because it is positively charged. However, both ΨRNA and TARpolyA RNA are negatively charged so that they cannot be observed directly on the raw mica. To overcome this repulsion between mica and RNAs, the mica had to be chemically treated to make its surface become positively

charged. There are usually two general approaches: using divalent cations as a bridge and chemical modification of mica [25-26]. The first major method is treatment of mica by divalent cations, including Mg^{2+}, Co^{2+}, Ni^{2+}, or Zn^{2+} [27-32]. The major feature of the divalent cation assisted technique to adsorb RNAs is the simplicity of the preparation of RNA samples. The potential drawback of the divalent cation assisted technique is that it is not clear if the introduced divalent cations will change the RNA structure through reaction with RNAs. Divalent cation Mg^{2+} was applied for both ΨRNA and TARpolyA RNA. However, it did not work very well. The second primary method is chemical modification of mica using polylysine, 3-aminopropyltriethoxy silane (APTES), or 1-(3-aminopropyl) silatrane (APS) [25, 33-34]. Polylysine did not work very well in the experiments and APS is not commercially available. Therefore, APTES was used in all the experiments to chemically modify the mica surface to make it become positively charged. The primary feature of the chemical modification of the mica method is that RNAs prepared on APTES treated mica are stable over a period of a few months without deterioration. While the disadvantage of chemical modification of mica method is that hydrolysis of APTES molecules occurs during real time AFM imaging in liquid, which leads to aggregation of APTES molecules [35]. The procedure of the preparation of APTES treated mica is as follows [25]. First, use double-sided tape to stick a 10mm in diameter mica (Ted Pella Inc., Redding, CA, USA) upon an AFM metal specimen disc with a diameter of 15mm. Second, use Scotch tape to cleave mica until the mica surface is complete and flat. Then, add 100μL APTES into a small plastic petri dish and place it at the bottom of a desiccator with a plastic net onto which the freshly cleaved mica is put.

Keep pumping to allow APTES to evaporate for 30 minutes. APTES treated mica is ready to use immediately or can be stored in a covered petri dish for later use within one month.

3.2.2 Calibration of AFM

Calibration of AFM probes had to be done before imaging HIV GagΔP6, RNAs, lipid molecules and their mixed complexes. Their sizes are either comparable to or smaller than the size of the tip. The comparable tip size will lead to feature broadening, which is one common type of widely known convolution effect [36]. The major factors with respect to the feature broadening are the pyrimadal geometry and curvature radius of the tip. The tips used have high aspect ratio and small tip radius. However, the tip size still had to be taken into account when analyzing the sizes of HIV GagΔP6, RNAs, lipids and their mixed complexes. Therefore, 2nm gold spheres were used to calibrate AFM probes due to the comparability of the heights of HIV GagΔP6 and two RNAs. Because the measured sample size also depends on the height of the sample, the actual tip size is given by equation (1) (Appendix A):

$$\lambda = L - 1.46D \qquad (1)$$

Where λ is the actual tip size, L is the measured size of the sample, and D is the height of the sample. The effective tip size t for any sample is the difference between measured size of the sample and the height of the sample, as given by equation (2)(please see Appendix A for more details):

$$t = L - D = \lambda + 0.46D \qquad (2)$$

For the 2nm gold spheres, after fitting to a normal distribution, the mean measured size (or diameter) $L = 7.56 \pm 0.09nm$, the mean height $D = 2.10 \pm 0.02nm$ which is the actual diameter of the 2nm gold particle, as shown in Figure 3-3. The total number of samples was 419 and the experiment was repeated twice. Therefore, the actual tip size $\lambda = 4.5nm$ according to the equation (1). The effective tip size for HIV GagΔP6 and other GagΔP6 complexes, including GagΔP6-ΨRNA, GagΔP6-PI(4,5)P2, PI(4,5)P2-ΨRNA-GagΔP6, is $t_{Gag} = 5.4nm$ given that the measured height of HIV GagΔP6 is 1.9nm according to equation (2). Similarly, the effective tip size for ΨRNA and TARpolyA RNA is $t_{RNA} = 5.0nm$ given that the measured height of both RNAs is 1.1nm.

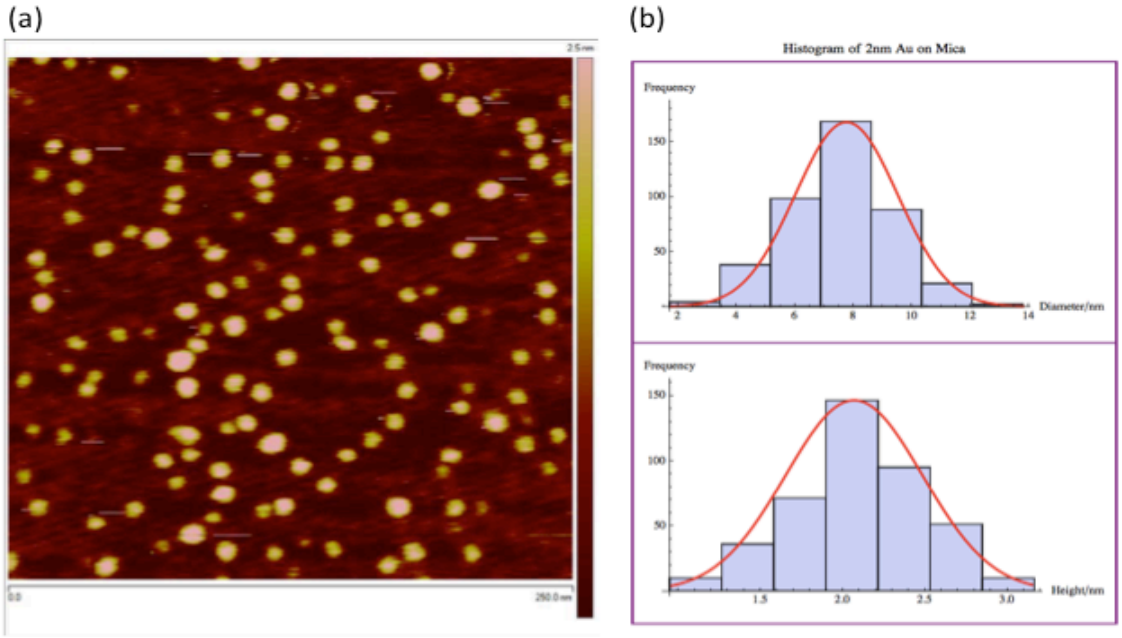

Figure 3-3 AFM tip calibration with 2nm Au Sphere. (a) AFM image of 2nm Au sphere on mica in tapping mode in liquid, the scan size is 250nm×250nm. (b) Histogram of 2nm Au sphere for the measured size (top) and the measured height (bottom). Shown in red are normal distribution fits to the peaks. The total number of samples was 419.

3.2.3 Setup of AFM

The procedure of setup of the multimode AFM Nanoscope IIIa operation in tapping mode in liquid is simple and given as following [37]. First, place a freshly cleaved mica (or APTES treated mica for RNAs and lipids) that is mounted on the AFM metal specimen disc by double-sided tape on the piezo scanner. Second, put a 30μL drop of following sample solution on the mica: (I) 0.5μM ΨRNA, (II) 0.5μM TARpolyA RNA, (III) 0.5μM GagΔP6, (IV) mixture of PI(4,5)P2-DPhPC-POPC (0.5μM : 5μM : 5μM) complex, (V) mixture of GagΔP6-ΨRNA (0.5μM : 0.5μM) complex, (VI) mixture of GagΔP6-TARpolyA RNA (0.5μM : 0.5μM) complex, (VII) mixture of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) complex, (VIII) and mixture of PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM) complex. Then, mount the cantilever probe upon the fluid cell and adjust it into proper position. Make sure the tip is completely immersed in the solution, which is critical for laser alignment when operating AFM in liquid. The laser signal will be lost if it is initially aligned in air and then submerged into the solution. Care is taken to fully immerse the cantilever probe to eliminate any optical interfaces. Align the laser by adjusting two knobs on top of the multimode scanner and the angle of the reflection mirror. Make sure the signal “SUM” is as large as possible, at least 4V for imaging in liquid. Adjust “VERT” and “HORZ” for vertical and horizontal position of the laser signal to make sure both of them are around 0V. For tapping mode, auto tune has to be implemented for the tip to get the correct resonant frequency and drive amplitude. For the sharp tip HI’RES-C19/CR-AU, it has a nominal resonant frequency around 65kHz in air and ~32.36kHz in liquid, which was obtained by auto tune. After the

auto tune is complete, proceed to the next step if drive frequency is around 32.36kHz. Before engaging, we need to check the setup of initial parameters, such as z limit, samples/line, scan size, scan rate, integral gain, proportional gain and amplitude setpoint. Set z limit to 1μm rather than the default value of 3.57μm. Reducing the z limit can increase the resolution in the vertical direction. Set samples/line to 512 rather than 256 to enhance the resolution in the horizontal direction. Set scan size to 1μm and scan rate to 2Hz at first. Initially, integral gain, proportional gain and amplitude setpoint are set to 0.4, 0.6 and 0.36V, respectively. Move the tip to approach the sample and then engage. When imaging adjust the amplitude setpoint such that the trace and retrace lines are matching. The force exerted on the sample should be as small as possible. Integral gain can be increased to enhance the signal to improve the resolution. However, the noise will increase if the integral gain is set too high. Proportional gain is usually set to 30% to 100% higher than the integral gain. Therefore, all the parameters aforementioned have to be adjusted collectively to achieve the best resolution.

For the dimension icon AFM, the procedure of setup is more complicated but the principle is very similar. There is no need to auto tune for the peak force mode. Instead of tuning integral gain, proportional gain and amplitude setpoint, feedback gain as well as peak force setpoint should be adjusted to acquire images with high resolution. In both tapping mode and peak force mode, the force exerted on samples should be as small as possible to prevent soft samples from being damaged.

3.2.4 AFM measurement of structure and statistical analysis

3.2.4.1 ΨRNA

For scenario (I) 0.5μM ΨRNA on APTES treated mica(+), the AFM image is shown in Figure 3-4. The motivation of measuring ΨRNA is to serve as a control before adding it into GagΔP6. In Figure 3-4(a), AFM imaging of ΨRNA shows that most of ΨRNA molecules seem to have inverted "L" shape. As shown in Figure 3-4(b), the most

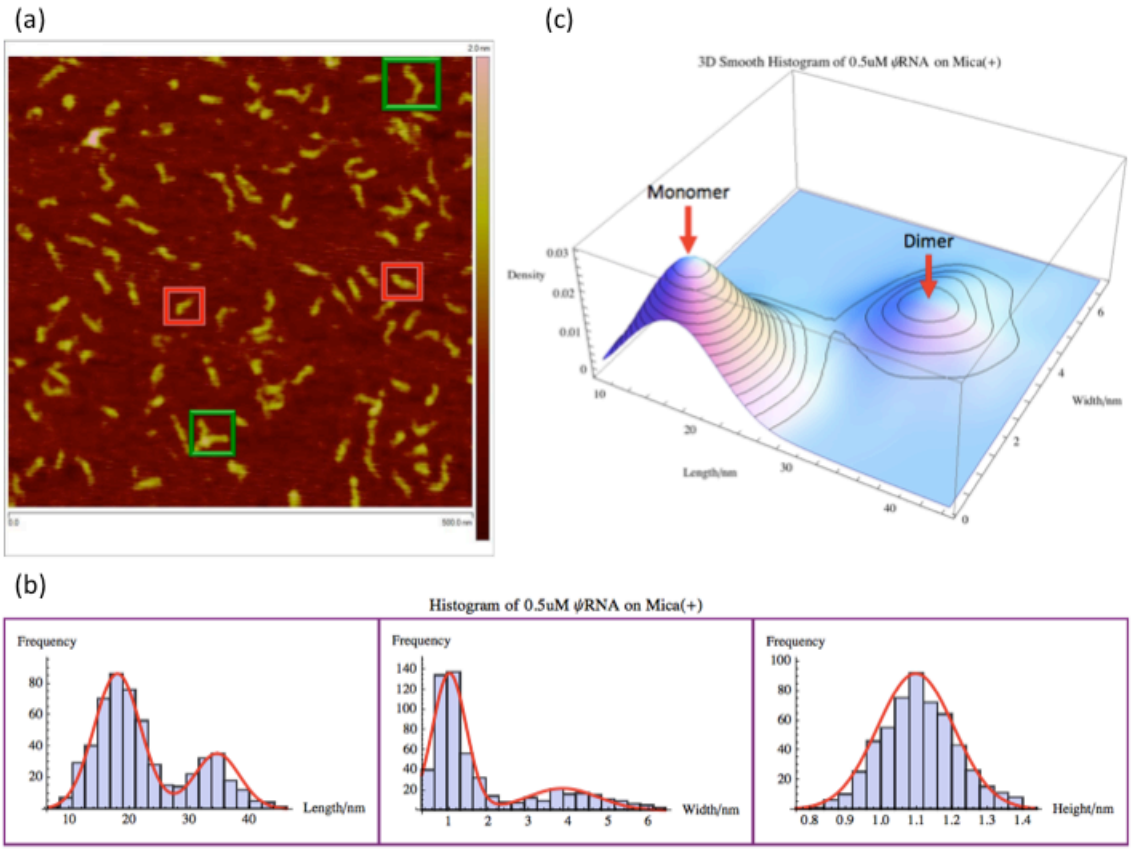

Figure 3-4 0.5μM ΨRNA on positively charged mica(+). (a) AFM image with a scan size of 500nm×500nm. A few ΨRNAs are boxed: monomer (red) and dimer (green). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer and dimer are indicated by red arrows. The total number of samples was 551.

frequent height is 1.1nm. The total number of samples was 551 and the experiment was repeated twice. This height is shorter than the height of double-stranded RNA found by Ares et al. and by Andersen et al. but higher than that found by Bussiek et al., and comparable to that found by Hansma et al. [4, 26, 38-39]. There are two peaks for length and two peaks for width. The three dimensional smooth histogram in Figure 3-4(c) shows that the length and width of monomer are less than those of dimer. After fitting each peak to a normal distribution for length and width, the percentages of monomer and dimer are 74% and 26%, respectively, as given in Table 3-2. The length of ΨRNA monomer is

Table 3-2 Statistics of AFM measurement of 0.5μM ψRNA on mica(+)

	Monomer	Dimer
Percentage	74%	26%
Length/nm	17.9 ± 0.2	34.6 ± 0.3
Width/nm	1.01 ± 0.02	3.8 ± 0.1
Height/nm	1.10 ± 0.01	

around 18nm given that ΨRNA contains 109 nucleotides so each nucleotide is only 0.17nm, which is much less than expected. RNAs most commonly adopt either A-form or B-form DNA conformation. The rise per base pair for A-form and B-form are 0.26nm and 0.34nm, respectively [40-42]. Given that most of ΨRNA nucleotides are self-paired with each other as shown in Figure 1-5, ΨRNA should be double-stranded with 55 base pairs and the rise per base pair is 0.33nm that resembles the B-form DNA conformation. The length of ΨRNA dimer is around 35nm, which is approximately twice as long as that of ΨRNA monomer. This means ΨRNA dimer consists of two ΨRNA monomers

connecting head to head. This assumption is reasonable given that DIS is located in the SL1 loop of ΨRNA as shown in Figure 1-5. The width of ΨRNA monomer is around 1nm consistent with the height/width of ΨRNA. The width of ΨRNA dimer is roughly 4nm that is much larger than two-fold of the width of ΨRNA monomer. This is probably because the junction of two ΨRNA monomers is wider than expected. Another explanation is due to the algorithm used to compute the dimension of a sample, which is discussed in Appendix B.

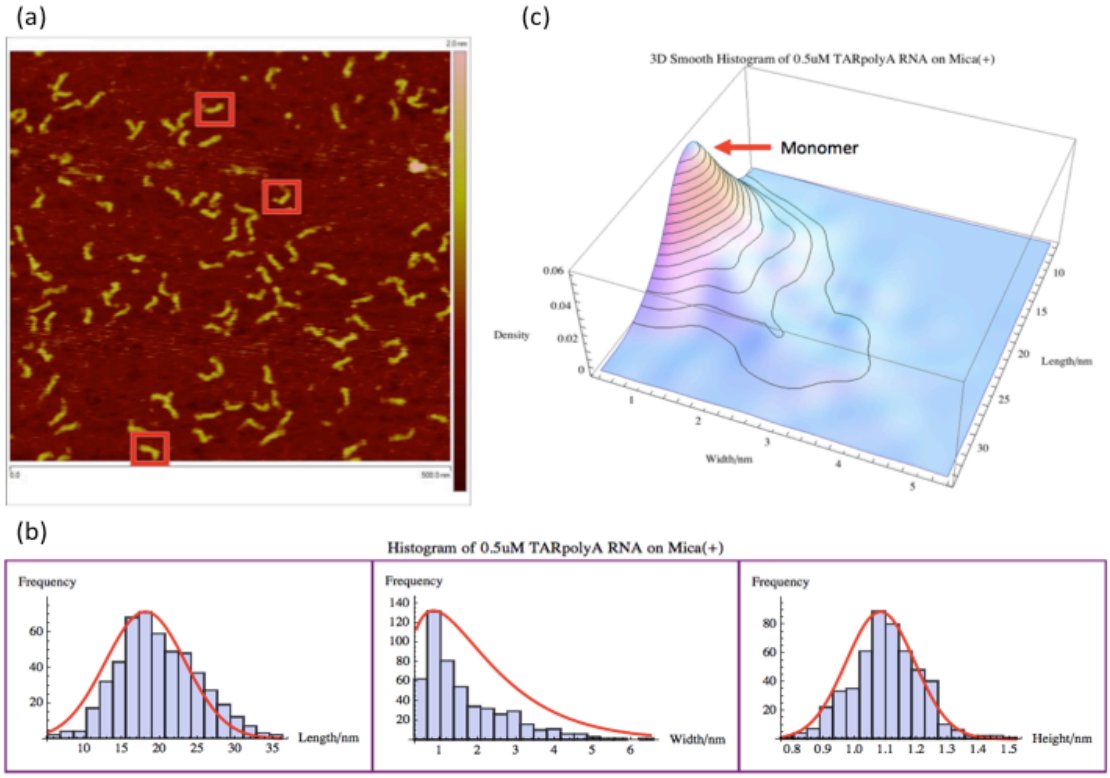

Figure 3-5 0.5μM TARpolyA RNA on positively charged mica(+). (a) AFM image with a scan size of 500nm×500nm. A few TARpolyA RNAs are boxed: monomer (red). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks for length and height. Width is fit to a gamma distribution due to its non-negativity and skewness to right. (c) Three dimensional smooth histogram, where monomer is indicated by a red arrow. The total number of samples was 504.

3.2.4.2 TARpolyA RNA

For scenario (II) 0.5μM TARpolyA RNA on APTES treated mica(+), the AFM image is shown in Figure 3-5. Similarly, the motivation of measuring TARpolyA RNA is also to serve as a control before adding it into GagΔP6. In Figure 3-5(a), AFM imaging of TARpolyA RNA shows that most of TARpolyA RNA molecules seem to have inverted "L" shape just like ΨRNA. As shown in Figure 3-5(b), the most frequent height of TARpolyA RNA is 1.1nm as well. The total number of samples was 504 and the experiment was repeated twice. This height is the same as that of ΨRNA in the AFM measurement. Unlike ΨRNA, TARpolyA RNA only has monomer as shown in Figure 3-5(b) and (c). The length of TARpolyA RNA monomer is around 17nm slightly less than that of ΨRNA monomer. This is reasonable given that ΨRNA contains 109 nucleotides while TARpolyA RNA only has 104 nucleotides. Similar to RNA, TARpolyA RNA should also be doubled-stranded as shown in Figure 1-4. It is noteworthy that the height of TARpolyA RNA peaks at 1nm with positive skewness. Therefore, we fit the height of TARpolyA RNA to gamma distribution, Gamma(α, β) with $\alpha = 1.7$ and $\beta = 1.2$, rather than normal distribution. According to gamma distribution, the mean should be equal $\alpha/\beta = 1.7/1.2 \approx 1.4nm$, which is consistent with the width peak around 1nm.

3.2.4.3 GagΔP6

For scenario (III) 0.5μM GagΔP6 on negatively charged mica(-), the AFM image is shown in Figure 3-6. The motivation of measuring GagΔP6 is to serve as a control

before addition of RNAs and lipids. In Figure 3-6(a), AFM imaging of GagΔP6 shows that most of GagΔP6 molecules have ellipsoidal shape rather than the expected rod-like

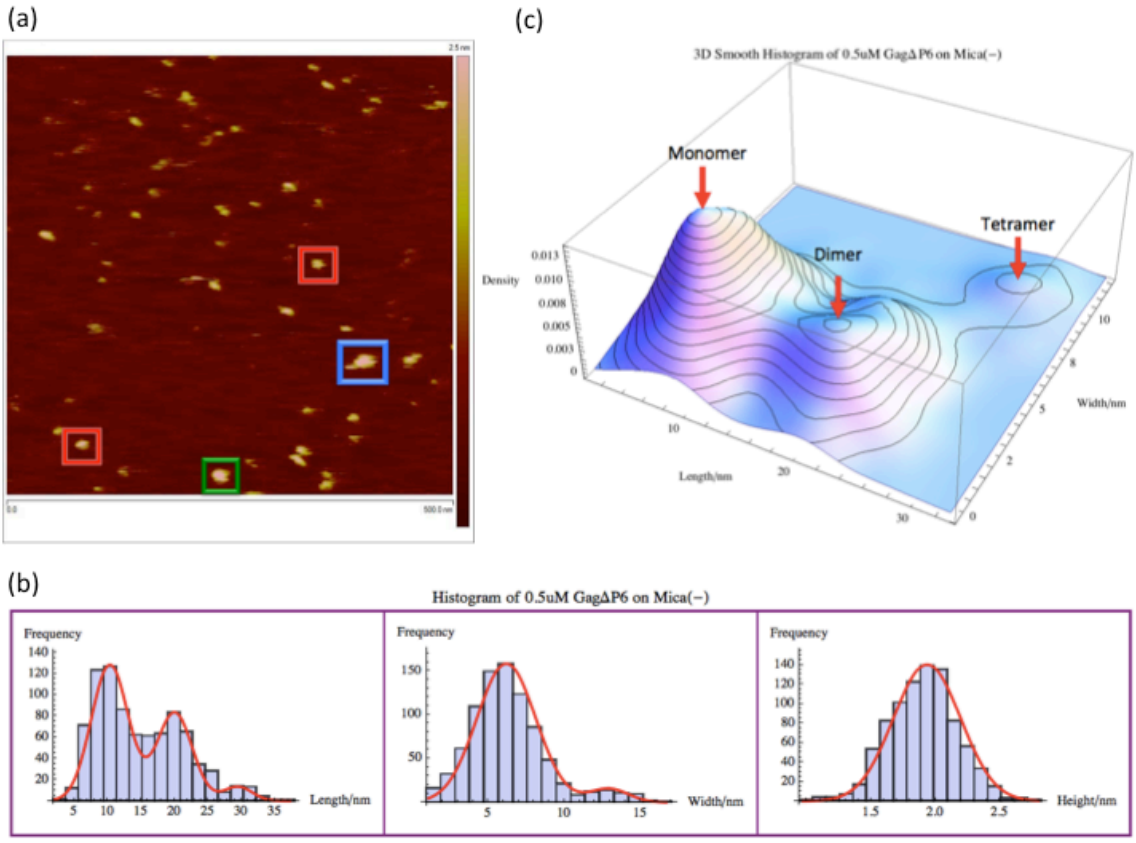

Figure 3-6 0.5μM GagΔP6 on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6s are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 858.

shape. Some studies shows HIV GagΔP6 is supposed to adopt a compact conformation such that MA and NC domains of GagΔP6 are close to each other even though both of them are positively charged [43-45]. The hydrodynamic radius, R_h, of GagΔP6 given by

three different hydrodynamic tests are 3.6nm, 3.8nm, and 4.1nm, respectively. The radius of gyraton, R_g, of GagΔP6 is best estimated to be 3.4nm by the small angle neutron scattering (SANS). The R_g of GagΔP6 when it is 25nm straight rod is supposed to be 7.2nm [45]. The average R_g of GagΔP6 in solution measured by SANS is also a monotonically increasing function of GagΔP6 concentration, with maximum of $R_g = 5nm$ at extremely high concentration, which means GagΔP6 molecules are in monomer-dimer equilibrium [44]. Our AFM results were consistent with these reports except that tetramers were also observed besides monomers and dimers. The presence of tetramers rather than trimmers was confirmed by gel electrophoresis as discussed in a later section. As shown in Figure 3-6(b), the most frequent height is 1.9nm which is consistent with the expectation that the diameter of Gag is around 2~3nm[43]. The total number of samples was 858 and the experiment was repeated twice. It might seem counterintuitive to have three peaks for length while having only two peaks for width. To have more profound insight in comprehending this phenomenon, a three dimensional smooth histogram is plotted as shown in Figure 3-4(c). It is clear that monomer and dimer have the same width so in the width histogram they are represented together and correspond to the first peak at around 6nm. The statistical analysis also confirmed this assumption the

Table 3-3 Statistics of AFM measurement of 0.5μM GagΔP6 on mica(-)

	Monomer	Dimer	Tetramer
Percentage	59%	35%	6%
Length/nm	10.3 ± 0.1	20.0 ± 0.2	29 ± 0.3
Width/nm	6.2 ± 0.1		12.9 ± 0.2
Height/nm	1.93 ± 0.01		

first peak of the width histogram is the sum of monomer and dimer as given in Table 3-3. After fitting each peak to a normal distribution for length and width, the percentages of monomer, dimer, and tetramer are 59%, 35%, and 6%, respectively. Monomer is around 10nm long and dimer is around 20nm long while both are about 6nm wide. Tetramer is around 29nm long and 13nm wide. Therefore, GagΔP6 monomer based on the shorted width should have a "C" like shape [43]. Again based on the AFM measured dimensions the model of the GagΔP6 dimer has two monomers connecting back to back through CA-CA interaction. GagΔP6 tetramer is formed by two dimers in some compact form. The proposed models of GagΔP6 monomer, dimer, and tetramer are shown in Figure 3-7.

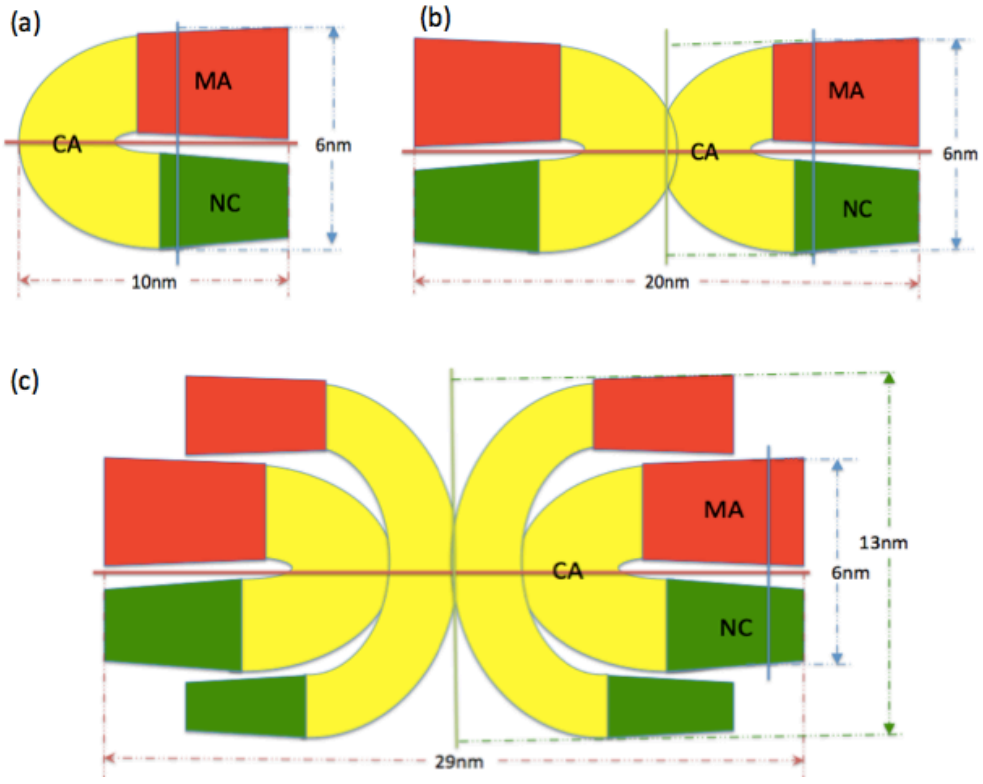

Figure 3-7 Models of GagΔP6. (a) Monomer, (b) Dimer, (c) Tetramer. MA domain is in red, CA domain is in yellow, and NC domain is in green.

2μM GagΔP6 on the negatively charged mica was also investigated. In contrast to 0.5μM GagΔP6, the percentages of GagΔP6 dimer and tetramer are higher when the concentration is 2μM. This also confirmed the conclusion that the weighted average size of GagΔP6 in solution is a monotonically increasing function of GagΔP6 concentration [44]. The later gel electrophoresis experiments also confirmed this conclusion.

3.2.4.4 PI(4,5)P2-DPhPC-POPC

For scenario (IV), a mixture of PI(4,5)P2-DPhPC-POPC (0.5μM : 5μM : 5μM) complex on APTES treated mica(+), the AFM result is shown in Figure 3-8. Our motivation is to measure PI(4,5)P2 itself as a control. But PI(4,5)P2 is much smaller than GagΔP6 given that PI(4,5)P2's mass is only about 1kDa. It is unable to use AFM to

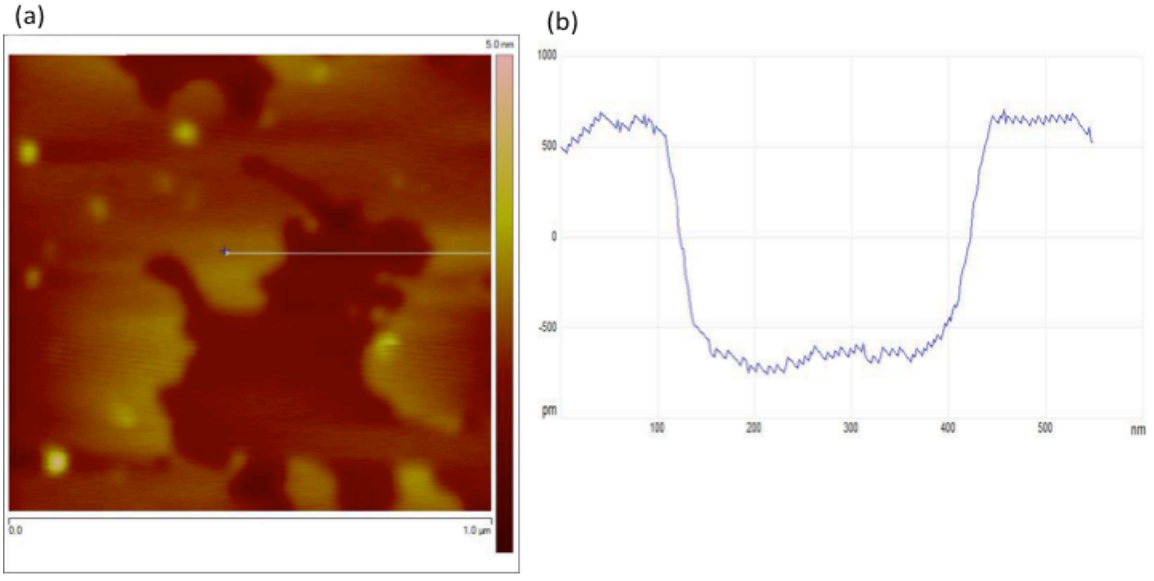

Figure 3-8 AFM image of lipid mixture of PI(4,5)P2-DPhPC-POPC (0.5μM : 5μM : 5μM) complex. (a) AFM image with a scan size of 1μm×1μm. (b) The line delineated in (a).

measure individual PI(4,5)P2 directly. So, we measured the mixture of PI(4,5)P2-DPhPC-POPC. Alternatively, lipid bilayers can also be easily formed on self-assembled monolayers (SAMs), which are most widely studied nanostructures [46-54]. In Figure 3-8(a), there are three distinguishable regions. The brightest spots correspond to PI(4,5)P2 because the concentrations of the other two lipids are 10 times higher than that of PI(4,5)P2 and the area of brightest spots is much less than the areas of another two regions. The areas of the other two regions are similar since their molar concentrations are the same and molecular weights are close to each other. At room temperature, DPhPC is in solid phase and POPC is in liquid phase. In Figure 3-8(b), the line crossed from yellow (high) domain into brown (low) domain with a thickness difference of about 1nm. Because the carbon chain of DPhPC is longer than that of POPC, the yellow domains correspond to DPhPC and the brown domains correspond to POPC. This 1nm difference in height between two different lipids is consistent with those values found by others [55-58].

3.2.4.5 GagΔP6-ΨRNA

For scenario (V), a mixture of GagΔP6-ΨRNA (0.5μM : 0.5μM) complex on negatively charged mica(-), the AFM result is shown in Figure 3-9. The motivation of measuring GagΔP6-ΨRNA complex is to investigate the effect of addition of specific ΨRNA to GagΔP6. According to current models the NC domain binds with ΨRNA and this interaction is specific and is a critical step in the formation of HIV [59]. Figure 3-9(a)

is the AFM image of GagΔP6-ΨRNA complex. As shown in Figure 3-9(b), the most frequent height of GagΔP6-ΨRNA complex is 1.9nm that is the same as the height of GagΔP6. The total number of samples was 895 and the experiment was repeated twice.

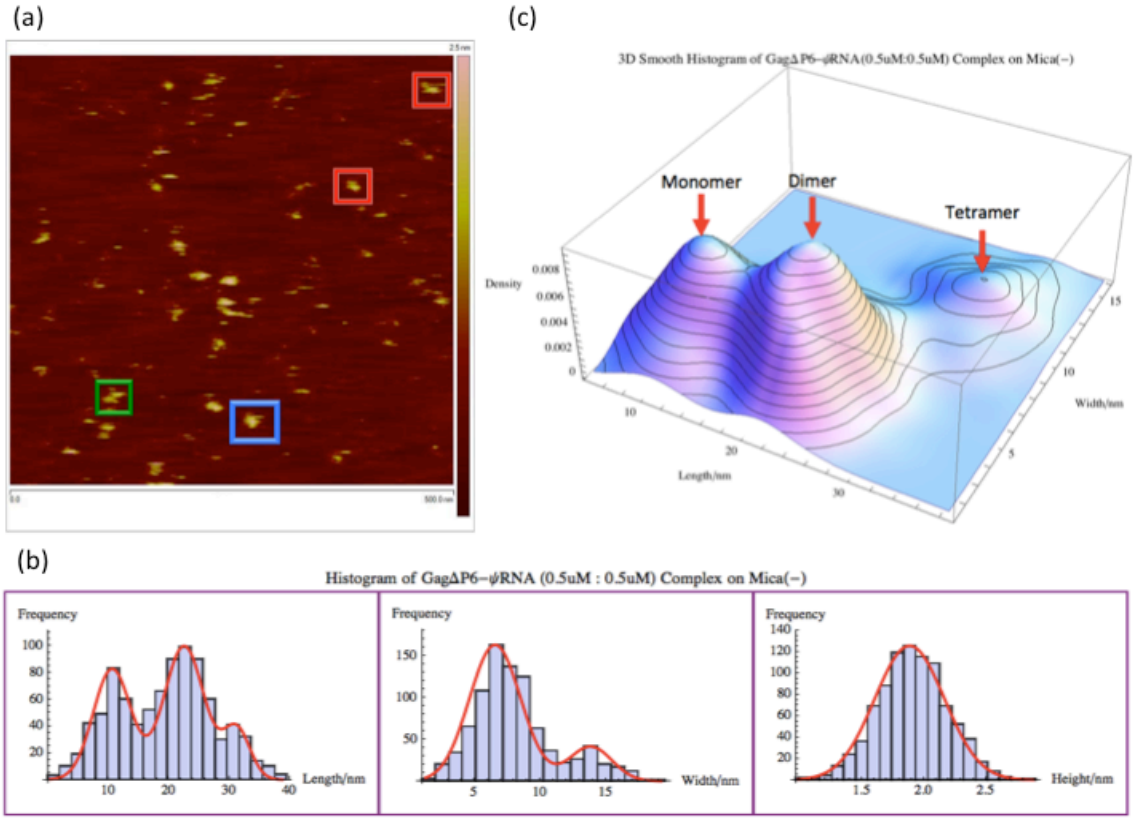

Figure 3-9 Mixture of GagΔP6-ΨRNA (0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6-ΨRNA complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 895.

This height is reasonable given that the height of GagΔP6 and ΨRNA are 1.9nm and 1.1nm, respectively. Similar to GagΔP6, GagΔP6-ΨRNA also has three peaks for length and two peaks for width. The difference is the ratio of dimer and tetramer increased

Table 3-4 Statistics of AFM measurement of GagΔP6-ψRNA (0.5μM : 0.5μM) on mica(-)			
	Monomer	Dimer	Tetramer
Percentage	35%	49%	16%
Length/nm	10.6 ± 0.2	22.4 ± 0.1	31.2 ± 0.2
Width/nm	6.8 ± 0.1		13.8 ± 0.1
Height/nm	1.90 ± 0.01		

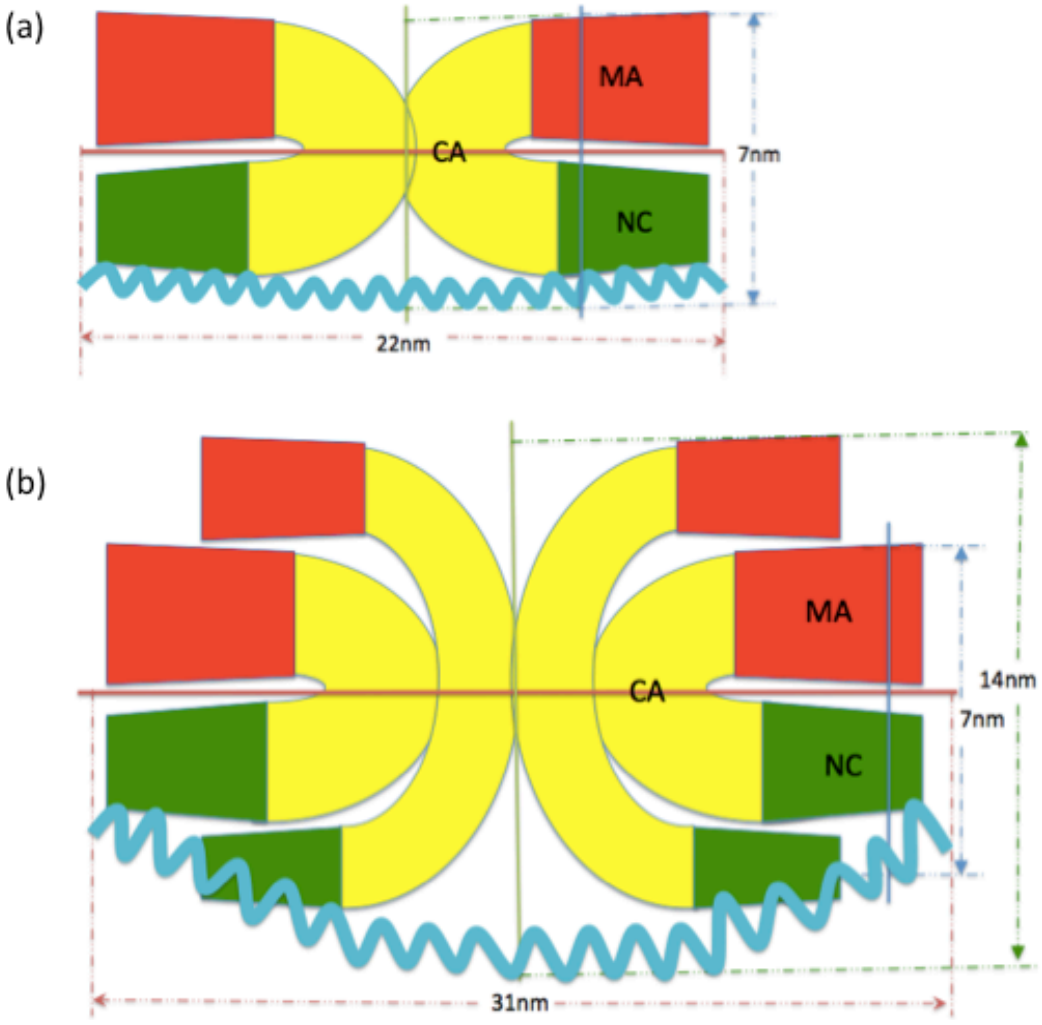

Figure 3-10 Models of GagΔP6-ΨRNA complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, and ΨRNA is in cyan.

compared to GagΔP6 itself. As shown in Figure 3-9(c), the dimer predominates over monomer as opposed to GagΔP6. After fitting each peak to a normal distribution for length and width, the statistical analysis shows the percentages of monomer, dimer, and tetramer are 35%, 49%, and 16%, respectively, as given in Table 3-4. The monomer decreased by 24%, the dimer increased by 14% and the tetramer increased by 10%. The length of monomer increased by roughly 1nm. The lengths of dimer and tetramer increased by about 2nm. The monomer and dimer still shared the width around 7nm. The width increased by about 1nm for monomer, dimer, and tetramer, which makes sense given the width of ΨRNA is around 1nm. The models for GagΔP6-ΨRNA complex dimer and tetramer are given in Figure 3-10. The increases in both length and width can be explained by the addition of ΨRNA given the size of ΨRNA in Table 3-2. The conclusion of the effect of addition of ΨRNA to GagΔP6 is that ΨRNA can bind with GagΔP6 and facilitate GagΔP6 multimerization given the increases in percentages of dimer and tetramer.

3.2.4.6 GagΔP6-TARpolyA RNA

For scenario (VI), a mixture of GagΔP6-TARpolyA RNA (0.5μM : 0.5μM) complex was measured. The motivation of measuring GagΔP6-TARpolyA RNA complex is to see the effect of addition of non-specific TARpolyA RNA to GagΔP6 and compare the difference between adding TARpolyA RNA and the case of addition of specific ΨRNA. It is unable to observe any statistical difference for GagΔP6-TARpolyA RNA

complex on negatively charged mica(-) in contrast to GagΔP6. In addition, particles observed on APTES treated mica(+) for GagΔP6-TARpolyA RNA complex resemble the shape of TARpolyA RNA, which means either TARpolyA RNA does not react with GagΔP6 or TARpolyA RNA reacts with GagΔP6 but not facilitate GagΔP6 multimerization. The experiments were repeated twice. Webb et al. reported HIV GagΔP6 can bind with both ΨRNA and TARpolyA RNA but with distinct binding mechanisms [59]. They proposed that HIV GagΔP6 binds with TARpolyA RNA through both MA and NC domains while with ΨRNA only through NC domain and leaves MA domain free to later interact with lipid membrane. Other studies also showed that HIV GagΔP6 can bind with both ΨRNA and non-Ψ RNAs but the selective binding with ΨRNA is more energetically favorable than other non-Ψ RNAs for HIV virus assembly [60-62]. The conclusion based on our data is that it is highly likely that HIV GagΔP6 interacts with ΨRNA and TARpolyA RNA through different mechanisms such that ΨRNA facilitates GagΔP6 multimerization while TARpolyA RNA does not.

3.2.4.7 GagΔP6-PI(4,5)P2

For scenario (VII), a mixture of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) complex on negatively charged mica(-), the AFM image is shown in Figure 3-11. The motivation of measuring GagΔP6-PI(4,5)P2 complex is to explore the effect of addition of lipid PI(4,5)P2 to GagΔP6. The current understanding is that both MA and NC domains of Gag can bind to PI(4,5)P2. The size is in between GagΔP6 and GagΔP6-ΨRNA. As

shown in Figure 3-11(b), the most frequent height of GagΔP6-PI(4,5)P2 complex is still 1.9nm, same as GagΔP6 and GagΔP6-ΨRNA. The total number of samples was 903 and the experiment was repeated twice. Unlike GagΔP6-ΨRNA, monomer is still predominant as shown in Figure 3-11(c). After fitting each peak to a normal distribution for length and width, the statistical analysis shows the percentage of monomer decreased by 12%, the percentages of dimer and tetramer both increased by 6% compared to the case of GagΔP6, as given in Table 3-5. The length and width of monomer, dimer,

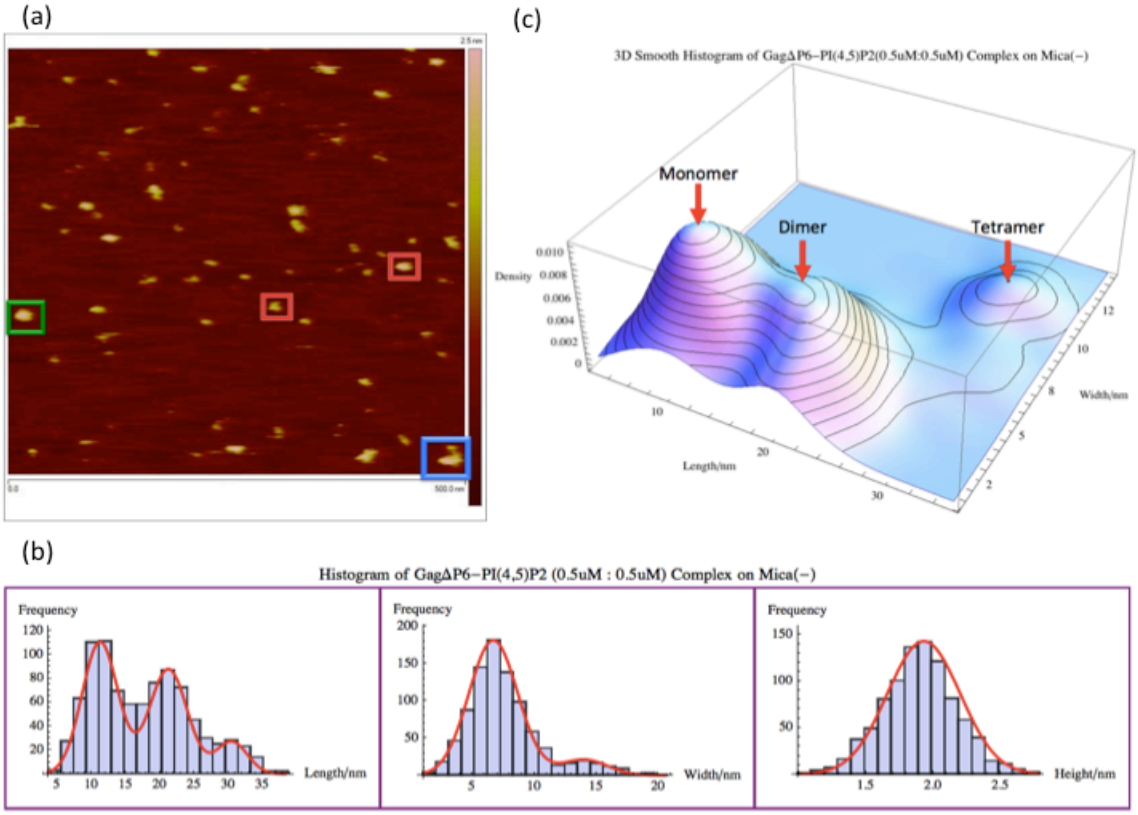

Figure 3-11 Mixture of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few GagΔP6-PI(4,5)P2 complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 903.

Table 3-5 Statistics of AFM measurement of GagΔP6-PI(4,5)P2 (0.5μM : 0.5μM) on mica(-)			
	Monomer	Dimer	Tetramer
Percentage	47%	41%	12%
Length/nm	11.2 ± 0.1	21.1 ± 0.1	30.4 ± 0.2
Width/nm	6.7 ± 0.1		14.0 ± 0.2
Height/nm	1.93 ± 0.01		

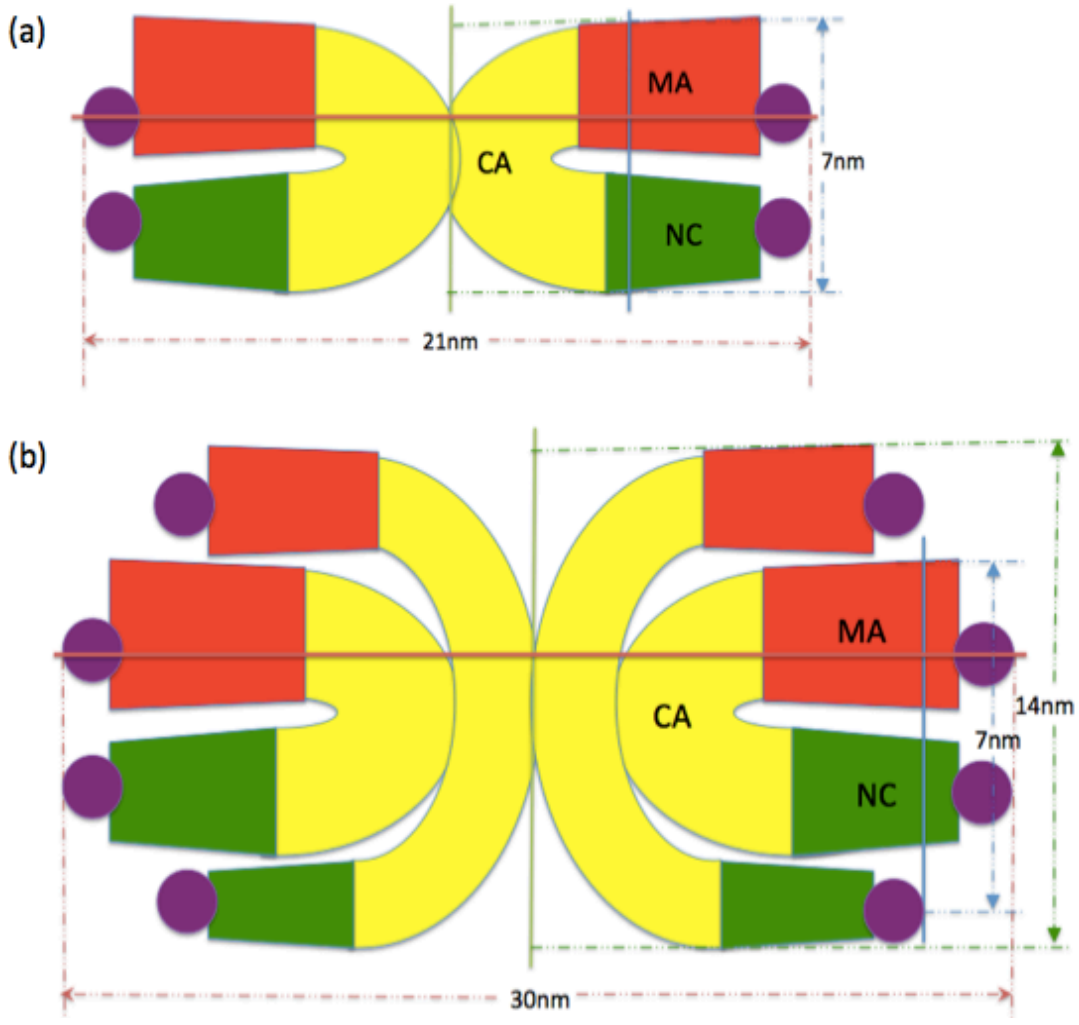

Figure 3-12 Models of GagΔP6-PI(4,5)P2 complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, and PI(4,5)P2 is in purple.

tetramer all increased by about 1nm. This is probably because of the size of PI(4,5)P2 attached to the ends of both MA and CA domains of HIV GagΔP6. The models for GagΔP6-PI(4,5)P2 complex dimer and tetramer are given in Figure 3-12. The increases in percentages of dimer and tetramer indicate that PI(4,5)P2 can bind with GagΔP6 as reported by other studies [63-69]. But the increases in percentages of dimer and tetramer are less than the case of addition of ΨRNA. Therefore, our conclusion is that PI(4,5)P2 facilitates GagΔP6 multimerization to a lesser extent compared to ΨRNA.

3.2.4.8 PI(4,5)P2-ΨRNA-GagΔP6

For scenario (VIII), a mixture of PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM) complex on negatively charged mica(-), the AFM result is shown in Figure 3-13. According to the prevailing model the lipid interacts with the MA domain of GagΔP6-ΨRNA complex leading to a conformational change [59]. In these experiments we study the effects of the inclusion of PI(4,5)P2 to the GagΔP6-ΨRNA complex. For the mixture of PI(4,5)P2-ΨRNA-GagΔP6, PI(4,5)P2 and ΨRNA were mixed first followed by addition of GagΔP6. Unlike the cases of GagΔP6, GagΔP6-ΨRNA, and GagΔP6-PI(4,5)P2, there are only two peaks for the length histogram as in Figure 3-13(b). The width histogram still has two peaks as before. To better understand the reason for missing one peak for the length histogram, a three dimensional smooth histogram is plotted as shown in Figure 3-13(c). It turns out that the dimer and tetramer have the same length at around 24nm. The total number of samples was 616 and the experiment was repeated

twice. This overlap is explained from the fact that there are only two peaks for the length histogram. The monomer and dimer have the same width around 7nm. After fitting each

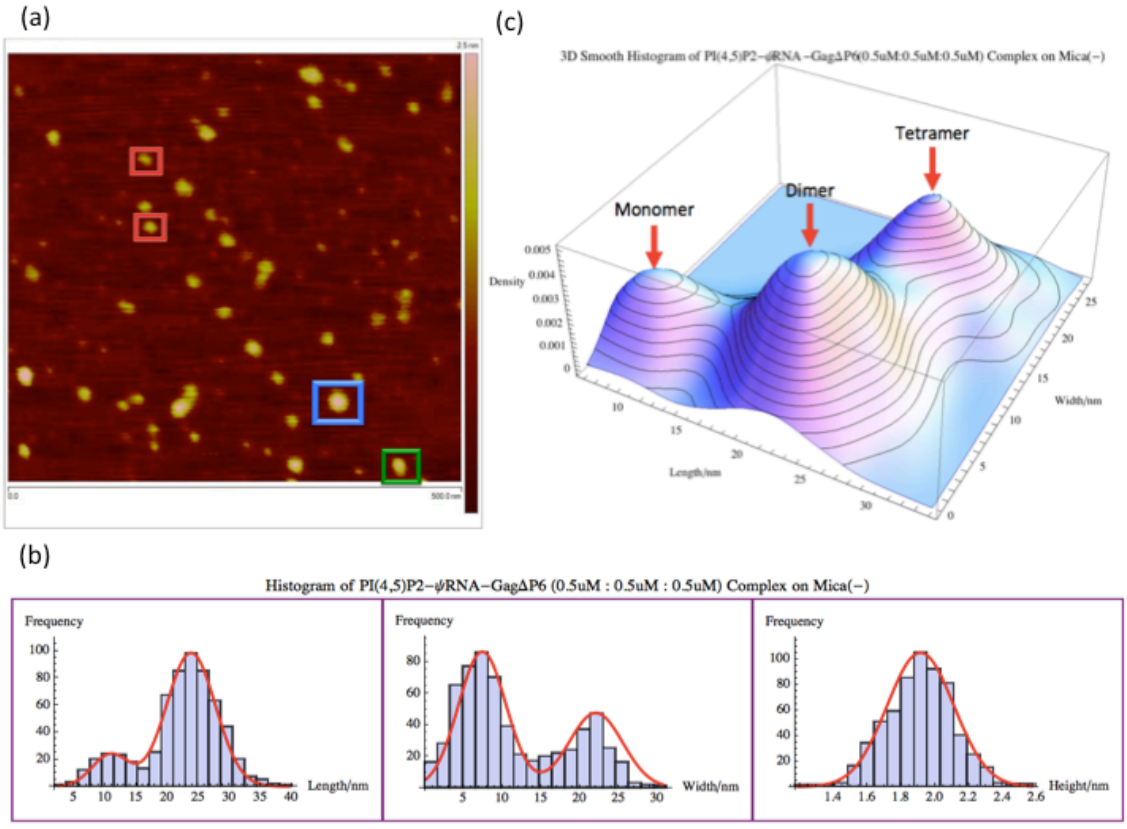

Figure 3-13 Mixture of PI(4,5)P2-ΨRNA-GagΔP6 (0.5μM : 0.5μM : 0.5μM) complex on negatively charged mica(-). (a) AFM image with a scan size of 500nm×500nm. A few PI(4,5)P2-ΨRNA-GagΔP6 complexes are boxed: monomer (red), dimer (green) and tetramer (blue). (b) Histogram for length (left), width (middle), and height (right). Shown in red are normal distribution fits to the peaks. (c) Three dimensional smooth histogram, where monomer, dimer, and tetramer are indicated by red arrows. The total number of samples was 616.

peak to a normal distribution for length and width, the distribution weights among monomer, dimer, and tetramer can be decoupled. The percentage of the first peak of the length histogram is 17% that is corresponding to the monomer. The percentage of the

second peak of the length histogram is 83% that is the sum of the dimer and tetramer. The percentage of the first peak of the width histogram is 68% that is the sum of the monomer and dimer. The percentage of the second peak of the width histogram is 32% that is the tetramer. Therefore, the percentages of monomer, dimer, and tetramer are 17%, 51%, and 32%, respectively, as given in Table 3-6. The substantial increases in the percentages of

Table 3-6 Statistics of AFM measurement of PI(4,5)P2-ψRNA-GagΔP6 (0.5μM:0.5μM:0.5μM) on mica(-)

	Monomer	Dimer	Tetramer
Percentage	17%	51%	32%
Length/nm	10.9 ± 0.3	23.8 ± 0.2	
Width/nm	7.4 ± 0.2		22.1 ± 0.2
Height/nm	1.91 ± 0.01		

dimer and tetramer indicate that both ΨRNA and PI(4,5)P2 can bind with GagΔP6 and collectively facilitate HIV GagΔP6 assembly as reported by other studies [37, 67]. Another noteworthy difference is the width of tetramer increased to 22nm rather than 13nm for GagΔP6 or 14nm for GagΔP6-ΨRNA and GagΔP6-PI(4,5)P2. This significant change of the shape indicates that some conformational changes occurred when both ΨRNA and PI(4,5)P2 are present [70-71]. The average spacing between GagΔP6 molecules in dimer and tetramer is about 7nm, which is close to the value found in other literature [72-73]. The proposed models for PI(4,5)P2-ΨRNA-GagΔP6 complex dimer and tetramer are shown in Figure 3-14. PI(4,5)P2-ΨRNA-GagΔP6 means PI(4,5)P2 and ΨRNA were mixed first followed by the addition of GagΔP6. When there are three elements, GagΔP6, ΨRNA, and PI(4,5)P2, the order of mixing them might matter. To see

if this is the case, other mixing orders were also studied, including GagΔP6-ΨRNA-PI(4,5)P2 and GagΔP6-PI(4,5)P2-ΨRNA. Regardless of the order of mixing, the size and size fractions of the monomers, dimers and tetramer were found to be the same. This is reasonable given that eventually the MA domain of GagΔP6 binds with the binding sites PI(4,5)P2 on the lipid membrane and the NC domain of GagΔP6 binds with the dimeric viral RNA in the complete immature HIV virion [74-77]. The conclusion for this case is when both ψRNA and PI(4,5)P2 are present, GagΔP6 undergoes conformational changes consistent with the model [59]. Given the order of the lipid PI(4,5)P2 or ΨRNA addition resulted in no changes to the different complex populations measured, we conclude that the proposed order of Gag interaction might not be critical to the formation of HIV.

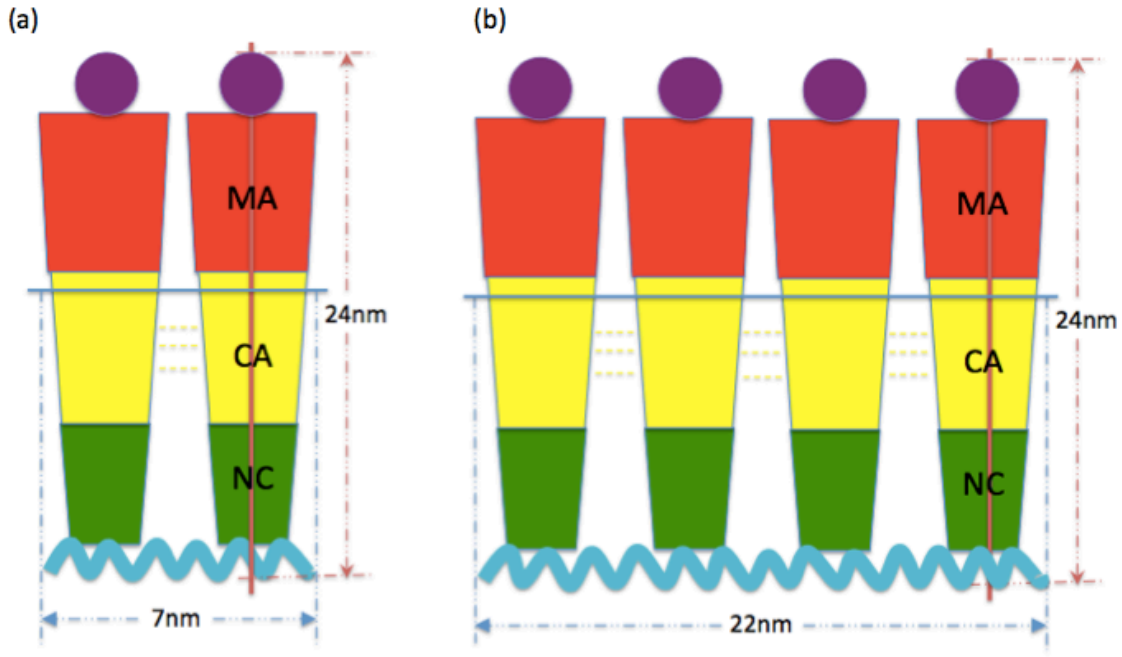

Figure 3-14 Models of PI(4,5)P2-ΨRNA-GagΔP6 complexes. (a) Dimer, (b) Tetramer. MA domain is in red, CA domain is in yellow, NC domain is in green, ΨRNA is in cyan, and PI(4,5)P2 is in purple.

3.3 Gel electrophoresis

3.3.1 Introduction to gel electrophoresis

Gel electrophoresis is a widely used technique for analysis of macromolecules and their fragments, including DNAs, RNAs, and proteins, by separating them based on their sizes and charges [78-85]. The principle of gel electrophoresis is that charged macromolecules travel through a porous substance under an electric field. The electric field generated from two opposite electrodes will generate electric force on charged molecules to move through the gel. Smaller molecules will migrate faster and travel a longer distance in a fixed time than larger molecules because smaller molecules have lower friction when passing through the pores of the gel. Molecules with different sizes and masses will have different speeds in passing through the pores of the gel and then form distinct bands on the gel. This is how the gel electrophoresis method sieves different molecules. After different molecules form distinct bands they can be compared with the standard marker in order to determine their molecular weights. There are a variety of types of gels. Two most commonly used gels are agarose and polyacrylamide gels. Agarose gels have lower resolving power than polyacrylamide gels. The advantage of agarose gels over polyacrylamide gels is that agarose gels have larger range of separation. Polyacrylamide gels are generally used for proteins given their very high resolving power resulting from the uniform pore size of polyacrylamide gels. The typical sizes of proteins can be separated by polyacrylamide gels range from 1-100kDa depending on the concentration of resolving gels [86-87]. The concentration of resolving gels usually ranges from 6%-15%. The stacking gel is typically 5%. Lower percentage gels have

larger pores that are better suited for resolving larger molecules, while higher percentage gels are used to resolve smaller molecules. Both native proteins and denatured proteins can be used in gel electrophoreisis. For large proteins denaturing might be necessary to get good resolution of the size. Depending on if molecules run in their native state wherein molecules preserve higher order structures or denaturated state where a chemical denaturant is added to break their high order structures, different procedure will apply. For denaturating proteins, sodium dodecyl sulfate polyacrylamide is used in the gel electrophoresis and is referred to as SDS-PAGE.

3.3.2 Protocol of gel electrophoresis

The SDS-PAGE usually comprises acrylamide, bisacrylamide, SDS, and a buffer with proper pH. In the experiments, tris-glycine SDS-PAGE was used for proteins with 6% resolving gel and 5% stacking gel. The recipes of Tris-Glycine SDS-PAGE are given in Table 3-7. Here SDS is sodium dodecyl sulfate, APS is ammonium persulfate, and

Table 3-7 Tris-Glycine SDS-PAGE recipe

6% resolving gels		5% stacking gels	
Components	Volume	Components	Volume
H_2O	5.3mL	H_2O	2.7mL
Acrylamide mix(30%)	2.0mL	Acrylamide mix(30%)	0.67mL
Tris(1.5M, PH 8.8)	2.5mL	Tris(1.5M, PH 8.8)	0.5mL
SDS(10%)	100μL	SDS(10%)	40μL
APS(10%)	100μL	APS(10%)	40μL
TEMED	8μL	TEMED	4μL
Total	10mL	Total	4mL

TEMED is N,N,N',N'-tetramethylethylenediamine. In general, one SDS molecule is approximately bound to two amino acids regardless of the sequences of polypeptides. Therefore, the migration of SDS bound proteins is proportional to the molecular weight of proteins. The procedure of SDS-PAGE of proteins used in the experiments is given as follows. First, prepare 10mL 6% resolving gel and 4mL 5% stacking gel based on the recipes for Tris-Glycine SDS-PAGE as given in Table 3-7. APS and TEMED should be added later for 5% stacking gel. Second, add about 7mL 6% resolving gel solution into the gel equipment Mini-PROTEAN II Cell (BIO-RAD, Hercules, CA, USA). Then, add 1mL isopropanol on top to remove air bubbles at the surface of resolving gel solution. Wait for 20 minutes and then remove the isopropanol. Add 40μL 10% APS and 4μL TEMED to the stacking gel solution. Place about 3mL 5% stacking gel solution above the resolving gel solution and put a 10-well Mini-Protean comb (BIO-RAD, Hercules, CA, USA) and wait for 30 minutes. Mix the sample solutions with dye bromophenol blue (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany) to the desired concentration. Seal all the sample solutions with parafilms and then boil them in a water bath for 5 minutes. Assemble the gel electrophoresis equipment properly. Add enough 1X SDS-PAGE loading buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) into the casket. Remove the comb gently. Add the dyed sample solutions into each comb slot carefully. Add enough 1X SDS-PAGE loading buffer into the chamber that surrounds the casket. Make sure there are bubbles generated at the bottom of the gel electrophoresis equipment. Set *V = 120V, I = 100mA, T = 75minutes*, and K in "Volts" mode. Start to run the gel and wait until the color has almost reached the bottom. Disassemble the equipment to get the gel

with distinct bands located at different positions. Use a plastic wedge plate to cut the top part of the gel. Put the gel into the transfer buffer (48 mM Tris, 39 mM glycine, 1.3mM SDS, 20%(v/v) methanol) in a clean container and then place it on a rotator with speed of 2 for 20 minutes. Cut one piece of Amersham Hybond 0.45μm PVDF blotting membrane (GE Healthcare Life Sciences, Pittsburgh, PA, USA) into the same size as the gel plate. Cut one corner to distinguish the orientation. Soak two pieces of blot papers with the transfer buffer in a clean container for 20 minutes. Soak the blotting membrane with pure methanol for 5 minutes. Then dump the methanol and add the transfer buffer to soak for 20 minutes. Place the blotting paper on the semi-dry transfer cell. Put the blotting membrane atop of the blotting paper. Place the gel plate on top of the blotting membrane. Put another blotting paper atop of the gel. Remove air bubbles by rolling a tube over the blotting paper. Add a little bit transfer buffer on top of the blotting paper. Set *V = 24V, I = 100mA, T = 28minutes*, and K in "Volts" mode, and start to run the gel. Add 0.5g dry milk to 10mL 1X Tris-buffered saline (TBS, 50 mM Tris-Cl, pH 7.5 150 mM NaCl) buffer to get 5% milk TBS. Put the blotting membrane into a clean container, add all the 5% milk TBS buffer. Incubate for 1 hour on a rocking platform with speed of 2 in a freezer. Discard the 5% milk TBS. Add 10mL reusable first antibody (goat anti-HIV P24 in 5% milk TBS, 1:500) and incubate overnight on a rocking platform with speed of 2 in a freezer. Wash the blotting membrane with 1X TBS buffer every 10 minutes for three times. Prepare 10mL 5% milk TBS with addition of 5μL second antibody (rabbit anti-goat secondary antibody). Add it into the blotting membrane and incubate for 90 minutes. Dump the milk TBS and wash the blotting membrane with 1X TBS buffer every 10

minutes for three times. Add enough developing solution (10mL 1M Tris-HCl(pH 9.5), 2mL 5M NaCl, 0.5mL 1M MgCl2, 33µL 50mg/mL NBT (nitro-blue tetrazolium chloride), and 16.5µL 50mg/mL BCIP (5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt) to completely cover the blotting membrane. Wait until bonds appear. Wash the blotting membrane with 1X TAE (40mM Tris, 20mM acetic acid, and 1mM EDTA (Ethylenediaminetetraacetic acid)) buffer. Dry the blotting membrane with nitrogen gas. Take a picture for the blotting membrane and analyze it using software ImagJ.

3.3.3 Analysis of gel electrophoresis outcomes

The motivation of implementing gel electrophoresis measurement is to justify whether the third peak in the AFM measurements correspoinds tetramers or trimers. The

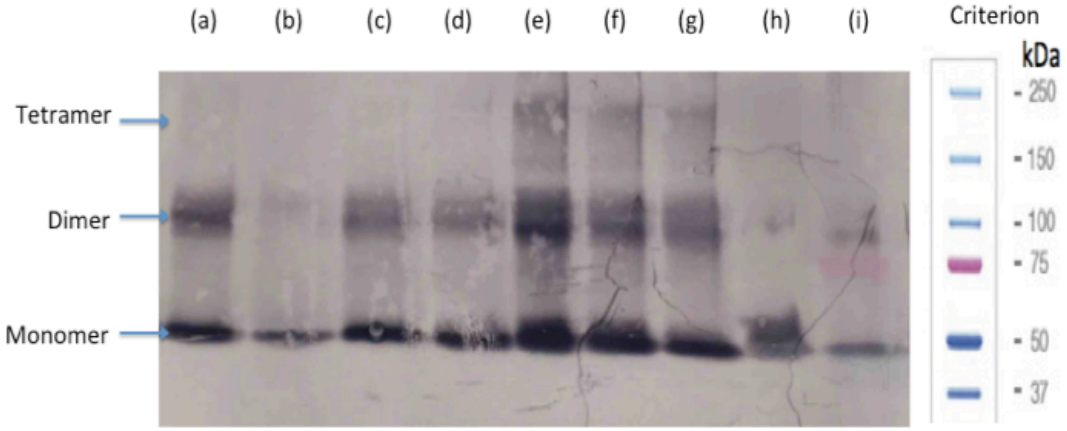

Figure 3-15 6% Tris-Glycine SDS-PAGE. (a) GagΔP6 (2µM). (b) GagΔP6 (0.5µM). (c) GagΔP6-ΨRNA (0.5µM : 0.5µM). (d) GagΔP6-ΨRNA (0.5µM : 2µM). (e) GagΔP6-ΨRNA (2µM : 0.5µM). (f) PI(4,5)P2-ΨRNA-GagΔP6 (0.5µM : 0.5µM : 0.5µM). (g) PI(4,5)P2-ΨRNA-GagΔP6 (2µM : 0.5µM : 0.5µM). (h) 50kDa Protein marker. (i) Protein standards. The rightmost is the criterion of molecular weight of protein standards. Bands corresponding to monomer, dimer, and tetramer are indicated on the left.

result of Tris-Glycine SDS-PAGE is shown in Figure 3-15. As the 50kDa protein marker indicates the bottom bands correspond to the 50kDa monomer. This is because the mass of Gag monomer is about 55kDa and the ratio of P6 segment to the entire Gag is 6/55 [88]. Therefore, the mass of GagΔP6 is around 50kDa. The middle bands are just above the 75kDa pink criteria and align with the 100kDa criteria, which means they are dimers. Since the approximate mass of the GagΔP6 dimer should be 100kDa. The top bands are in between the criterion 150kDa and 250kDa, which means they are tetramers with approximate molecular weight of 200kDa or 236kDa if including one ΨRNA. This result confirmed what we found in AFM measurement were tetramers rather than trimers. As shown in Figure 3-15(a) and (b), the percentages of dimer and tetramers increased as the concentration of GagΔP6 increased from 0.5μM to 2μM. This result confirmed the conclusion that the average radius of gyration Rg of GagΔP6 in solution measured by SANS is a monotonically increasing function of GagΔP6 concentration. This is reasonable because the size of the tetramer is greater than the size of the dimer, which is greater than the size of the monomer [37]. As shown in Figure 3-15(b) and (c), or (a) and (d), the percentages of dimer and tetramers increased with the addition of ΨRNA for the same concentration of GagΔP6. This also confirmed the AFM measurement where ΨRNA can bind with GagΔP6 and facilitate GagΔP6 multimerization. As shown in Figure 3-15(c) and (f) and (g), the percentages of dimer and tetramers increased even more when both ΨRNA and PI(4,5)P2 were added. Similar to ΨRNA, the more PI(4,5)P2 the more favorable for GagΔP6 multimerization. Base on the result of gel electrophoresis, our conclusion is the third peak in the AFM measurements were tetramers.

3.4 Summary

In this dissertation, the interaction of GagΔP6 with ΨRNA and PI(4,5)P2 was investigated by CM measurement, AFM measurement, and gel electrophoresis measurement. In the CM measurement, GagΔP6 selectively bound with PI(4,5)P2 on the leaflet of GUVs. As the AFM measurement showed GagΔP6 can not only bind with either ΨRNA or PI(4,5)P2, but simultaneously interact through a conformational change. Either ΨRNA or PI(4,5)P2 can facilitate GagΔP6 multimeration. In contrast to ΨRNA, PI(4,5)P2 facilitates GagΔP6 multimeration to a lesser extent. When both ΨRNA and PI(4,5)P2 are present, GagΔP6 multimeration can be further enhanced. As confirmed by the gel electrophoresis measurement, the third peak in the AFM measurement corresponds to tetramer rather than trimer. Selective and cooperative binding of HIV GagΔP6 with both ΨRNA and PI(4,5)P2 leads to a high local concentration of GagΔP6. GagΔP6 localization in turn results in GagΔP6 undertaking conformational changes that is crucial for HIV virus assembly.

3.5 References

[1] Binnig G., Quate C. F., and Gerber Ch.. (1986). Atomic force microscope. Phys. Rev. Lett. 56, 930.

[2] Müller D.J., Sass H.J., Müller S.A., Büldt G., and Engel A.. (1999). Surface structures of native bacteriorhodopsin depend on the molecular packing arrangement in the membrane. J. Mol. Biol., 285, 5, 1903-1909.

[3] Blanchard C.R.. (1996). Atomic force microscopy. Chem. Educ., 1, 5, 1-8.

[4] Meyer E.. (1992). Atomic force microscope. Prog. Surf. Sci., 41, 3-49.

[5] Ares P., Fuentes-Perez M.E., Herrero-Galán E., Valpuesta J.M., Gil. A., Gomez-Herrero J., and Moreno-Herrero F.. (2016). High resolution atomic force microscopy of double-stranded RNA. Nanoscale, 8, 23, 11818-11826.

[6] Pyne A., Thompson R., Leung C., Roy D., and Hoogenboom B.W.. (2014). Single-molecule reconstruction of oligonucleotide secondary structure by atomic force microscopy. Small, 10, 16, 3257-3261.

[7] Müller D.J., Fotiadis D., Scheuring S., Müller S.A., and Engel A.. (1999). Electrostatically balanced subnanometer imaging of biological specimens by atomic force microscope. Biophys. J., 76, 2, 1101-1111.

[8] Yamashita H., Voïtchovsky K., Uchihashi T., Contera S.A., Ryan J.F., Ando T.. (2009). Dynamics of bacteriorhodopsin 2D crystal observed by high-speed atomic force microscopy. J. Struct. Biol., 167, 2, 153-158.

[9] Casuso I., Khao J., Chami M., Paul-Gilloteaux P., Husain M., Duneau J.P., Stahlberg H., Sturgis J.N., and Scheuring S.. (2012). Characterization of the motion of membrane proteins using high-speed atomic force microscopy. Nature Nano., 7, 525-529.

[10] Butta H.J., Cappellab B., and Kappla M.. (2005) Force measurements with the atomic force microscope: Technique, interpretation and applications. Surf. Sci. Rep., 59, 1-6, 1-152.

[11] Cappella B. and Dietler G.. (1999). Force-distance curves by atomic force microscopy. Surf. Sci. Rep., 34, 1-3, 1-104.

[12] Lee C., Wei X., Kysar J.W., and Hone J.. (2008). Measurement of the elastic properties and intrinsic strength of monolayer graphene. Science, 321, 5887, 385-388.

[13] Fuhrmann A., Staunton J.R., Nandakumar V., Banyai N., Davies P.C., and Ros R.. (2011). AFM stiffness nanotomography of normal, metaplastic and dysplastic human esophageal cells. Phys. Biol., 8, 1, 015007.

[14] Cabrera A.P., Bhaskaran A., Xu J., Yang X., Scott H.A., Mohideen U., and Ghosh K.. (2016). Senescence increases choroidal endothelial stiffness and susceptibility to complement injury: implications for choriocapillaris loss in AMD. Invest Ophthalmol Vis Sci., 57, 14, 5910-5918.

[15] Cabrera A.P., Bhaskaran A., Xu J., Yang X., Scott H.A., Mohideen U., and Ghosh K.. (2016). Basement membrane stiffening promotes retinal endothelial activation associated with diabetes. FASEB J., 30, 2, 601-611.

[16] Xu D., Watt G.D., Harb J.N., and Davis R.C.. (2005). Electrical conductivity of ferritin proteins by conductive AFM. Nano Lett., 5, 4, 571-577.

[17] Moy V.T., Florin E.L., and Gaub H.E.. (1994). Adhesive forces between ligand and receptor measured by AFM. Colloids Surf. A Physicochem. Eng. Asp., 93, 343-348.

[18] Dai Q., Vollmer R., Carpick R.W., Ogletree D. F., and Salmeron M.. (1995). A variable temperature ultrahigh vacuum atomic force microscope. Rev. Sci. Instrum., 66, 5266.

[19] Radenović A., Bystrenová E., Libioulle L., Taborelli M., DeRose J.A., and Dietler G.. (2003). A low-temperature ultrahigh vacuum atomic force microscope for biological applications. Rev. Sci. Instrum., 74, 1022.

[20] Gysin U., Rast S., Kisiel M., Werle C., and Meyer E.. (2011). Low temperature ultrahigh vacuum noncontact atomic force microscope in the pendulum geometry. Rev. Sci. Instrum., 82, 023705.

[21] Wickramasinghe H.K.. (1989). Scanned-probe microscopes. Sci. Am., 261, 4, 98-105.

[22] Rojas O.J.. (2002). Adsorption of polyelectrolytes on mica. Ency. Surf. Coll. Sci., 517-535.

[23] Müller D.J., Amrein M., and Engel A.. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. J. Struct. Biol., 119, 2, 172-188.

[24] Müllerab D.J., Engela A., and Amreinc M.. (1997). Preparation techniques for the observation of native biological systems with the atomic force microscope. Biosens. Bioelectron., 12, 8, 867-877.

[25] Lyubchenko Y.L., Shlyakhtenko L.S., and Ando T.. (2011). Imaging of nucleic acids with atomic force microscopy. Methods, 54, 2, 274-283.

[26] Bussiek, M., Schöne, A. and Nellen, W.. (2014). Atomic force microscopy imaging and force spectroscopy of RNA. Handbook of RNA Biochemistry: Second, Completely Revised and Enlarged Edition, 527-546.

[27] Pastré D., Piétrement O., Fusil S., Landousy F., Jeusset J., David M.O., Hamon L., Le Cam E., and Zozime A.. (2003). Adsorption of DNA to mica mediated by divalent counterions: a theoretical and experimental study. Biophys. J., 85, 4, 2507-2518.

[28] Hansma H.G. and Laney D.E.. (1996). DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. Biophys. J. 70, 4, 1933-1939.

[29] Zheng J., Li Z., Wu A., and Zhou H.. (2003). AFM studies of DNA structures on mica in the presence of alkaline earth metal ions. Biophys. Chem., 104, 1, 37-43.

[30] Thomson N.H., Kasas S., Hansma H.G., and Hansma P. K.. (1996). Reversible binding of DNA to mica for AFM imaging. Langmuir, 12, 24, 5905-5908.

[31] Kienberger F., Costa L.T., Zhu R., Kada G., Reithmayer M., Chtcheglova L., Rankl C., Pacheco A.B., Thalhammer S., Pastushenko V., Heckl W.M., Blaas D., and Hinterdorfer P.. (2007). Dynamic force microscopy imaging of plasmid DNA and viral RNA. Biomaterials, 28, 15, 2403-2411.

[32] Fay M.J., Walter N.G., and Burke J.M.. (2001). Imaging of single hairpin ribozymes in solution by atomic force microscopy. RNA, 7, 6, 887-895.

[33] Hansma H.G., Golan R., Hsieh W., Lollo C.P., Mullen-Ley P., and Kwoh D.. (1998). DNA condensation for gene therapy as monitored by atomic force microscopy. Nucleic Acids Res., 26, 10, 2481-2487.

[34] Bezanilla M., Manne S., Laney D.E., Lyubchenko Y.L., and Hansma H.G.. (1995). Adsorption of DNA to Mica, Silylated Mica, and Minerals: Characterization by Atomic Force Microscopy. Langmuir, 11, 2, 655-659.

[35] Hsieh S., Chao W.J., and Hsieh C.W.. (2009). Improved performance of aminopropylsilatrane over aminopropyltriethoxysilane as an adhesive film for anchoring gold nanoparticles on silicon surfaces. J. Nanosci. Nanotechnol., 9, 5, 2894-2901.

[36] Raposo M., Ferreira Q., and Ribeiro P. A.. (2007) A guide for Atomic Force Microscopy Analysis of Soft-Condensed Matter. Mod. Res. Educ. Top. Microsc.， 758-769.

[37] Pfreundschuh M., Martinez-Martin D., Mulvihill E., Wegmann S., and Muller D.J.. (2014). Multiparametric high-resolution imaging of native proteins by force-distance curve–based AFM. Nat. Protoc., 9, 1113–1130.

[38] Andersen E.S., Contera S.A., Knudsen B., Damgaard C.K., Besenbacher F., Kjems J. (2004). Role of the trans-activation response element in dimerization of HIV-1 RNA. J. Biol. Chem., 279, 21, 22243-22249.

[39] Hansma H.G., Oroudjev E., Baudrey S., and Jaeger L.. (2003). TectoRNA and 'kissing-loop' RNA: atomic force microscopy of self-assembling RNA structures. J. Microsc., 212, 3, 273-279.

[40] Harvey S.C.. (2015). The scrunchworm hypothesis: transitions between A-DNA and B-DNA provide the driving force for genome packaging in double-stranded DNA bacteriophages. J. Struct. Biol., 189, 1, 1-8.

[41] Watson J.D. and Crick F.H.C.. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature, 171, 737-738.

[42] Franklin R.E. and Gosling R.G.. (1953). Molecular configuration in sodium thymonucleate. Nature, 171, 740-741.

[43] Rein A., Datta S.A.K., Jones C.P., and Musier-Forsyth K. (2011). Diverse interactions of retroviral Gag proteins with RNAs. Trends. Biochem. Sci., 36, 373-380.

[44] Datta S.A., Heinrich F., Raghunandan S., Krueger S., Curtis J.E., Rein A., and Nanda H.. (2011). HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. J. Mol. Biol., 406, 2, 205-214.

[45] Datta S.A.K., Curtis J.E., Ratcliff W., Clark P.K., Crist R.M., Lebowitz J., Krueger S., and Rein A. (2007). Conformation of the HIV-1 Gag protein in solution. J. Mol. Biol., 365, 812-824.

[46] Love J.C., Estroff L.A., Kriebel J.K., Nuzzo R.G., and Whitesides G.M. (2005). Self-assembled monolayers of thiolates on metals as a form of nanotechnology. Chem. Rev., 105, 4, 1103-1170.

[47] Budvytyte R., Valincius G., Niaura G., Voiciuk V., Mickevicius M., Chapman H., Goh H.G., Shekhar P., Heinrich F., Shenoy S., Lösche M., and Vanderah D.J.. (2013). Structure and properties of tethered bilayer lipid membranes with unsaturated anchor molecules. Langmuir, 29, 27, 8645-8656.

[48] Ulman A.. (1996). Formation and structure of self-assembled monolayers, Chem. Rev., 96, 4, 1533-1554.

[49] Lingler S., Rubinstein I., Knoll W., and Offenhäusser A.. (1997). Fusion of small unilamellar lipid vesicles to alkanethiol and thiolipid self-assembled monolayers on gold. Langmuir, 13, 26, 7085-7091.

[50] Sun W., Kewalramani S., Hujsak K., Zhang H., Bedzyk M.J., Dravid V.P., and Thaxton C.S.. (2015). Mesophase in a thiolate-containing diacyl phospholipid self-assembled monolayer. Langmuir, 31, 10, 3232-3241.

[51] Lahiri J., Kalal P., Frutos A.G., Jonas S.J., and Schaeffler R.. (2000). Method for fabricating supported bilayer lipid membranes on gold. Langmuir, 16, 20, 7805-7810.

[52] Vockenroth I.K., Rossi C., Shah M.R., and Köper I.. (2009). Formation of tethered bilayer lipid membranes probed by various surface sensitive techniques. Biointerphases, 4, 2, 19-26.

[53] McGillivray D.J., Valincius G., Vanderah D.J., Febo-Ayala W., Woodward J.T., Heinrich F., Kasianowicz J.J., and Löscheb M.. (2007). Molecular-scale structural and functional characterization of sparsely tethered bilayer lipid membranes. Biointerphases, 2, 1, 21-33.

[54] Castellana E.T. and Cremer P.S.. (2006). Solid supported lipid bilayers: From biophysical studies to sensor design. Surf. Sci. Rep., 61, 10, 429-444.

[55] Mou J., Yang J., and Shao Z.. (1995). Atomic force microscopy of cholera toxin B-oligomers bound to bilayers of biologically relevant lipids. J. Mol. Biol., 248, 3, 507-512.

[56] Saslowsky D.E., Lawrence J., Ren X., Brown D.A., Henderson R.M., and Edwardson J.M.. (2002). Placental alkaline phosphatase is efficiently targeted to rafts in supported lipid bilayers. J. Biol. Chem., 277, 30, 26966-26970.

[57] Rinia H.A. and de Kruijff B.. (2001). Imaging domains in model membranes with atomic force microscopy. FEBS Letters, 504, 3, 194, 199.

[58] Rinia H.A., Snel M.M.E., van der Eerden J.P.J.M., and de Kruijff B.. (2001). Visualizing detergent resistant domains in model membranes with atomic force microscopy. FEBS Letters, 501, 1, 92, 96.

[59] Webb J.A., Jones C.P., Parent L.J., Rouzina I., and Musier-Forsyth K.. (2013). Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging. RNA, 19, 8, 1078-1088.

[60] Comas-Garcia M, Davis S.R., and Rein A.. (2016). On the selective packaging of genomic RNA by HIV-1. Viruses, 8, 9, 246.

[61] Didierlaurent L., Racine P.J., Houzet L., Chamontin C., Berkhout B., and Mougel M.. (2011). Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. Nucleic Acids Res., 39, 20, 8915-8927.

[62] Houzet L., Paillart J.C., Smagulova F., Maurel S., Morichaud Z., Marquet R., and Mougel M.. (2007). HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res., 35, 8, 2695-2704.

[63] Ono A., Ablan S.D., Lockett S.J., Nagashima K., and Freed E.O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. PNAS, 101, 41, 14889-14894.

[64] Saad J.S., Miller J., Tai J., Kim A., Ghanam R.H., and Summers M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. PNAS, 103, 30, 11364-11369.

[65] Carlson L.A. and Hurley J.H.. (2012). In vitro reconstitution of the ordered assembly of the endosomal sorting complex required for transport at membrane-bound HIV-1 Gag clusters. PNAS, 109, 42, 16928-16933.

[66] Chukkapalli V., Hogue I.B., Boyko V., Hu W.S., and Ono A. (2008). Interaction between the human immunodeficiency virus type 1 Gag matrix domain and phosphatidylinositol-(4,5)-bisphosphate is essential for efficient gag membrane binding. J. Virol., 82, 5, 2405-2417.

[67] Shkriabai N., Datta S.A., Zhao Z., Hess S., Rein A., and Kvaratskhelia M. (2006) Interactions of HIV-1 Gag with assembly cofactors. Biochemistry, 45, 13, 4077-4083.

[68] Alfadhli A., Barklis R.L., and Barklis E. (2009). HIV-1 matrix organizes as a hexamer of trimers on membranes containing phosphatidylinositol-(4,5)-bisphosphate. Virology, 387, 2, 466-472.

[69] Freed E.O.. (2006). HIV-1 Gag: Flipped out for PI(4,5)P2. PNAS, 103, 30, 11101-11102.

[70] Spearman P., Horton R., Ratner L., and Kuli-Zade I.. (1997). Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. J. Virol., 71, 9, 6582-6592.

[71] Ono A., Demirov D., and Freed E.O.. (2000). Relationship between human immunodeficiency virus type 1 Gag multimerization and membrane binding. J. Virol., 74, 11, 5142-5150.

[72] Wright E.R., Schooler J.B., Ding H.J., Kieffer C., Fillmore C., Sundquist W.I., and Jensen G.J.. (2007). Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J., 26, 8, 2218-2226.

[73] Briggs J.A., Riches J.D., Glass B., Bartonova V., Zanetti G., and Kräusslich H.G. (2009). Structure and assembly of immature HIV. PNAS, 106, 27, 11090-11095.

[74] Ganser-Pornillos B.K., Yeager M., and Pornillos O. (2012). Assembly and architecture of HIV. Adv. Exp. Med. Biol., 726, 441-465.

[75] Sundquist, W.I. and Kräusslich, H.G. (2012). HIV-1 Assembly, budding, and maturation. CSH Perspect. Med., 2, 7, a006924.

[76] Freed E.O. (2015). HIV-1 assembly, release and maturation. Nat. Rev. Microbiol., 13, 484–496.

[77] Zhou W., Parent L.J., Wills J.W., and Resh M.D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J. Virol., 68, 4, 2556-2569.

[78] Maizel J.V.. (2000). SDS polyacrylamide gel electrophoresis. Trends Biochem. Sci., 25, 12, 590-592.

[79] Lee P.Y., Costumbrado J., Hsu C.Y., and Kim Y.H.. (2012). Agarose gel electrophoresis for the separation of DNA fragments. J. Vis. Exp., 62, 3923.

[80] Smith S.B., Aldridge P.K., Callis J.B.. (1989). Observation of individual DNA molecules undergoing gel electrophoresis. Science, 243, 4888, 203, 206.

[81] Rio D.C., Ares Jr M., Hannon G.J., and Nilsen T.W.. (2010). Polyacrylamide gel electrophoresis of RNA. Cold Spring Harb. Protoc.,2010, 6.

[82] Rio D.C., Ares Jr M., Hannon G.J., and Nilsen T.W.. (2010). Nondenaturing agarose gel electrophoresis of RNA. Cold Spring Harb. Protoc., 2010, 6.

[83] Walker J.M.. (1984). Gradient SDS polyacrylamide gel electrophoresis of proteins. Methods Mol. Biol., 1, 57-61.

[84] O'Farrell P.H.. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem., 250, 10, 4007-4021.

[85] Weber K., and Osborn M.. (1969). The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem., 244, 16, 4406-4412.

[86] Schägger H. and von Jagow G.. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem., 166, 2, 368-379.

[87] Schägger H.. (2006). Tricine–SDS-PAGE. Nat. Protoc., 1, 16-2.

[88] Bell N.M., and Lever A.M.L. (2013). HIV Gag polyprotein: processing and early viral particle assembly. Trends Microbiol., 21, 136-144.

Appendix A

AFM tip calibration

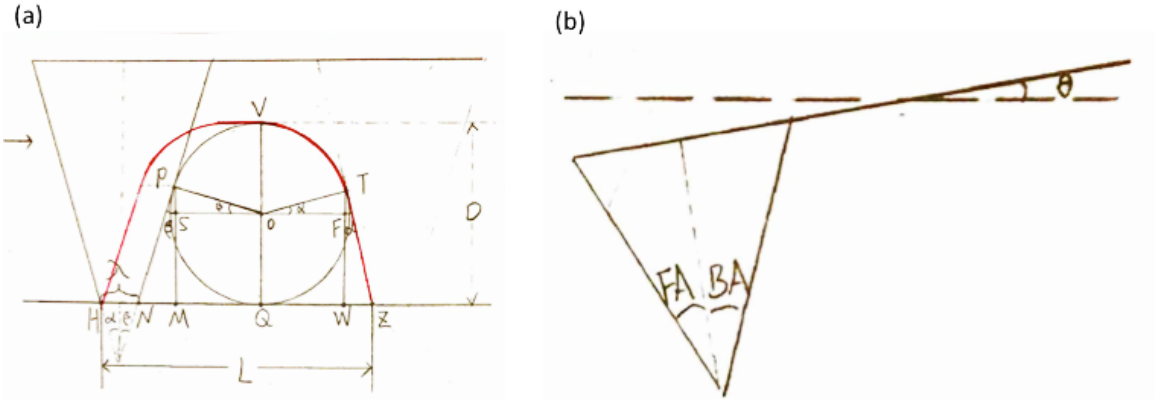

Figure A-1 AFM tip calibration. (a) Schematic representation of the tip and a spherical sample. The red curve is the trajectory of the tip when it scans from left to right as indicated by the arrow on left. (b) The configuration of the tip after it is placed into the AFM probe holder (fluid cell if imaging in liquid).

As shown in Figure A-1(a), λ is the actual tip size, L is the measured size of the sample, D is the height of the sample which is a sphere. α and β are effective front and back angles of the tip which are related to actual front angle FA and back angle BA of the tip FA through the tilt angle, θ, of the AFM probe holder (fluid cell if imaging in liquid) as shown in Figure A-1(b).

$$\alpha = FA + \theta$$

$$\beta = BA - \theta$$

$$HN = \lambda$$

$$OP = \frac{D}{2}$$

$$PS = OP \sin\beta$$

$$OS = OP \cos\beta$$

$$MQ = OS = \frac{D}{2} \cos\beta$$

$$NM = PM \tan\beta = (PS + SM) \tan\beta = \frac{D}{2} \tan\beta (\sin\beta + 1)$$

$$NQ = NM + MQ = \frac{D}{2} \tan\beta (\sin\beta + 1) + \frac{D}{2} \cos\beta = \frac{D}{2} (\sec\beta + \tan\beta)$$

Similarly,

$$QZ = \frac{D}{2} (\sec\alpha + \tan\alpha)$$

$$L = HN + NQ + QZ = \lambda + \frac{D}{2} (\sec\alpha + \tan\alpha + \sec\beta + \tan\beta)$$

For $\theta = 10^0$, $FA = BA = 20^0$, then $\alpha = 30^0$ and $\beta = 10^0$, hence,

$$L \approx \lambda + 1.46D$$

Therefore, the actual tip size is

$$\lambda = L - 1.46D \quad (1)$$

And the effective tip size t for a sample with the height of D is

$$t = L - D = \lambda + 0.46D \quad (2)$$

Appendix B

Algorithm of computation of the sample size

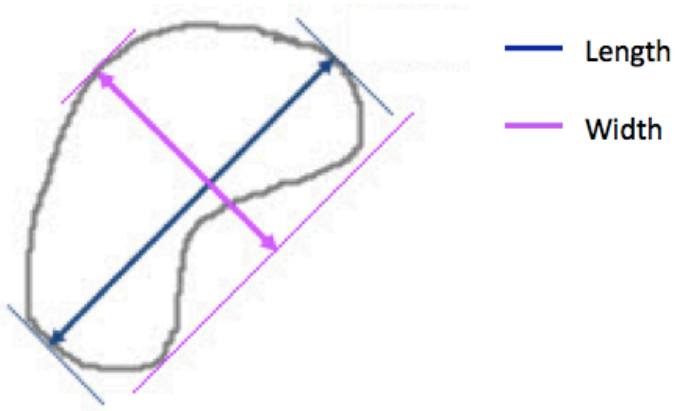

Figure B-1 Schematic diagram of a sample with defined length and width.

The size of the sample is computed by a MATLAB script. First, a proper threshold value is selected to compute the height of the sample. The threshold value is a cutoff to avoid the background noises. The height is the distance from the highest point of the sample to the cutoff plane. After the cutoff plane is set, the boundary of the sample can be obtained from the intersection of the sample and the cutoff plane. From that the length of the sample is defined as the distance between two furthest points on the boundary. The width is defined as the distance between the two parallel lines restricting the object perpendicular to the direction of the length.

Appendix C

MATLAB script to analyze raw AFM image data

```
clear all;

syms N threshold filter elpcut range tipsize;

N=512; % number of pixeles

range=1000.0; %scan size in nm

pixel=range/N; % each pixel size

tipsize=2.0; %nm

threshold=0.1; %preliminary height cutoff in nm

area=40.0; % minimum gag size in nm^2

filter=area/pixel/pixel; % minimum gag size in pixel^2

elpcut=10.0; %length/width ratio cutoff

lmin=2.0; % actual minimum length in nm

lmax=50.0; %actual maximum length in nm

wmin=2.0; % actual minimum width in nm

wmax=30.0; % actual maximum width in nm

hmin=0.5; % actual minimum height in nm

hmax=3.5; % actual maximum height in nm
```

```
files = dir('*.txt');
filename = {files.name};
nf=length(filename);
fileID = fopen('data.csv','w');
for k=1:nf
    disp(files(k).name);
    file=fopen(files(k).name);
    img_char=fscanf(file,'%c');
    img_num=str2num(img_char);
    fclose(file);
    for i=1:N
        for j=1:N
            img_raw(i,j)=img_num(j+(i-1)*N,1);
        end
    end
    img_raw=transpose(img_raw);
    img_raw=rot90(img_raw);
    raw_ave=sum(sum(img_raw))/N/N;
    fprintf('%s%f\n','raw data average=',raw_ave)
```

```
figure
subplot(2,2,1);
image(img_raw,'CDatamapping','scaled')
title('raw data')
colorbar
count=0;
for i=1:N
   for j=1:N
      if img_raw(i,j)<threshold/calibration/10
         img_base(i,j)=img_raw(i,j);
         img_data(i,j)=0.0;
         count=count+1;
      else img_base(i,j)=0.0;
          img_data(i,j)=img_raw(i,j);
      end
   end
end
base_ave=sum(sum(img_base))/count;
fprintf('%s%f\n','base plane average=',base_ave)
```

```
subplot(2,2,2);
image(img_base,'CDatamapping','scaled')
title('base plane')
colorbar
img_data(1,:)=0.0;
img_data(:,1)=0.0;
subplot(2,2,3);
image(img_data,'CDatamapping','scaled')
title('preliminary selected data')
colorbar
label=zeros(N);
k=0;
for i=2:N
   for j=2:N
      k=max(max(label))+1;
      if img_data(i,j)==0.0
         label(i,j)=0.0;
      elseif img_data(i-1,j)==0.0&&img_data(i,j-1)==0.0
          label(i,j)=k;
```

```
        elseif img_data(i-1,j)*img_data(i,j-1)==0.0
            label(i,j)=max(label(i-1,j),label(i,j-1));
        else label(i,j)=min(label(i-1,j),label(i,j-1));
            idx=find(label==max(label(i-1,j),label(i,j-1)));
            label(idx)=min(label(i-1,j),label(i,j-1));
        end
      end
   end
   max_temp=max(max(label));
   disp(['preliminary total number=' num2str(max_temp)])
   k_new=1;
   edge=0.0;
   %[number,peak_x,peak_y,peak_z,long axis,short axis]
   for k=1:max_temp
        idx=find(label==k) ;
        edge=ismember(k,label(2,:))+ismember(k,label(:,2))+ismember(k,label(N,:))+
ismember(k,label(:,N));
        [l,s]=dimension(idx,N);
        [peak_z,I]=max(img_data(idx));
```

```
        peak_idx=idx(I);
        l=l*pixel-tipsize; %long axis=length in nm
         s=s*pixel-tipsize; %short axis=width in nm
         h=peak_z-base_ave;  %peak_z=height in nm
         if size(idx,1)<=filter||edge>0.0||l/s>elpcut||l<=s||l<=lmin||l>=lmax||s<=wmin||
s>=wmax||h<=hmin||h>=hmax
                label(idx)=0.0;
         else disp(['now, process point number: ' num2str(k_new)])
                label(idx)=k_new;
                %DATA(k_new,1)=k_new; % label No.
                %DATA(k_new,2)=ceil(peak_idx/N); %peak_x
                %DATA(k_new,3)=mod(peak_idx,N);  %peak_y
                  DATA(k_new,1)=l; %length
                DATA(k_new,2)=s; %width
                DATA(k_new,3)=h; %height
                 k_new=k_new+1;
        end
    end
    max_filtered=max(max(label));
```

```
    disp(['filtered total number=' num2str(max_filtered)])
    subplot(2,2,4);
    image(label,'CDatamapping','scaled')
    title('label matrix')
    colorbar
    DATA=transpose(DATA);
    fprintf(fileID,'%10s,%10s,%10s\r\n','Length/nm','Width/nm','Height/nm');
    fprintf(fileID,'%5.1f,%5.1f,%5.1f\r\n',DATA);
    clear DATA;
end
fclose(fileID);
DATA=importdata('data.csv');
figure
    subplot(1,3,1);
    histogram(DATA(:,1),15)
    title('Histogram of Length')
    xlabel('Length/nm') % x-axis label
    ylabel('Frequency') % y-axis label
    subplot(1,3,2);
```

```
histogram(DATA(:,2),15)

title('Histogram of Width')

xlabel('Width/nm') % x-axis label

ylabel('Frequency') % y-axis label

subplot(1,3,3);

histogram(DATA(:,3),15)

title('Histogram of Height')

xlabel('Height/nm') % x-axis label

ylabel('Frequency') % y-axis label
```

Appendix D

Mathematica script to plot histogram and fit distribution

```
ClearAll["Global`*"];
SetDirectory[NotebookDirectory[]];
data = Import["data.csv"];
n = Length[data];
bins = 20;
cutoff = 0; (*cutoff in height in nm*)
t = 5.4; (*effective tip size for GagΔP6.*)
For[i = 1, i <= n, i++, data[[i]][[1]] = Abs[data[[i]][[1]] - t]; data[[i]][[2]] =
Abs[data[[i]][[2]]] - t];
l = Table[Abs[data[[i]][[1]]], {i, 1, n}];(*l= GagΔP6 length in nm*)
w = Table[Abs[data[[i]][[2]]], {i, 1, n}];(*w= GagΔP6 width in nm*)
h = Table[data[[i]][[3]] + cutoff, {i, 1, n}];(*h= GagΔP6 height in nm*)
lh = Histogram[l, {"Raw", bins}, AxesLabel -> {"Length/nm", "Frequency"}];
wh = Histogram[w, {"Raw", bins}, AxesLabel -> {"Width/nm", "Frequency"}];
hh = Histogram[h, {"Raw", bins}, AxesLabel -> {"Height/nm", "Frequency"}];
GraphicsRow[{lh, wh, hh}, PlotLabel -> "Histogram of 0.5μM GagΔP6 on Mica(-)"];
(*Length fitting*)
n1 = 590;
n2 = 350;
```

```
n3 = 60;
sortdata = Sort[data, #1[[1]] < #2[[1]] &];
monomer = Table[sortdata[[i]], {i, 1, n1}];
l1 = Table[monomer[[i]][[1]], {i, 1, n1}];
lh1 = Histogram[l1, bins, AxesLabel -> {"Length/nm", "Frequency"}];
dimer = Table[sortdata[[i]], {i, n1 + 1, n1 + n2}];
l2 = Table[dimer[[i]][[1]], {i, 1, n2}];
lh2 = Histogram[l2, bins, AxesLabel -> {"Length/nm", "Frequency"}];
trimer = Table[sortdata[[i]], {i, n1 + n2 + 1, n}];
l3 = Table[trimer[[i]][[1]], {i, 1, n3}];
lh3 = Histogram[l3, bins, AxesLabel -> {"Length/nm", "Frequency"}];
peak1 = Mean[l1];
s1 = StandardDeviation[l1];
peak2 = Mean[l2];
s2 = StandardDeviation[l2];
peak3 = Mean[l3];
s3 = StandardDeviation[l3];
distl = MixtureDistribution[{n1, n2, n3}, {NormalDistribution[peak1, s1],
NormalDistribution[peak2, s2], NormalDistribution[peak3, s3]}];
(*Width fitting*)
nw1 = n1 + n2;
nw2 = n – nw1;
```

```
sortWidth = Sort[data, #1[[2]] < #2[[2]] &];
width1 = Table[sortWidth[[i]], {i, 1, nw1}];
w1 = Table[width1[[i]][[2]], {i, 1, nw1}];
wh1 = Histogram[w1, {"Raw", bins}, AxesLabel -> {"Width/nm", "Frequency"}];
width2 = Table[sortWidth[[i]], {i, nw1 + 1, n}];
w2 = Table[width2[[i]][[2]], {i, 1, nw2}];
wh2 = Histogram[w2, {"Raw", bins}, AxesLabel -> {"Width/nm", "Frequency"}];
peakw1 = Mean[w1];
sw1 = StandardDeviation[w1];
peakw2 = Mean[w2];
sw2 = StandardDeviation[w2];
distw = MixtureDistribution[{nw1, nw2}, {NormalDistribution[peakw1, sw1],
NormalDistribution[peakw2, sw2]}];
(*Height fitting*)
peakh = Mean[h];
sh = StandardDeviation[h];
disth = MixtureDistribution[{n, 0}, {NormalDistribution[peakh, sh],
NormalDistribution[peakh, sh]}];
lhf = Show[lh, Plot[1.45 n*PDF[distl, x], {x, Min[l], Max[l]}, PlotStyle -> {Red,
Thick}]];
whf = Show[wh, Plot[0.86 n*PDF[distw, x], {x, Min[w], Max[w]}, PlotStyle -> {Red,
Thick}]];
```

```
hhf = Show[hh, Plot[n*PDF[disth, x], {x, Min[h], Max[h]}, PlotStyle -> {Red, Thick}]];
GraphicsRow[{lhf, whf, hhf},
 PlotLabel -> "Histogram of 0.5μM GagΔP6 on Mica(-)", Frame -> All, FrameStyle ->
Purple]
lengthWidth = Table[{data[[i]][[1]], data[[i]][[2]]}, {i, 1, n}];
Histogram3D[lengthWidth, ChartElementFunction -> "GradientScaleCube", AxesLabel ->
{"Length/nm", "Width/nm", "Frequency"}, PlotLabel -> "3D Histogram of 0.5μM
GagΔP6 on Mica(-)"];
SmoothHistogram3D[lengthWidth, AxesLabel -> {"Length/nm", "Width/nm",
"Density"}, PlotLabel -> "3D Smooth Histogram of 0.5μM GagΔP6 on Mica(-)"]
```