

STRUCTURAL STUDIES OF AUTONOMOUS PARVOVIRUS CAPSIDS TOWARDS  
UNDERSTANDING THE MECHANISMS OF CELL RECOGNITION AND ENDOSOMAL  
TRAFFICKING

By

SUJATA HALDER

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

UNIVERSITY OF FLORIDA

2012

© 2012 Sujata Halder

To my parents, Debasree Saha Halder and Gokul Chandra Saha Halder for always  
believing in my abilities

## ACKNOWLEDGMENTS

I would like to begin by thanking my parents; Debasree and Gokul Chandra Saha Halder for the sacrifices they made to provide me with great opportunities and I am forever grateful because they continue to love, support and encourage me even after my countless mistakes. I thank my brother, Kunal Halder and grandparents, especially Dr. Rabindranath Basu for timely advice and much needed financial help. I would like to thank my uncle and aunt, Hari Peri and Rama Peri for being the ideal guardians and for their warmth and words of advice. I feel extremely indebted and honored to be mentored by Dr. Mavis Agbandje-McKenna who believed in my abilities and patiently nurtured the scientist in me. I am always grateful for the encouragement and opportunities she has provided me and the personal counseling sessions. I would also like to thank Dr. Robert McKenna for being an inspiring teacher and his great sense of humor. My committee members, Dr. James B. Flanagan, Dr. Nicholas Muzyczka, Dr. David C. Bloom, and Dr. Grant McFadden have been extremely encouraging and helpful and provided critical input. I would also like to thank my collaborators, Dr. Peter Tattersall (Yale University), Dr. Susan F. Cotmore (Yale University), Dr. Nathalie Salome (Program Infection and Cancer, Germany), Dr. David F. Smith (Emory University School of Medicine) and the beamline scientists at CHESS (Cornell High Energy Synchrotron Source) for the resources and time invested in my work. I greatly appreciate the help offered by the administrative staff of the IDP (Interdisciplinary Program), BMB (Biochemistry and Molecular Biology) and International Student Center. I am grateful to be a member of ‘The McKenna Lab’ which has been my ‘home’ away from home. Dr. Lakshmanan Govindasamy, Dr. Hyun-Joo Nam, Dr. Antonette Bennett, Dr. Brittney L. Gurda deserve special thanks for teaching me the techniques used in the

lab and also for invaluable advice; Balendu Avvaru, Bala Venkat, Harald Messer, Lawrence Tartaglia and Lauren Drouin for it helps to have great friends in the lab; Robert Ng and Yue Liu for always being there when I needed your help and for ‘adopting’ me and my husband; the newcomers in the lab for the youthfulness and helping the ‘experienced’ me. I would like to thank the undergraduate students that I have taught over the years, especially Omar Shakeel, Caroline Zapiec and Jay Jackson for being patient with me. I also thank my IDP buddy Mansi Parekh who helped me immensely in my first year in IDP. I am grateful to have been blessed with amazing friends; Swati Debroy for being present at all the special events in my life including the first day I landed in Gainesville and for undoubtedly being the perfect elder sister; Priyanka Roy for her fun-filled company in my initial years at University of Florida; Subhajit Sengupta, Kiranmoy Das, Arkendu Chatterjee, Souvik Bhattacharya for pampering me like your younger sister; Dipti Patel for the invaluable help in adjusting to a new country; Debpriya Dutta, Sugata Hazra, Ranjana Sarma and Deepak Kar for always being there when I needed you. Last but definitely not the least, my heartfelt thanks to my friend and husband Sourish Dasgupta whose unconditional love and support has helped me tremendously in this memorable journey.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	9
LIST OF FIGURES.....	10
LIST OF ABBREVIATIONS.....	12
ABSTRACT .....	15
CHAPTER	
1 INTRODUCTION .....	17
Overview of Viruses.....	17
<i>Parvoviridae</i> Family .....	19
Taxonomic Classification.....	20
Genome Architecture .....	21
Capsid Composition and Structure.....	23
Infectious Pathway of Parvoviruses .....	28
Receptor Binding.....	29
Intracellular Trafficking and Structural Transitions Induced in the Virion .....	32
Tissue Tropism and Pathogenicity Determinants.....	40
Receptor Binding and Virulence Determination .....	43
Parvoviruses as Gene Therapy Vectors .....	46
Significance .....	48
2 ANALYSIS OF MINUTE VIRUS OF MICE (MVM) RECEPTOR COMPLEXES TOWARDS UNDERSTANDING THE MECHANISMS OF TISSUE TROPISM.....	60
Background.....	60
Experimental Methods .....	62
Cell Lines.....	62
Generation of Full and Empty MVM capsids .....	63
Recombinant Virus Production and Purification .....	64
Crystallization and Data Collection.....	65
Structure Solution.....	66
Sialylated Glycan Microarray (SGM) Preparation and Virus Screening .....	68
Glycan Profiling of Permissive Cell Lines .....	70
Results.....	70
Virus Purification and Crystallization .....	70
Structural Analysis of MVM Capsid-Receptor Interactions .....	71
Binding of MVM Viruses to SGM .....	74
Glycan Profiling of Cell Lines.....	75

Discussion .....	80
Capsid-Receptor Interactions Dictate Infectious Outcome .....	80
Recognition of Sialic Acid Derivatives by MVM Viruses .....	85
Cellular Glycan Profiling Validates Glycan Screening Data .....	88
Summary .....	92
<b>3 STRUCTURAL CHARACTERIZATION OF H-1PV: INSIGHTS INTO CAPSID-GENOME INTERACTIONS AND RECEPTOR BINDING .....</b>	<b>106</b>
Background.....	106
Experimental Methods .....	109
Cell Lines.....	109
Virus Production and Purification .....	109
Glycan Array Analysis .....	110
Crystallization .....	112
Diffraction Data Collection and Processing .....	112
Structure Determination.....	114
Structural Alignment Comparison.....	118
Results.....	119
Virus Purification and Crystallization .....	119
Glycans Recognized by H-1PV .....	119
H-1PV Capsid Structure .....	120
Structural Studies of H-1PV and Glycan Complexes.....	124
Discussion .....	125
Recognition of Sialylated Glycans by H-1PV .....	125
Structure of H-1PV .....	127
Structural Comparison to Other Autonomous Parvoviruses .....	128
DNA-Binding Pocket.....	130
Conformational Variation Between Full and Empty Capsids .....	134
Cis-Peptide Bonds.....	135
Glycan Binding Site on H-1PV Capsid .....	136
Summary .....	137
<b>4 PARVOVIRUS CAPSID DYNAMICS ASSOCIATED WITH ENDOSOMAL TRAFFICKING .....</b>	<b>153</b>
Background.....	153
Experimental Methods .....	155
Virus Production and Purification .....	155
Crystallization, Data Collection and Processing .....	156
Structure Determination.....	157
Results.....	159
Discussion .....	160
<b>5 SUMMARY AND FUTURE DIRECTIONS .....</b>	<b>166</b>
<b>LIST OF REFERENCES .....</b>	<b>173</b>

BIOGRAPHICAL SKETCH..... 203

## LIST OF TABLES

<u>Table</u>		<u>page</u>
1-1	Parvoviruses, their receptors, and hosts .....	50
1-2	VP2 amino acid differences between MVMP and MVMi .....	51
1-3	Forward mutations in the MVMi strain with a fibrotropic phenotype.....	51
2-1	Data processing and refinement statistics .....	96
3-1	Data processing and refinement statistics .....	139
3-2	RMSD in C alpha positions for the parvovirus structures .....	140
3-3	Amino acid sequence comparison for autonomous parvoviruses at nucleotide binding pocket .....	140
3-4	Amino acid sequence comparison in cytosine binding pocket.....	141
3-5	The ssDNA sequence present in H-1PV, CPV and MVMi .....	141
4-1	Data processing and refinement statistics .....	163

## LIST OF FIGURES

<u>Figure</u>		<u>page</u>
1-1	Genome architecture of the parvovirus genus .....	52
1-2	The parvovirus capsid VP structure and capsid surface topology. ....	53
1-3	Capsid features of <i>Parvovirinae</i> subfamily members.....	54
1-4	A schematic of the life cycle of parvovirus, MVM.....	55
1-5	Sialic acid binding site in MVMP .....	56
1-6	Structural clustering of MVMP/i amino acid differences .....	57
1-7	MVM capsid structures at ~3.5 Å resolution .....	58
1-8	Glycans recognized by the MVM viruses.....	59
2-1	Virus purification .....	97
2-2	Glycan binding sites on the MVMP and MVMP-K/I capsids .....	98
2-3	Glycan binding sites on the MVMi capsids .....	99
2-4	Superposition of glycans on MVM capsid .....	100
2-5	Sialylated derivatives recognized by the MVM viruses .....	101
2-6	Data from the SGM array .....	102
2-7	N-glycan expression on A9, EL4 T and NB324K cell lines .....	103
2-8	Polar and non-polar glycolipid expression on A9, EL4 T and NB324K cell lines .....	104
2-9	O-glycan expression on A9, EL4 T and NB324K cell lines.....	105
3-1	Virus purification, crystallization and diffraction .....	142
3-2	Glycan array data for H-1PV .....	143
3-3	Structure of H-1PV .....	144
3-4	ssDNA observed in H-1PV virions .....	145
3-5	ssDNA topology observed in members of the parvovirus genus .....	146

3-6	ssDNA-protein interactions in autonomous parvoviruses .....	147
3-7	Cytosines observed in the full H-1PV structure .....	148
3-8	Conformational variation between empty and full capsids .....	149
3-9	Location of the cis-peptide bonds observed in H-1PV structure .....	150
3-10	Sialic acid binding site on the H-1PV capsid .....	151
3-11	Comparison of VP2 structures of H-1PV, MVM, PPV, FPV and CPV.....	152
4-1	MVMP capsid dynamics on the exterior surface .....	164
4-2	Capsid dynamics in the interior of the MVMP capsid .....	165

## LIST OF ABBREVIATIONS

AAV	adeno-associated
ADV	aleutian mink disease virus
APS	Advanced Photon Source
BNL	Brookhaven National Lab
BPV1	bovine parvovirus 1
CAR	coxsackie and adenovirus receptor
CFG	consortium for functional glycomics
CHESS	Cornell High Energy Synchrotron Source
CryoEM	cryo-electron microscopy
CP	capsid protein
CPV	canine parvovirus
CsCl	cesium chloride
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
EDTA	ethylene diamine tetraacetic acid
EGFR	epidermal growth factor receptor
EM	electron microscopy
FGFR1	fibroblast growth factor receptor 1
FPV	feline panleukopenia virus
H-1PV	human tumor isolated 1 parvovirus
HBoV	human bocavirus
HGFR	hepatocyte growth factor receptor
HIV	human immunodeficiency virus
HS	heparan sulfate

HSPG	heparan sulfate proteoglycan
HSV	herpes simplex virus
ITR	inverted terminal repeats
kDa	kilodalton
LamR	laminin receptor
mRNA	messenger ribonucleic acid
MVM	minute virus of mice
Neu5Ac	N-acetylneuraminic acid
NES	nuclear export signal
NHS	N-hydroxysuccinimide
NLM	nuclear localization motif
NLS	nuclear localization signal
NPC	nuclear pore complex
NS	non-structural
NT	nucleotides
ORF	open reading frame
PBS	phosphate buffer saline
PDGFR	platelet-derived growth factor receptor
PEG	polyethylene glycol
PLA <sub>2</sub>	phospholipase A <sub>2</sub>
PPV	porcine parvovirus
RFU	relative fluorescence units
RMSD	root mean square deviation
RNA	ribonucleic acid
RT	room temperature

SIA	sialic acid
SGM	sialylated glycan microarray
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
ssDNA	single-stranded deoxyribonucleic acid
SV40	simian virus 40
TfR	transferrin receptor
VLPs	virus like particle
VP	viral protein
VP1u	VP1 unique N-terminal region
WT	wild type

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

STRUCTURAL STUDIES OF AUTONOMOUS PARVOVIRUS CAPSIDS TOWARDS  
UNDERSTANDING THE MECHANISMS OF CELL RECOGNITION AND ENDOSOMAL  
TRAFFICKING

By

Sujata Halder

August 2012

Chair: Mavis Agbandje-McKenna

Major: Medical Sciences – Biochemistry and Molecular Biology

Minute Virus of Mice (MVM) and H-1 Parvovirus (H-1PV), members of the *Parvoviridae*, share 66% capsid viral protein sequence identity, are oncotropic, and utilize cell surface sialic acid (SIA) for infection. Here, MVM and H-1PV served as models to characterize the molecular determinants of cell recognition and host range for the parvovirus capsid. In addition, MVM capsid dynamics associated with endosomal trafficking were probed.

MVM virus strains screened on a derivatized glycan microarray identified 9-O-methylation as an additional component of their SIA recognition, and glycan cell profiling confirmed a role for sialylated glycans in infection. Structural studies of MVM strains complexed with sialylated glycan receptors were used to define the role of specific amino acids in the interaction that dictates cell recognition. These residues were observed in the depression at the capsid icosahedral twofold axis and are associated with host pathogenicity determination. The structure of H-1PV, determined to aid the mapping of SIA site(s), is similar to that of other parvoviruses, especially MVM, except for differences at previously defined variable loops. H-1PV capsids screened on a

glycan array revealed specificity for α2-3 linked SIA glycans, including the SIA-Le<sup>X</sup> tumor cell marker motif, similar to previous observations for MVM. Structural studies of the H-1PV complexed with SIA glycans, including one with a 3'SIA-Le<sup>X</sup> motif, also identified the same twofold depression as MVM, as its SIA binding site. This observation indicates that parvovirus capsids utilize common regions as determinants of cell recognition, pathogenicity, and host range. Ordered ssDNA in a conserved binding site inside the H-1PV virion pointed to a potential role in packaging and capsid stability. Structural studies of MVM capsids at endosomal pHs showed two conformational changes: one in the capsid interior, which alters capsid-DNA interaction(s), possibly in preparation for DNA uncoating; and the other on the capsid exterior at the fivefold axis channel region increasing its diameter, possibly for VP1 N-terminus externalization. Information arising from these studies on cellular recognition and the specific capsid-receptor interactions that dictate oncotropism could aid the improvement of the efficacy of viral vectors based on these two viruses.

## CHAPTER 1

### INTRODUCTION

#### **Overview of Viruses**

Viruses are obligate intracellular parasites that are the causative agents of severe diseases and can infect all domains of life, from bacteria, archaea to eukaryotes. They have evolved a variety of strategies to exploit the host cell's replication machinery for a successful infection. Viruses may be of different shapes and sizes, but the basic structure called the nucleocapsid consists of the genomic material (DNA or RNA) enclosed within a protective protein shell (or capsid). Enveloped viruses incorporate lipids (and glycoproteins) from the host cell membranes during assembly as either an external envelope surrounding the nucleocapsid or internal envelope surrounding the genetic material, while non-enveloped viruses lack any lipid bilayer membrane. The shapes of viruses range from helical and icosahedral to more complex structures. The packaged genome encodes for the structural capsid/viral proteins (CPs or VPs) that make up the capsid and the auxiliary replication and regulatory proteins that are required for its propagation. Viruses may be classified on the basis of their morphology, including size, shape, capsid symmetry and presence or absence of an envelope; and biological properties, including genetic material, host range, replication strategy, mode of transmission, and pathogenicity.

The key initial step of virus infection is virus attachment onto cell surface receptors, followed by penetration through (non-enveloped viruses) or fusion with (enveloped viruses) the cellular plasma membrane to enter the cytosol, and culminates in the release and targeting of the viral genome to the host cell replication compartment. Many viruses with a DNA genome must enter the nucleus, whereas RNA viruses, with a

few exceptions, replicate in the cytosol. A number of factors, such as virulence, host permissivity, and the host immune response dictate the outcome of infection. Viruses, in particular those that undergo rapid mutation, such as RNA viruses, are also known to be able to switch hosts, which is a major health risk. Various host defense mechanisms exist that limit viral replication, but some viruses have evolved to circumvent these defenses and commandeer factors involved in host-defense or host-signaling pathways as part of their infection strategy. The VPs that assemble the capsid perform a wide variety of functions required during the viral life cycle, including host recognition, internalization, intracellular trafficking, genome replication, capsid assembly, genome encapsidation, and progeny virion release for re-infection. The high resolution three-dimensional (3D) images of viruses infecting phylogenetically diverse organisms or with different replication strategies and genome structures surprisingly show common structural features, implying conservation of commonly utilized structural motifs (reviewed in (162)).

The study of the virus life cycle and the intricate relationships with their hosts has resulted in several key developments. They have identified targets for the development of anti-viral drugs, such as the drug acyclovir that blocks transcription of *Herpesvirus* (HSV) DNA; and also vaccines such as the Sabin vaccine that stimulates the host immune system to develop long-term adaptive immunity against *Poliovirus* (107, 245). Much of the basic concepts and tools of molecular biology have been derived from the study of viruses, such as the discovery by Alfred Hershey and Martha Chase in 1952 that genes are composed of DNA, and the use of viruses as sources of enzymes and vectors for protein production. The integration of viral genome into host genome plays a

major role in host evolution and provides opportunities for new viruses to emerge that are optimized for replication in the host (reviewed in (103)). The relative ease of genetic manipulation of viruses has also provided insights into the host cell machinery. Viruses such as *Human Papillomavirus* (HPV) that induce host cell transformation, also serve as model systems to explore the role of complex signal transduction pathways in cell growth and differentiation, while viruses (e.g., *Myxoma virus*) that control abnormal cell growth are being developed as vectors for anti-cancer therapy (284, 292).

Understanding the mechanisms of various capsid-mediated interactions during an infection, such as cell surface receptor recognition, genome packaging, and host immune response would provide information necessary to design virus-based vectors that have specific tissue targeting capability, package a therapeutic gene, and have low immunogenic response. For example, mutating the capsid residues involved in binding to a specific receptor and inserting a peptide that recognizes a different receptor would allow the vector to bind alternate receptors, or mutating the residues that interact with a neutralizing antibody would now make the vector less immunogenic (10, 37). For the single-stranded (ss) DNA *Parvoviridae*, little is known about the structural determinants of the steps involved in infection, particularly for members of the *Parvovirus* genus. This project aims to begin to fill this gap by employing two related members, Minute Virus of Mice (MVM) and H-1PV, as models to dissect the molecular mechanisms of tissue tropism and capsid dynamics associated with endosomal trafficking.

### ***Parvoviridae Family***

The *Parvoviridae* consists of small (260Å), non-enveloped viruses that package an ~4-6 kb linear ssDNA genome within a T=1 icosahedral capsid (27). The parvovirus biology is dominated by its small size that restricts its coding capacity. They lack

accessory proteins that might induce resting cells to enter S-phase and also lack a duplex transcription template and therefore, must wait for the host to enter the S-phase before they can seize the cell's DNA replication machinery for their own replication. Although limited by their size and coding capacity and lack of any motility, they can infect a wide range of hosts, including insects, rodents, dogs, and humans. The *Parvoviridae* is highly diverse and consists of both pathogenic and non-pathogenic members.

### Taxonomic Classification

On the basis of host range, this broad family is divided into two subfamilies: the *Parvovirinae*, which infect vertebrates and include the five genera *Amdovirus*, *Bocavirus*, *Dependovirus*, *Erythrovirus*, and *Parvovirus*; and the *Densovirinae*, which infect insects and other arthropods and subdivided into four genera *Iteravirus*, *Brevidensovirus*, *Densovirus*, and *Pefudensovirus*. This study will focus on the members of the *Parvovirinae*, and especially the parvovirus genus. The type species of each of the *Parvovirinae* genera are; amdovirus: *Aleutian mink Disease Virus* (ADV); bocavirus: *Bovine Parvovirus 1* (BPV1); dependovirus: *Adeno-Associated Virus* serotype 2 (AAV2); erythrovirus: *Human Parvovirus B19* (B19V); and parvovirus: *Minute Virus of Mice* (MVM) (26). Members of the dependovirus genus rely on co-infection with a complex helper virus from a different taxonomic family (such as *Adenovirus* (Ad), HSV, HPV or *Vaccinia Virus*) for productive infection, and in the absence of the helper virus establish a latent infection in the host that can be re-activated by the introduction of helper virus (38). The other genera contain 'autonomous' parvoviruses that can replicate independently of helper viruses but require cellular factors expressed transiently during the S-phase for their DNA replication. Since these viruses replicate

productively only in actively dividing host cell populations, lethal infections occur in fetal or neonatal hosts, or involve tissues that remain actively dividing in adult life such as cells of the gut epithelium or leukocyte lineages. Pathogenic members, such as B19V and MVM cause severe disease in young and immunocompromised adults, while nonpathogenic members establish asymptomatic but persistent infections (232). The severity of disease depends on the virus and host factors, such as age and susceptibility. The parvovirus genus contains four distinct subgroups: (1) ‘rodent virus’ species that contains 3 clades (a) MVM, (b) *Mouse Parvovirus 1* (MPV1), (c) Rat virus group that includes *Rat Minute Virus 1* (RMV1), *H-1 Parvovirus* (H-1PV), *Kilham Rat Virus* (KRV), and *Lulli*, an ‘orphan’ virus; (2) an outlying *Rat Parvovirus 1* (RPV1) branch; (3) *Feline Panleukopenia Virus* (FPV) and *Canine Parvovirus* (CPV); and (4) *Porcine Parvovirus* (PPV) (293). As mentioned above, this study utilized MVM and H-1PV.

## **Genome Architecture**

The common genomic structure of parvoviruses consists of two open reading frames (ORFs) flanked by palindromic sequences (120 to ~550 nucleotides (NT) in length) that can fold into hairpin structures and are required for DNA replication and packaging. The terminal hairpins may be related in sequence and structure, and are referred to as inverted terminal repeats (ITRs) (e.g., members of the dependovirus and erythrovirus genera), or may be different in sequence and structure (e.g., members of the parvovirus genus). Unlike the dependoviruses which package both strands of the ssDNA genome with equal frequency into different capsids, some parvoviruses package predominantly the negative strand which is complimentary to the mRNA (e.g., MVM, CPV, FPV, PPV), while others encapsidate strands of either polarity in equimolar (e.g.,

AAVs, Lull) or different proportions (e.g., BPV1) (55). This differential encapsidation is dictated by the terminal hairpins. The 5' end ORF (*rep* or *ns*) in the complementary DNA strand encodes non-structural proteins (referred to as Rep in the dependoviruses and NS in the autonomous parvoviruses) that are important for genome replication and packaging, while the 3' ORF (*cap*) encodes structural viral proteins (VP) which assemble the capsid (76).

The MVM and H-1PV genome are organized into two transcription units with promoters located at map units (m.u) 4 (P4) and 38 (P38) (Figure 1-1) (14, 78, 239). The early P4 promoter codes for the non-structural regulatory proteins NS1 and NS2, derived from R1 and R2 transcripts, respectively. NS1(83 kDa) is a nuclear phosphoprotein that has helicase, ATPase, endonuclease, and sequence-specific DNA-binding activities essential for genome amplification and progeny excision (reviewed in (70, 218)) and is also the major mediator of cytotoxicity (12, 48, 80, 217). NS1 upregulates the P4 promoter itself in a positive feedback mechanism. The NS2 protein consists of three isoforms, NS2P, NS2Y, and NS2L (23 to 28 kDa) derived from alternate splicing of R2 transcript (71). NS2 is not essential for infection in certain non-murine permissive cells (such as transformed human cell lines), but is indispensable for the infection of natural murine hosts (51, 206). Multiple functions have been attributed to NS2 such as capsid assembly (66), DNA replication, message translation, virus production (206, 207), and nuclear egress of progeny virions (33, 96, 200, 220). NS2 also appears to be important in viral pathogenesis and tropism, since mutations in the MVM genome that resulted in increased NS2 levels contributed to host-range switching by MVM strains (63, 79), and because an NS2-defective virus was unable to infect

newborn mice (44). In addition, NS2 enhances NS1-mediated cytotoxicity (39, 176). The late P38 promoter is transactivated (~100 fold) by NS1 and drives the synthesis of R3 transcripts that encode VP1 and VP2 (58, 89, 295). VP1 (83 kDa) and VP2 (64 kDa) are produced by alternative splicing from the same mRNA and expressed at a ratio of 1:5, and capsid maturation results in the generation of VP3 (61 kDa) by post translational cleavage of ~20 amino acids from the N-terminus of VP2 following genome packaging (76). The entire sequence of VP3 is contained within VP2 which is in turn contained within VP1 that has a unique N-terminal region (VP1u) of 142 amino acids (295, 296). VP1u harbors an active phospholipase A<sub>2</sub> (PLA<sub>2</sub>) activity required for endosomal exit and important for infectivity (91, 102, 115, 191, 335). The cleavage of VP2 to VP3 can be mimicked by trypsin digestion, although peptide mapping indicates that this cleavage site is not the same as that *in vivo*. The same proteolytic site is present in VP1 but is inaccessible to cleavage, suggesting different structure disposition of these common sequences (296).

### Capsid Composition and Structure

**Capsid composition.** The number of VPs encoded by the *cap* gene and used to assemble the capsid differs between members of the *Parvovirinae* and also depends on the capsid maturation stage. ADV and B19V contain only VP1 and VP2; AAV2 and MVM virions contain VP1, VP2, and VP3; and BPV1 contains VP1, VP2, VP3, and VP4 (9, 68, 76, 143, 175, 256). Sixty copies of these VPs assemble the capsid with T=1 icosahedral symmetry (53). VP1 is always the minor component while the smallest VP is always the major component in all virus capsids. MVM empty capsids (devoid of DNA) contain VP1 and VP2 in a predicted ratio of 1:5, while full capsids (DNA containing) are composed of VP1, VP2 and VP3 in a ratio of 1:1:10 (295). Empty,

recombinant MVM virus-like particles (VLPs) that are morphologically and antigenically similar to native, empty capsids can be produced by baculovirus expression of the VP2 in Sf9 insect cells (132). The large-scale production of the VLPs enables their exploitation as vehicles for antigen presentation in vaccine development, and also for various structural and biophysical assays (16, 42, 168, 187, 260, 267, 268).

The parvovirus VP is capable of performing a variety of functions during the viral life cycle ranging from cell-surface receptor binding, endosomal entry and trafficking, cytoplasmic processing, nuclear entry, capsid self-assembly, genome encapsulation, capsid maturation, nuclear export of infectious virus progeny, and host immune response evasion (reviewed in (5, 6)). The relatively small genome size has allowed the use of genetic manipulation for the structure-function annotation of the VPs/capsid. The parvoviruses have evolved to utilize specific functional motifs in the N-terminal regions of VPs such as the phospholipase A<sub>2</sub> (PLA<sub>2</sub>) function in the VP1u for endosomal escape during infection (91, 102, 115, 191, 335), nuclear localization signal (NLS) in the VP1u and VP2 N-terminal region (in dependoviruses) for nuclear entry (122, 145, 186, 279, 310, 312), and nuclear export signal (NES) in the VP2 N-terminal region of certain autonomous parvoviruses for nuclear exit of progeny virions (194). The common VP1/2/3 core is involved in receptor attachment, tropism, nuclear entry, capsid assembly, and host antigenic response (5).

**Capsid structure.** Towards understanding the capsid structural features that dictate the multiple functions of the VP/capsid during cellular infection, the three-dimensional structures for several members of the *Parvoviridae* family have been determined using X-ray crystallography and/or cryo-electron microscopy (cryo-EM) and

image reconstruction (cryo-reconstruction) (3, 4, 7, 53, 119, 123, 158, 171, 177, 184, 198, 210, 214, 225, 275, 276, 306, 314, 328, 330). In all these structures only ~520 residues of the C-terminal overlapping VP region have been observed. The signal-rich N-terminal extensions of the VPs, including VP1u, ~40-60 residues of VP2 and the first ~15-24 residues of VP3 have been un-resolved, except for B19V. Low copy numbers of VP1 and VP2 in the mature virions or differential conformations adopted by the VP1/2/3 N-termini, which is incompatible with the 60-fold icosahedral averaging applied during structure determination, could result in the lack of N-terminal VP ordering (53). A conserved glycine-rich portion in the VP2 N-terminus has been modeled in the electron density observed within the fivefold pores of DNA-containing CPV and MVM virions (7, 306, 329). The un-observed N-terminal sequences are proposed to be positioned inside the capsid consistent with the location of the first N-terminal residue observed in crystal structures. Cryo-reconstruction studies in which the structures of empty AAV capsids containing VP1-3 were compared to capsids assembled without VP1 and/or VP2 and to infectious virions suggest that the N-terminal regions of VP1 and VP2 are located in the interior of the capsid underneath the icosahedral twofold axis (170, 225, 280). For B19V, cryo-EM reconstruction of wild type (wt) virions and empty particles and comparison to recombinant VP2 containing VLPs showed VP2 N-termini exposed on the capsid surface (155, 158). The N-terminal peptides perform essential functions at different steps in the infectious life cycle as mentioned before, and become sequentially exposed in response to successive cellular signals following virus internalization (35, 191, 254).

The structure of the ordered common VP region is highly conserved for the *Parvoviridae*, even for members that are only ~20% identical at the amino acid sequence level, such as AAV2 and B19V. The core is composed of an eight-stranded (designated  $\beta$ B to  $\beta$ I from N to the C-terminus) antiparallel  $\beta$ -barrel motif and an  $\alpha$ -helix ( $\alpha$ A) which are conserved in all parvovirus structures determined to date (Figure 1-2A and B). This  $\beta$ -barrel motif is also observed in other icosahedral viruses with very low sequence similarity (24). An additional  $\beta$ -strand,  $\beta$ A, involved in antiparallel interactions with  $\beta$ B, is also present in all the *Parvovirinae* structures. In the *Densovirinae*,  $\beta$ A strand undergoes domain swapping and interacts in an antiparallel fashion with the  $\beta$ B of the twofold related monomer (156, 275). Loop insertions between the  $\beta$ -barrel strands, which also contain small stretches of  $\beta$ -strand structure and a few  $\alpha$ -helices, form the remainder of the VP structure and contribute to capsid surface topology. The loops are named after the strands between which they are inserted, for example, the GH loop is inserted between  $\beta$ -strands  $\beta$ G and  $\beta$ H. These loops vary in length, sequence and structure between members in the same and different genera, and have been defined as variable regions (VRs) 1-8 for the members of the parvovirus genus (168) and as VR I-IX for members of the dependovirus genus (119). The  $\beta$ -sheet forms the contiguous shell, while the elaborate loop insertions form characteristic features at and around the icosahedral two-, three-, and five-fold symmetry axes (Figure 1-2C). There is (I) a depression (dimple) at each twofold axis, (II) a single protrusion (mound) at the threefold axis of members of the parvovirus genus, or three separate protrusions surrounding a depression at the threefold axis in members of the amdovirus, bocavirus, dependovirus, and erythrovirus genera, (III) a cylindrical channel at each fivefold axis,

(IV) a ‘canyon-like’ depression surrounding the cylindrical channel, and (V) a “wall” between the twofold and fivefold depressions (Figure 1-3). Exceptions to this general surface topology are seen in the *Densovirinae* members that have smoother capsids as a result of smaller loop insertions between their  $\beta$ -strands (53). The twofold axis (in all the parvovirus structures) is created by the loop after  $\beta$ I from two twofold symmetry related VP monomers. The conserved  $\alpha$ -helix forms the wall of this depression. The floor of the twofold depression is the thinnest region of the capsid, being only one polypeptide chain thick and contains the smallest number of intermonomer interactions (Figure 1-2C). The single threefold protrusion in members of the parvovirus genus are created from six loops (within the GH loop), two from each threefold symmetry related VP monomer; each of the three separate protrusions in the other *Parvovirinae* members are created by three loop regions (within the GH loop) from two VP monomers. The channel at the fivefold axis that connects the inside and the outside of the capsid is created by clustering of  $\beta$ -ribbons in the DE loop of five symmetry related monomers. The floor of the depression around the channel is lined by the HI loop that resembles a flower petal extending from the adjacent fivefold symmetry related monomer and forms the most extensive fivefold related VP contacts. A DNA binding pocket, conserved in the parvovirus genus, is observed in a depression bound by the BIDG  $\beta$ -sheet in the interior of the capsid. Based on the differences in the surface topologies, the parvoviruses were divided into three structural groups (225), and members of the *Parvovirinae* subfamily fall into group I and III, while group II comprised of members of the *Densovirinae* subfamily. Group I capsids have a mound like protrusion at the threefold axis and a wider and shallow twofold depression and are comprised of members of the parvovirus

genus. The group III capsids have three distinct spike-like protrusions at the threefold axis and narrower and deeper depression at the twofold axis and are comprised of members of the amdovirus, bocavirus, dependovirus, and erythrovirus genera. The protrusions are more pronounced in ADV and the AAVs compared to B19V and *Human Bocavirus 1* ((HBoV1), member of the bocavirus genus) . B19V and HBoV1 appear to share characteristics of members of the parvovirus and dependovirus genera and have flatter protrusions. The fivefold channel is conserved in the group I and III capsid structures. In B19V and HBoV1, the wall between the twofold and fivefold depressions are at almost the same height as its threefold protrusions, and form a continuous rim surrounding the canyon. Mutagenesis, biochemical, and structural studies have shown that the VRs play important roles in the viral life cycle, such as receptor recognition, tissue tropism, pathogenic outcome, and antigenic response (reviewed in (5, 6)).

### **Infectious Pathway of Parvoviruses**

The parvovirus capsids serve as robust delivery vehicles that penetrate two cellular barriers, the plasma membrane and nuclear membrane, and deliver the encapsidated genome to the nucleus for replication (Figure 1-4). Virions are stable in the presence of lipid solvents and survive exposure to pH between 3 and 9 and incubation at ~60°C for one hour (132, 168, 192, 213, 334). However, biochemical studies demonstrate that the capsid is metastable, and undergoes minor conformational transitions following receptor attachment and internalization to display functional motifs required to maneuver through the endocytic pathway to the nucleus for genome replication (35, 178, 191). Several research groups have characterized various steps in the life cycle of the parvoviruses using biochemical and cell-based assays but there is very little 3D information on the capsid structural dynamics involved in the infectious

process. The parvovirus structures discussed in the previous section represent only low-energy state conformations. The data presented in chapter 4 aims to fill this dearth by structural characterization of the capsid dynamics associated with endocytic trafficking. The various steps in the life cycle of the parvoviruses have been reviewed in the following sections.

## **Receptor Binding**

Recognition of cell surface receptor by a virus is the first step of infection and a key parameter of tropism and pathogenesis. The primary receptors and co-receptors utilized for cellular recognition and internalization, respectively, during parvoviral infection are listed in Table 1-1. For most of the parvoviruses only the glycan component of the glycoprotein or glycolipid receptor is known, hence the focus of this study will be on the glycoconjugate receptors. Glycans (carbohydrate polymers) are the major constituents of the cell surface and may be conjugated to cell surface proteins or membrane lipid head groups to form glycoproteins and glycolipids, respectively, or are present as glycosaminoglycan (GAG) chains attached to proteoglycans (222). The immense variability of the glycan structures expressed between different species and also between different tissues in the same species creates vast diversity in viral-tissue tropism. The most common glycoepitopes contain terminal sialic acid (SIA) or sulfated oligosaccharide motifs of GAGs (e.g., heparan sulfate (HS)). Sialic acids comprise a family of structurally diverse monosaccharides derived from neuraminic acid, a nine-carbon sugar. There are more than 50 natural analogues of SIA that result from modifications to the carbohydrate backbone, of which N-acetylneuraminic acid (Neu5Ac) is the most common. In general, the amino group attached to C-5 is N-acetylated (Neu5Ac), N-glycolylated (N-glycolylneuraminic acid; Neu5Gc), or removed

(deaminoneuraminic acid; Kdn). The hydroxyl groups may be free, esterified (acetylated, lactylated, sulfated, phosphorylated), or etherified (methylated), thus leading to increased chemical diversity (149). Structural heterogeneity among sialylated glycans can arise from variations in the glycosidic linkage positions associated with the Neu5Ac residues, which may be linked  $\alpha$ 2-3 or  $\alpha$ 2-6, principally to galactose (Gal) or N-acetylgalactosamine (GalNAc) residues, or linked  $\alpha$ 2-8 or  $\alpha$ 2-9 to adjacent SIA residues (11, 309). The carboxyl group at position 1 confers a negative charge on SIA under physiological conditions. They commonly occur at the non-reducing termini of oligosaccharide chains that are attached to glycoconjugates. HS is a highly sulfated, linear polysaccharide built up from disaccharide units of glucosamine and uronic acid. HS proteoglycans (HSPG) are the major components of extracellular matrix. The sulfate groups impart high charge density to HS and contribute to non-specific virus binding through electrostatic interactions (183). Binding to HS often involves positively charged patches on capsid surface (219, 223). However, in some cases, HS acts as a specific receptor, as in the case of HSV and *Foot and Mouth Disease Virus* (FMDV) (109, 273). A few examples of neutral virus-receptor glycans, such as histo-blood group epitopes have also been identified.

Biochemical studies utilizing neuraminidase and proteinase K treatment have shown that MVM requires SIA containing glycoproteins for cell recognition and infection (78, 188). SIA is also an important attachment factor for other parvoviruses infecting different species, such as BPV1 (28, 142, 297), H-1PV (10), PPV (34), AAV1 (324, 325), AAV4, AAV5 (148, 270, 315), CPV and FPV, although the SIA-CPV and SIA-FPV interactions are not essential for infection in certain cell types (19). CPV and FPV use

their respective transferrin receptor (TfR) on canine and feline cells for infection (19, 136, 230). Eliminating the glycosylation site on the canine TfR expanded its ability to bind both FPV and CPV (116, 117, 229). HSPG serves as the cell surface receptor for AAV2 and the closely related AAV3b (290). AAV6, which is closely related to AAV1 is able to utilize HS or SIA depending on the cell type being infected (324, 325). AAV9 utilizes terminal galactose as a receptor (23, 272). *Bovine AAV* (B-AAV) utilizes gangliosides for transduction (264) and chitotriose, a trimer of  $\beta$ 1-4 linked N-acetyl glucosamine found on gp96, for cellular transcytosis (85). B19V binds to the glycolipid erythrocyte P antigen (globoside) on erythroid progenitor cells but requires  $\alpha_5\beta_1$  integrin and Ku80 as co-receptors for cellular entry (43, 205, 316, 317). The protein receptors/co-receptors that have been identified for the parvoviruses, such as the  $\alpha_v\beta_5$  integrin and growth factor receptors are listed in Table 1-1 and reviewed in detail in Halder *et al.* (126).

The recognition sites for the glycoconjugate receptors, conformed only on the assembled capsid have been identified using biochemical, molecular, and structural approaches for several parvoviruses, such as MVM, CPV, FPV, B19V, AAV2, AAV5 and AAV1/AAV6. The depression at the twofold axis, also known as the dimple was identified as the SIA binding site on MVM, and this is proximal to the CPV and FPV determinants of SIA binding to erythrocytes (188, 303) (Figure 1-5). Significantly, the residues determining *in vitro* tropism and *in vivo* pathogenicity for MVM, such as K368 and I362 are localized in the vicinity of this SIA binding pocket. The CPV-TfR attachment site has been mapped to the wall between the two and fivefold depressions (118, 135, 136). Structural studies utilizing cryo-reconstruction mapped the globoside

receptor attachment site for B19V to the depression at the threefold axis (57, 154). The AAV2-HS binding site has been mapped to basic residues located at the wall of the protrusions surrounding the threefold axis (160, 178, 219, 223). Binding of B19V virions to globoside triggers the externalization of VP1u (35). Structural studies on the AAV2-HS complex reported structural rearrangement of the HI loop which was proposed to be related to the opening of the fivefold channel (178). Protease susceptibility assays did not detect any changes in the CPV capsid structure upon TfR binding (213).

The 3D structure determination studies of wt and mutant MVM capsid-glycan receptor complexes presented in chapter 2 attempts to provide a structural understanding of the role of the specific amino acid residues in receptor recognition and specificity. Data presented in chapter 3, identifies the sialic acid motifs recognized by H-1PV and examines the structural interactions involved in receptor attachment. This information allows for a comparison of the mechanisms of host recognition between the two members of the parvovirus genus and will also aid in the development of these viruses as vectors with specific tissue tropism.

### **Intracellular Trafficking and Structural Transitions Induced in the Virion**

MVM, CPV, AAV2 and AAV5 are the best characterized viruses with respect to endosomal trafficking. MVM will be the main focus of this section and will be compared to CPV, AAV2 and AAV5 when variations are observed in their pathways. Although the parvoviruses use different cellular receptors and the attachment sites vary, most of them are internalized by receptor-mediated endocytosis, predominantly by dynamin-dependent clathrin mediated endocytosis (34, 93, 231) (Figure 1-4). Alternate mechanisms of entry might be used; for example, AAV2 utilizes the clathrin-independent carriers/GPI-anchored-protein-enriched endosomal compartment (CLIC/

GEEC) pathway, and AAV5 utilizes the caveolar endocytic pathway (17, 215). Some of the AAVs have been demonstrated to use selective receptor-mediated vesicle transcytosis to penetrate barrier cell layers (83). AAV2-integrin binding and clustering has been shown to activate cell signaling pathways that enhance virus uptake (150, 261).

Particle-to-infectivity ratio of most parvoviruses seems to be high, implying that most particles that bind and enter the cell fail to navigate to the nucleus. Endosomal trafficking of parvovirus capsids is reported to be a slow and rate limiting process in viral transduction in several cell types (94, 127, 128, 131). Following uptake, parvoviral particles are trafficked through the endocytic pathway and delivered to early endosomes, late endosomes, recycling endosomes and lysosomes (20, 86, 87, 130, 191, 231, 253, 288). CPV capsids appear to remain associated with the TfR in recycling endosomes (230, 231, 288). This suggests that *in vitro* biochemical and structural studies characterizing the role of the capsid in the infectious life cycle be conducted in the presence of the receptors (if known). Confocal microscopy studies have shown that in addition to trafficking through the endocytic pathway, AAV2 can localize to the Golgi and endoplasmic reticulum prior to nuclear entry (144, 227), and similarly AAV5 has been detected in the Golgi compartment (17, 18). Further studies are required to characterize the role of capsid-host cellular factor(s) interactions in facilitating these alternative trafficking routes. Structural studies of capsid-receptor complexes in chapter 2 and chapter 3 are the first step towards fulfilling this aim. CPV and MVM move through the cytoplasm to the nucleus by microtubule mediated processes, as shown by the disruption of the microtubule network with nocodazole or microinjection of antibodies

against the microtubule based motor protein (129, 133, 191, 231, 253, 261, 286, 288, 311, 313). MVM and PPV have also been shown to be dependent on an intact microfilament network as disruption of the actin filament network with cytochalasin B/D and latrunculin treatment blocks cytoplasmic trafficking (34, 253, 261).

Endosomal acidification is known to be essential for the infection of all parvoviruses, since lysosomotropic drugs such as bafilomycin A1 and chloroquine, and NH<sub>4</sub>Cl that interfere with the endosomal pH, block infection (20, 21, 92, 128, 231, 253, 311). The low pH environment of the endosome triggers the externalization of the VP1u for its PLA<sub>2</sub> activity essential for endosomal exit, while the capsid remains intact (65, 91, 102, 115, 191, 283, 335). The PLA<sub>2</sub> activity is reported to induce a transient pore formation or permeability change in endosomal membranes rather than complete endosomal lysis as smaller dextrans (molecular weight of 3000) coendocytosed with CPV were released from the vesicles while large dextrans (molecular weight of 10,000) and α-sarcin were retained in the vesicles (231, 287). Treatment of CPV particles *in vitro* with the acidic pH of the endosome induces VP1u exposure (287). However, for MVM, cleavage of VP2 N-termini to VP3 is a prerequisite for VP1u externalization at low pH *in vitro* (65, 100). And, in the case of AAV, low pH alone is not sufficient to mediate VP1u extrusion, suggesting a requirement for yet unidentified cellular factors in this capsid transition (170). The VP1u externalization in MVM, CPV and AAV2 may be mimicked *in vitro* by exposure to heat or urea (65, 67, 170, 312). Heat treatment of AAV virions to 65° C and empty capsids to 75° C is able to mimic endosomal conditions required to release the VP1u (170). For MVM, VP1u eviction is only seen in full capsids but not in empty particles when heated *in vitro*, however a recent report observed VP1u

extrusion in empty capsids by *in vivo* immunofluorescence staining and *in situ* hybridization (65, 191). The differences observed in the *in vitro* and *in vivo* studies for MVM, suggests that artificial heat treatment experiments cannot directly reproduce the *in vivo* stimuli but have enabled biochemical and structural characterization of capsid dynamics (65, 67, 170, 312). A combination of factors, such as pH, receptor binding or interaction with other host factors is implicated to play a role in these capsid conformational changes. It was earlier believed that the VP1u of B19V is always exposed on the capsid surface but recent studies show that it is true only for B19V VLPs, and the VP1u of B19V virions is buried but can be exposed *in vitro* by heat or low pH treatments (254, 257). Although the extrusion of VP1 N-terminus is a common feature among all parvoviruses and is essential for infectivity, the exact mechanism that triggers this event is still unclear. As has been reported for other PLA<sub>2</sub> domains, the VP1u is predicted to be α-helical in nature (170) . Mutagenesis and biochemical studies indicate that the channel at the fivefold axis serves as the VP1u extrusion route for AAV2 and MVM (32, 100, 101). Such dynamic flexibility has been reported for other viruses, such as *Poliovirus*, where externalization of VP1 N-terminus upon receptor binding is essential for infectivity (36, 108) and the myristoylated VP4 polypeptide found on the inside surface of the virion is released early in infection while the particles remain intact (179). Structural studies of AAV8, CPV and MVM capsids (in chapter 4 of this study) under pH conditions that mimic the environment encountered in the endocytic pathway have detected conformational changes which will be discussed in detail in chapter 4 (208, 274).

Following endosomal release, there is likely further processing in the cytoplasm because microinjection of virions with/without exposed N-termini does not confer a nuclear translocation phenotype (231, 279, 287, 288). It has been shown that phosphorylation of capsid surface exposed tyrosines followed by ubiquitination targets AAV2 and AAV5 capsids for proteasomal degradation (332, 336, 337). Thus, co-administration of proteasome inhibitors enhances AAV2 and AAV5 transduction efficiency (92, 94, 203, 331-333). However, for MVM, CPV and PPV, treatment with proteasome inhibitors was detrimental to infection, and specifically, the chymotrypsin-like activity of proteasome appeared necessary for infection (253, 255). PPV capsid proteins were ubiquitinated early during infection but no evidence of particle ubiquitylation or degradation was observed for CPV and MVM (34, 253, 255).

The exact mechanism of genome translocation into the nucleus following endosomal release is unclear for the parvoviruses, as are the determinants of capsid uncoating. The VP1 N-termini in MVM (and the VP1/2 N-termini in AAVs) contains nuclear localization signals (NLSs) required for nuclear trafficking (122, 145, 186, 279, 312, 323). Theoretically, the 260 Å diameter capsids of the parvoviruses should be able to pass through the nuclear pore complex (NPC) intact, and this has been seen for CPV capsids during nuclear entry and for newly synthesized MVM capsids during viral egress (194, 200, 313). Other studies report that virus uncoating occurs during endosomal trafficking and the viral DNA enters the nucleus devoid of viral protein (130, 191). There is also evidence indicating that MVM might not enter the nucleus through the NPC, and an alternative nuclear entry strategy involving partial disruption of the nuclear membrane by host caspase 3 has been proposed (59-61). Similarly, for AAV2,

majority of the studies indicate that genome uncoating occurs in the nucleus (20, 146, 261, 279, 326), however there are reports that uncoating may occur before or during nuclear entry (190).

MVM is the best characterized member with respect to capsid requirements of genome release. Studies with MVM suggest that the low endosomal pH and divalent cation depletion in the cytoplasm together with receptor binding induce 3' to 5' genome uncoating thus exposing the 3' end of the viral DNA to polymerases (67). The pH dependence of the simultaneous externalization of VP1u and viral DNA from intact capsids without particle disassembly has been observed *in vivo* (191). Similarly, *in vitro* heat or urea treatment induces VP1u exposure and genome release from intact capsids (65, 100, 312). Mutational studies of the residues at the base of the fivefold channel control genome uncoating, thus, implicating the cylinder as the likely genome extrusion portal (74, 100).

Following genome delivery, NS1 (for autonomous parvoviruses) or Rep78/68 (for AAVs) initiated replication proceeds via a series of duplex intermediates in a rolling-hairpin mechanism primed by replication origins at each end of the linear genome (70). The replicated genome is transcribed, and then translated in the cytoplasm and the VPs are reported to be transported back into the nucleus in the form of intermediates for capsid assembly. Mutagenesis and structural studies suggest that VP trimers are the stable assembly intermediates for MVM capsids (236, 246, 248). The NLS in the VP1u and the structural nuclear localization motif (NLM) localized in the common β-barrel region of the MVM capsid, target the expressed VPs to the nucleus (185, 186). In the case of AAVs, a transiently expressed 23kDa protein called assembly activating protein

(AAP) targets VPs into the nucleolus where capsid assembly is proposed to occur (280, 281). Mutagenesis studies have identified charged residues involved in symmetry interface interactions (6) and residues in the HI loop to be important for capsid assembly (88, 122).

MVM and AAV2 are the best characterized parvoviruses with respect to genome packaging. The replicated DNA is packaged in a 3' to 5' direction into the preformed empty capsid using energy provided probably by NS1 or Rep52/40 and requires the terminal hairpins as packaging signals (31, 54, 69, 95, 164, 327). It has been shown that ~24 nucleotides of the 5' end of the packaged viral genome is outside the virions and covalently attached to NS1 or Rep (69, 241). Mutational analysis on MVM implicate the fivefold channel as the portal for genome encapsidation, in addition to VP1u exposure and genome release (32, 101, 240, 323). For AAV2, in addition to the fivefold channel, the DE and HI loop are also shown to be involved in genome packaging (32, 323). Crystal structures of MVM and CPV infectious virions have shown the presence of ordered genomic ssDNA inside the capsid with the bases interacting with side chains of the capsid amino acid residues (7, 306). The DNA-capsid interactions involve residues that surround the interior capsid surface near the two- and threefold axes and mutagenesis of these residues impaired infectivity (247). Unlike in AAVs, there is no ordered DNA in empty capsid structures in MVM or CPV which suggests that the capsid-genome interactions in MVM and CPV are specific (7, 119, 177, 210, 214, 329). This explains the low tolerance of autonomous parvoviruses, such as MVM and H-1PV as compared to the AAV, to replacement of wild-type genome sequence with a foreign sequence for vector production (161, 174). In chapter 3, the ssDNA-capsid interactions

in the context of H-1PV will be analyzed in an effort to provide an understanding of genome packaging which would also aid in the engineering of improved H-1PV based gene therapy vectors.

Post-capsid assembly and genome packaging, the virions must exit the nucleus and traffic to the cell surface for a second round of infection. The exit mechanism for the parvoviruses is best characterized for MVM. The VP2 N-termini possesses a phosphoserine-rich nuclear export signal (NES) that interacts with the exportin molecule chromosome region maintenance 1 protein (CRM1) to traffick packaged virions out of the nucleus (193, 194). Full capsids are released from the cell with all their VP2 N-termini intact but capsid maturation in the extracellular environment or during re-entry into a new host cell involves the proteolytic removal of ~25 amino acids from most of the VP2 N-termini to generate VP3, thus removing the nuclear export signal (100, 194). It has been shown that the VP2 N-terminus is exposed in DNA-containing capsids prior to any treatment and is susceptible to trypsin digestion. However, in empty capsids and VLPs, this sequence is sequestered but heat treatment can trigger its exposure (65, 132, 191). It has been suggested that following DNA packaging, a structural shift occurs in the capsid leading to the exposure of VP2 N-termini. The VP2 to VP3 cleavage primes the pH dependent externalization of the VP1u and genome as shown by the studies on the fivefold channel base mutants of MVM (100), suggesting that the channel serves as the site for VP2 externalization for cleavage to VP3 in autonomous parvoviruses that undergo a maturation step. Interestingly, cryo-electron EM and X-ray crystallographic studies of B19V reveal that its fivefold channel is narrower, consistent

with the observation that B19V does not undergo VP2 to VP3 cleavage and that the VP2 N-termini is exposed on the capsid surface (3, 57, 155).

From the available parvovirus capsid structures at physiological pH it is understood that the N-terminal extensions of VP1 and VP2 cannot be externalized without structural rearrangement. Although it is appreciated that capsid dynamics play a central role in parvovirus cell binding, entry, trafficking, genome release, and egress following assembly, few studies have addressed the three-dimensional visualization of these structural transitions. Chapters 2 and 3 focus on identifying the capsid regions involved in receptor binding, and chapter 4 characterizes the structural transitions encountered in the endocytic pathway.

### **Tissue Tropism and Pathogenicity Determinants**

The parvovirus genus includes viruses that infect many different species, including mice, cats, dogs, and pigs, and share medium (~50%) to very high (~98%) sequence identity. For several members of the parvovirus genus, there are distinct tissue tropism and pathogenic differences between highly homologous strains. FPV infects cats but not dogs (136, 230, 235, 304). CPV-2 emerged as a host range variant of FPV in 1978 and could infect dogs but not cats (136, 304, 305). In 1979, CPV-2 was replaced by an antigenic variant called CPV-2a that could infect both dogs and cats (137, 228, 233, 234). Since the 1980s, several antigenic variants of CPV-2a, designated CPV-2b and CPV2c that share the same host range have been identified (233). CPV can bind to both canine and feline TfR while FPV can only bind feline TfR. However, alteration of the glycosylation site (N383K) in the apical domain of canine TfR enabled it to bind to FPV (228, 229). Several PPV strains have been distinguished by pronounced differences in their tissue tropism and *in vivo* pathology, although their VP2 proteins are

~99% identical (25). NADL-2 is the attenuated vaccine strain of PPV that is non-pathogenic, but can be lethal if injected *in utero*, while Kresse, IAF-A54, IAF-76, and NADL-8 are the virulent strains (25). In the case of ADV, the highly pathogenic ADV-Utah 1 strain replicates poorly in cell culture, whereas the non-pathogenic ADV-G strain replicates permissively in Crandell feline kidney (CrFK) cells.

MVM, the focus of this study, has two well characterized strains, the prototype strain (MVMp) and the immunosuppressive strain (MVMi) that are reciprocally restricted for growth in each other's cell type, despite sharing 97% sequence identity and being serologically indistinguishable. MVMp was originally isolated from a murine adenovirus stock and replicates efficiently in A9 mouse fibroblasts, whereas MVMi, which was recovered from an EL4 T-cell lymphoma, replicates in mouse T lymphocytes and hematopoietic precursors (294). However, both viruses can replicate efficiently in the human transformed cell line, NB324K (SV40 transformed human newborn kidney fibroblast cells). The viruses differ by only 14 amino acids (of 587) in VP2, all of which are ordered in the crystal structure except residue 10 (Table 1-2). Twelve of the 13 ordered residues are located throughout the primary amino acid sequence but are clustered on the assembled 3D capsid at and around the icosahedral twofold axis and on the shoulder of the threefold protrusions (Figure 1-6). These regions also differ between other highly homologous parvoviral strains, such as wt CPV and wt FPV, or wt CPV and its host range mutants.

Mutagenesis and selective plaque assays map the MVM *in vitro* tropism determinant to residues 317 and 321 in VP2 (15, 112, 196) and forward second-site mutations conferring fibrotropism to MVMi to residues 399, 460, 553, and 558 when

either 317 or 321 are mutated (7) (Table 1-3). The switch to fibrotropism for MVMi requires both an equivalent region of the MVMi capsid protein gene and a segment of the non structural protein genes that results in an increase in NS2 levels (63, 79). The MVMi strain, inoculated by the oronasal route in newborn mice is pathogenic and replicates in endothelia, neuroblasts (244), and hematopoietic stem cells, and in adult Severe Combined Immunodeficient (SCID) mice, causes acute leucopenia (269), while MVMP infection is asymptomatic (163). However, the MVMP intravenously inoculated into SCID mice, evolved into virulent variants which carried one of the three mutations (V325M, I362S, and K368R) in the VP capsid protein, which caused a systemic lethal disease when introduced by the oronasal route (259). Residues 362 and 368 differ between MVMP and MVMi, but these variants remained fibrotropic *in vitro* without any genetic changes in the capsid gene that control MVM tropism *in vitro*. The reintroduction of the highly virulent MVMP mutants into SCID mice caused lethal leukopenia, reflecting the pattern of MVMi infection (189).

Significantly, the residues determining *in vitro* tropism (317 and 321), *in vivo* pathogenicity (325, 362, 368), conferring fibrotropism on MVMi (399, 460, 553, 558), and those associated with the development of leukopenia (321, 551, 575) show local surface structure variability between the strains and are localized in the vicinity of the sialic acid binding pocket at the twofold axis (188, 189) (Figure 1-5, 1-6 and 1-7). For PPV, CPV, and FPV, tissue tropism and pathogenicity determinants have also been mapped to this region (reviewed in (5)), thus highlighting the utilization of common autonomous capsid regions for similar functions. Also, the residues controlling CPV and

FPV host range and virulence show local structural variations on the capsid surface and colocalize with the predicted transferrin receptor binding site (135, 136).

The co-localization of structurally variable tropism/pathogenicity determinants with receptor recognition sites suggests that the disparities in phenotypes involves differential MVM-glycan receptor interaction and/or utilization. Previous cell binding studies have shown that both MVM strains compete for binding to either cell type, arguing against restriction at the level of a cell surface receptor (282). Hybrid cells from fibroblasts and lymphocytes can propagate both strains of MVM (294), suggesting that the block is due to the lack of a differentiation-dependent cellular factor in restrictive cells. It was suggested that the point of restriction was post entry and conversion of genomic ssDNA to replicative form (RF) intermediates, but prior to viral genome transcription (134). It has also been suggested that the block is prior to RF DNA replication and likely due to a block in uncoating (242). However, glycan array screening and affinity assays have demonstrated (as detailed in the next section) that MVMP, MVMi and the virulent MVMP mutants show differences in SIA receptor interactions and specificity (188, 189, 209). The tissue tropism and pathogenicity of a virus may be regulated by virus-receptor interactions or post-entry virus-host cellular factor interactions. The work in chapter 2 is aimed at structurally characterizing the role of minor MVM capsid variations in regulating receptor recognition and dictating cell recognition and pathogenicity.

### **Receptor Binding and Virulence Determination**

The MVMP and MVMi strains have served as ideal models to address questions related to the role of cell surface receptor recognition in determination of tissue tropism and pathogenicity between highly homologous parvovirus strains. These two viruses

have been shown to require cell surface SIA for infection as well as to compete with each other for cell binding even in the restrictive cell line (77, 282). Thus, to identify possible differences in SIA structures or other carbohydrates recognized by the MVM viruses, the interactions of VLPs of MVMP, MVMi, and the three MVMP virulent mutants, MVMP-I362S, MVMP-K368R, and MVMP-I362S/K368R (MVMP-K/I), were studied on a glycan microarray with 180 different naturally occurring and synthetic glycans (209). All of the MVM viruses specifically recognized  $\alpha$ 2-3 sialylated glycans linked to a common Gal $\beta$ 1-4GlcNAc moiety: Neu5Aca2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-4Gal $\beta$ 1-4GlcNAc (3'SIA-LN-LN or 3'SIA(LN)<sub>2</sub>), Neu5Aca2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc (3'SIA-LN-LN-LN or 3'SIA(LN)<sub>3</sub>), and Neu5Aca2-3Gal $\beta$ 1-4(Fuca1-3)GlcNAc $\beta$ 1-3Gal $\beta$ 1-4(Fuca1-3)GlcNAc (3'SIA-Le $\times$ -Le $\times$ -Le $\times$  or 3'SIA(Le $\times$ )<sub>3</sub>). In addition, MVMi showed expanded recognition to multisialylated glycans with terminal  $\alpha$ 2-8 linkages: Neu5Aca2-8Neu5Aca2-8Neu5Aca ((SIA)<sub>3</sub>), Neu5Aca2-8Neu5Aca2-3Gal $\beta$ 1-4Glc (GD3), Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3Gal $\beta$ 1-4Glc (GT3), and Neu5Aca2-8Neu5Aca2-3(GalNAc $\beta$ 1-4)Gal $\beta$ 1-4Glc (GD2) (Figure 1-8). Interestingly, the virulent MVMP-K368R mutant that contains the MVMi residue type also recognized GT3, indicating that the residue 368 plays a role in the multisialic acid recognition observed for MVMi.

Notably, other than the extended recognition of multisialic acid by MVMi and MVMP-K368R, all the capsids showed a specificity for  $\alpha$ 2-3 linked terminal SIA and not  $\alpha$ 2-6 linked SIA. The 3'SIA-LN motif is commonly found in N- and O-linked glycoproteins, which are abundant in most cell surface glycoproteins. The recognition of the 3'SIA-Le $\times$  motif, which is a known tumor cell marker by all MVM viruses, explains

MVM's tropism for transformed cells (139, 151, 152, 182). Interactions with the multisialylated  $\alpha$ 2-8 glycans that are abundant in brain glycoproteins likely mediate MVMi's neurotropism (138, 262, 263). Also, GD3 gangliosides and 9-O-acetylated form of GD3 are present on a majority of T lymphocytes and may mediate MVMi's tropism for these cells (105).

Based on cell binding assays and neuraminidase treatments, *in vitro* cytotoxicity in A9 fibroblasts, glycan array screening and BIACore surface plasmon resonance (SPR) studies, it was shown that the virulent MVMP mutants had a lower affinity for the sialic acid component of the receptor, than MVMP or MVMi (188, 209, 259). The double recombinant mutant MVMP-I362S/K368R showed the lowest affinity but increased pathogenicity. This suggests that the affinity/avidity of the interactions with the SIA containing receptor modulates parvovirus virulence. Similarly, in *Polyomavirus* single amino acid changes that reduced SIA receptor affinity increased viral spread and disease severity (22). There are other viruses, including *Influenza Virus* and *Theiler's Virus* where a single or few amino acid changes modulate receptor binding and virulence. For Influenza virus, a single amino acid mutation in the receptor binding pocket of hemagglutinin causes the virus to switch specificity from Neu5Aca2-3Gal to Neu5Aca2-6Gal-terminal residues (250), and capsid mutations affecting SIA binding control neurotropism in the Theiler's Virus (141). Since, the K368R mutant bound to GT3 with lower affinity than MVMi, but the MVMP-I362S/K368R double mutant didn't bind to GT3, it indicated that more than a single amino acid substitution and the other 14 amino acids that differ between MVMP and MVMi near the receptor binding site, may be involved in the additional recognition of  $\alpha$ 2-8 linked sialic acids by MVMi. Also, the

virulent MVM<sub>p</sub> mutants remained fibrotropic *in vitro*, suggesting a difference in virulence pattern between the MVM<sub>p</sub> mutants and MVM<sub>i</sub>.

In a recent study to characterize the adaptive host range of MVM, capsid adaptations at the SIA binding site (mutations at residues 334, 384, 554, 578) resulted in the emergence of a host range variant, called F1, which was able to infect rat fibroblasts (Figure 1-6) (99). This observation suggests that in addition to controlling virulence and pathogenic outcome, the MVM receptor binding pocket also plays a role in adaptations to a new host. Chapter 2 presents data on the structural verification of the binding sites of the glycan receptors recognized by MVM<sub>p</sub>, MVM<sub>i</sub> and the MVM<sub>p</sub> virulent mutants in an effort to define the specificity and role of the capsid-glycan interactions in the infectious life cycle. Also, the glycan composition of the three cell lines that are permissive or restrictive for infection by MVM<sub>p</sub> or MVM<sub>i</sub> was analyzed in chapter 2 to correlate the natural expression of the glycans identified in the glycan array screening to the differences in tissue tropism.

### **Parvoviruses as Gene Therapy Vectors**

The autonomously replicating parvoviruses predominantly propagate in rapidly dividing cells due to their dependence on both cellular proliferation factors expressed transiently during the S-phase and the differentiation of host cells. These viruses, for example, MVM<sub>i</sub> and H-1PV, were first isolated from tumor tissue and were then believed to be oncogenic (300). But, it was later observed that the rodent parvoviruses display oncopreferential cytotoxic activity *in vitro* and also possess an oncosuppressive potential, inhibiting the formation of spontaneous and chemical or virus-induced tumors *in vivo* and *in vitro* (226, 251). The cytotoxicity has been attributed, in part, to NS1 (39). The intracellular oncotropism is mainly based on enhanced P4 promoter activity via

binding of ATF/CreB, Ets and SP1 transcription factors in transformed cells (110, 237).

Since transformed cells are intrinsically deficient in antiviral mechanisms, MVM's oncotropism might be related to its failure to mount an antiviral type I interferon (IFN) response in transformed cells, similar to other oncotropic viruses (120, 147). Thus, it is already known that specific host cellular factors are involved in the oncotropic properties of MVM and H-1PV but the mechanisms underlying tumor cell recognition which is the first step towards a successful infection, and oncosuppression are not yet fully understood. It has been shown that MVM viruses recognize the 3'SIA-Le<sup>x</sup> tumor cell marker and in chapter 2 the interactions of this glycan with the MVM capsid surface residues were analyzed (209). Chapter 3 presents data towards identifying the glycan receptor that explains H-1PV's recognition of tumor cells and structural characterization of the capsid-glycan interactions.

The rodent parvoviruses, such as MVMP, H-1PV and Lulli can persistently infect their natural hosts, do not integrate their genome into cellular chromosomes, are non pathogenic in adult animals, and efficiently infect human cell lines (reviewed in (30)). These characteristics make these viruses attractive candidate vectors for anticancer gene therapy, particularly for cytoreductive and immunogene therapy approaches to target tumor cells. H-1PV based vectors are currently being tested for antineoplastic effects in preclinical studies (1, 172). The AAVs have low toxicity, no known pathology, broad tropism and can establish long-term transgene expression, thus making them also an excellent choice for gene therapy vectors. Currently clinical trials are underway with AAV vectors packaging therapeutic genes for the treatment of several diseases,

including Leber's congenital amaurosis, hemophilia B, cystic fibrosis, Alzheimer's disease, arthritis, lipoprotein lipase deficiency, Parkinson's disease (204, 211, 285).

### **Significance**

The structures of the capsid VPs that are assembled into virions control various steps in the infectious life cycle, such as, host cell receptor recognition, endosomal trafficking, capsid assembly, and genome encapsidation. For the homologous MVM strains minor structural variations due to the amino acid differences in the capsid VPs result in altered virus-SIA receptor interactions and utilization which results in pronounced differences in their host range, tissue tropism and pathogenic outcomes of infection. Several research groups have characterized some of the steps in the infectious life cycle but there is very little structural information on the capsid dynamics or the capsid regions involved in the structural transitions associated with the life cycle. This project aims to utilize MVM and H-1PV as models to decipher the capsid dependent mechanisms for host cell recognition and structural transitions required for a successful infection.

In chapter 2, structure determination studies on the capsid-receptor complexes for the MVM strains and mutants was conducted to map the receptor binding site on the MVM capsids and provide further insight into the differential mechanisms of receptor recognition by these highly homologous strains and single site mutants that results in differences in tissue tropism and pathogenicity. Cellular glycan profiling analysis on the cell lines permissive for the MVM viruses was conducted to verify the expression of the glycans that had been shown to interact with MVM by glycan array screening and SPR studies. Also, MVM viruses were screened on an array consisting of SIA derivatives to provide further chemical information on the identity of the MVM glycan receptor. MVM

serves as a tractable model for studying emerging viral pathogens, allowing us to track subtle capsid changes that confer a disease phenotype as well as adaptations to a new host. Chapter 3 presents data on the identification of the glycan receptor for H-1PV and structural characterization of the capsid-glycan interactions and also capsid-genome interactions. Chapter 2 and 3 provide information on the capsid regions utilized for recognition of tumor cells by MVM and H-1PV and its role in dictating oncotropism. So, chapter 2 focuses on understanding the mechanism of differential tropism of MVMP (fibrotropic) and MVMi (fibrotropic and neurotropic), and chapter 2 and 3 utilize MVM and H-1PV as models to examine the mechanism of their oncotropism. A detailed understanding of MVM and H-1PV's oncotropism, will aid in the development of these viruses for tumor cell targeted gene delivery applications. In chapter 4, the dynamic character of the parvovirus capsid that allows for the sequential exposure of domains required to mediate successive interactions with the host, was probed by structure determination studies of MVM capsids at pHs (pH 6.0, 5.5, and 4.0) that mimic the endocytic pathway. The parallel study of these two members of the family and the homologous strains provides a comprehensive view of the mechanisms of infection and enables us to identify the common or divergent links. This study stems from the need to structurally characterize the mechanisms of host cell recognition and capsid dynamics associated with endosomal trafficking to fill in the information gaps in this field.

Table 1-1. Parvoviruses, their receptors, and hosts

Virus	Receptors	Co-receptors	Host	References
AAV1	$\alpha$ 2-3 and $\alpha$ 2-6 N-linked sialic acid		Human	(325)
AAV2	HSPG <sup>a</sup>	Integrin $\alpha$ 5 $\beta$ 1, $\alpha$ V $\beta$ 5, FGFR1 <sup>b</sup> , HGFR <sup>c</sup> , LamR <sup>d</sup>	Humans	(8, 13, 153, 243, 289, 290)
AAV3	HSPG <sup>a</sup>	HGFR <sup>c</sup> , LamR <sup>d</sup> , FGFR1 <sup>b</sup>	Humans	(8, 29, 180)
AAV4	$\alpha$ 2-3 O-linked sialic acid		NHP <sup>e*</sup>	(148)
AAV5	$\alpha$ 2-3 and $\alpha$ 2-6 N-linked sialic acid	PDGFR <sup>f</sup>	Humans	(84, 148, 270, 315)
AAV6	$\alpha$ 2-3 and $\alpha$ 2-6 N-linked sialic acid, HSPG <sup>a</sup>	EGFR <sup>g</sup>	Humans	(319, 324, 325)
AAV8		LamR <sup>d</sup>	NHP <sup>e</sup>	(8)
AAV9	Galactose	LamR <sup>d</sup>	Humans	(8, 23, 272)
Bovine AAV	Gangliosides, Chitotriose		Bovine	(85, 264)
ADV	ADV binding protein		Mink	(106)
BPV1	Sialic acid	Glycophorin A	Bovine	(28, 142, 297)
B19V	Erythrocyte P antigen	Integrin $\alpha$ 5 $\beta$ 1, ku80	Human	(43, 205, 316, 317)
MVM	Sialic acid		Rodent	(77, 209)
CPV and FPV	Sialic acid	Transferrin receptor	Cat, dog	(77, 136, 230)
PPV	Sialic acid		Swine	(34, 77)

<sup>a</sup>HSPG=heparan sulfate proteoglycan; <sup>b</sup>FGFR1=fibroblast growth factor receptor1;

<sup>c</sup>HGFR=hepatocyte growth factor receptor; <sup>d</sup>LamR=37/67-kDa laminin receptor;

<sup>e</sup>NHP=non-human primate; <sup>f</sup>PDGFR=platelet-derived growth factor receptor;

<sup>g</sup>EGFR=epidermal growth factor receptor. This table was adapted from Halder *et al.* (126).

Table 1-2. VP2 amino acid differences between MVMP and MVMi

	10	160	232	317	321	362	366	368	388	402	410	440	455	551
MVMP	Ser	Leu	Val	Thr	Gly	Ile	Val	Lys	Ser	Ser	Lys	Asn	Ala	Ala
MVMi	Gly	Ser.	Ile	Ala	Glu	Val	Met	Arg	Ala	Asn	Arg	Asp	Thr	Val
Position	N	S	S	SA	S	S	S	S	B	S	S	B	B	

\*S=Surface; B=Buried; SA=Solvent accessible; N=not ordered in structure

Table 1-3. Forward mutations in the MVMi strain with a fibrotropic phenotype

Site directed mutation	A317T	A317T	A317T	E321G	E321G	E321G
Selected forward mutation	D399G	D399A	D553N	A317T	S460A	Y558H

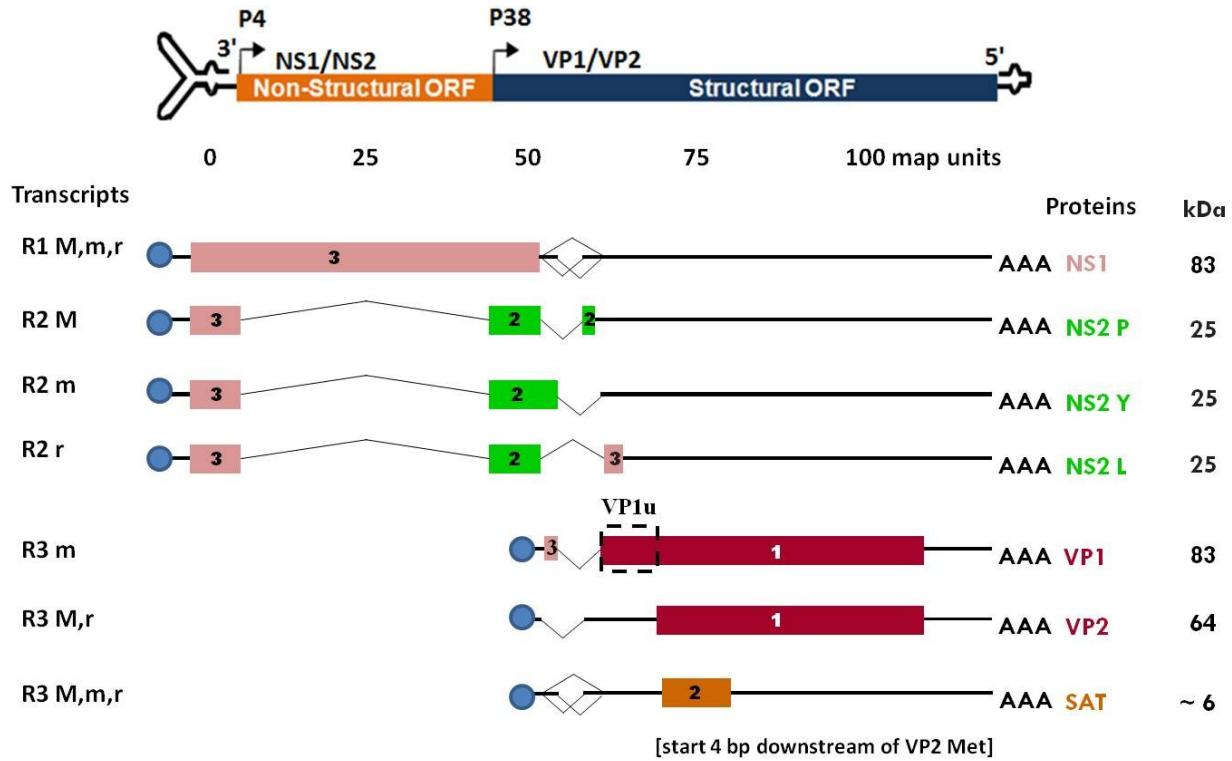
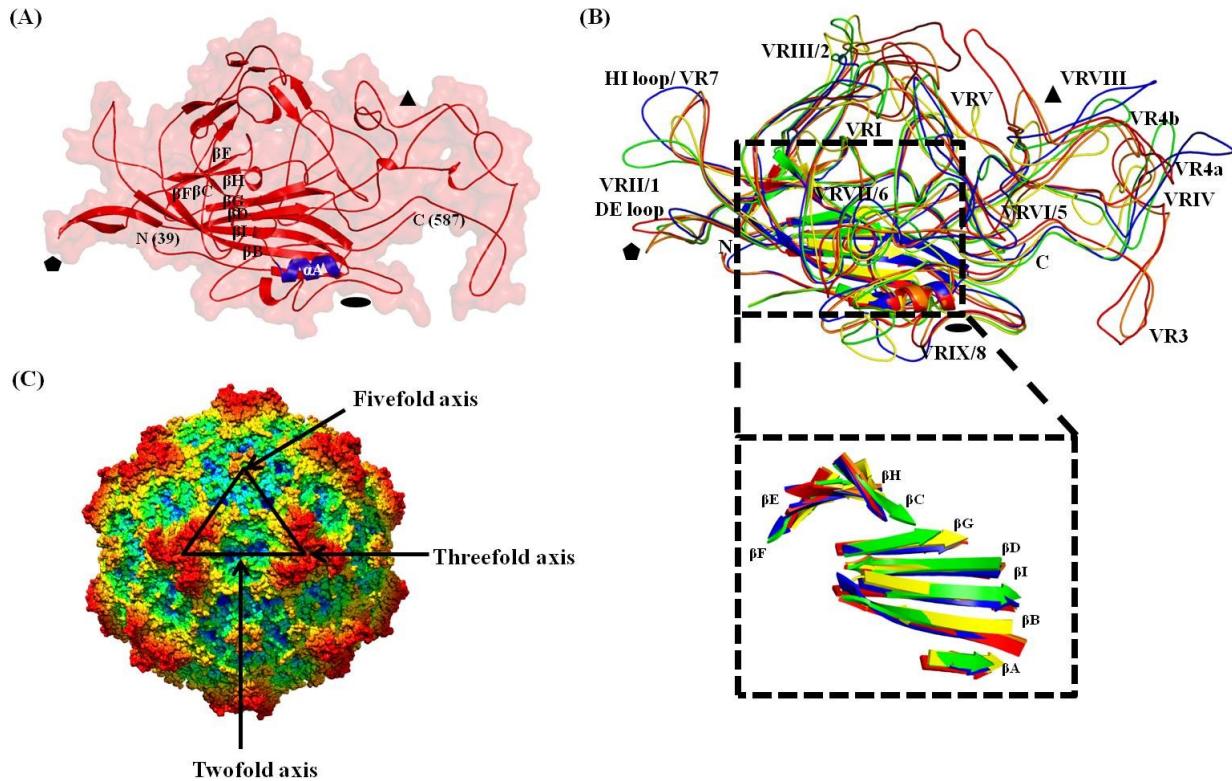


Figure 1-1. Genome architecture of the parvovirus genus. The single-stranded, negative-sense DNA genome of the parvovirus genome with the terminal folded hairpin structures is shown. The two viral promoters, P4 and P38 are shown by rightward arrows, and the mature, cytoplasmic transcripts R1, R2, and R3 are displayed below, with a blue sphere indicating the capped 5' ends and AAA denoting their polyadenylated tails. ORFs specifying the viral gene products, named on the right, are displayed in different shades according to their reading phase, and their spliced-out introns are represented by the thin-lined carets. The dashed box denotes the VP1u region involved in entry functions.



**Figure 1-2.** The parvovirus capsid VP structure and capsid surface topology. (A) Ribbon diagram of MVMp VP2 superimposed on a semi-transparent surface representation illustrating  $\beta$ -strand, helical, and loop regions. Conserved  $\beta$ -strands  $\beta$ B to  $\beta$ I,  $\alpha$ -helix A, and the first N-terminal residue modeled (39) and the C-terminal residue (587) are labeled. The approximate icosahedral twofold (filled oval), threefold (filled triangle), and fivefold (filled pentagon) axes are shown. (B) Conserved secondary structure superposition of VP for one member from every genus in *Parvovirinae* subfamily is shown: ADV (red), HBoV (yellow), AAV2 (blue), B19V (green) and MVMp (orange). Atomic coordinates for AAV2, MVMp, and B19V were obtained from RCSB protein database (PDB accession numbers 1lp3, 1z14, and 1s58, respectively). The ADV and HBoV images were generated from pseudo-atomic coordinates built into cryo-reconstructions (123, 198). The N-terminus (N), C-terminus (C), variable regions (VRI-IX, VR1-8), DE, and HI loops are labeled. The boxed region is shown below, depicting just the  $\beta$ A and  $\beta$ -barrel motif ( $\beta$ BIDG- $\beta$ CHEF) conserved in all parvovirus VP structures determined to date. (C) Surface representation of MVMp used to illustrate the topological features of the parvovirus capsid surface. The image is depth-cued (blue-cyan-green-yellow-orange-red) to show regions at the shortest radial distance to capsid center in blue and those at the furthest radial distance in red. A viral asymmetric unit is depicted by a black triangle bound by a fivefold axis and two threefold axes divided by a line drawn through the twofold axis. (A) and (B) were generated using the PyMol program (82), and (C) was generated using the UCSF-Chimera program (238).

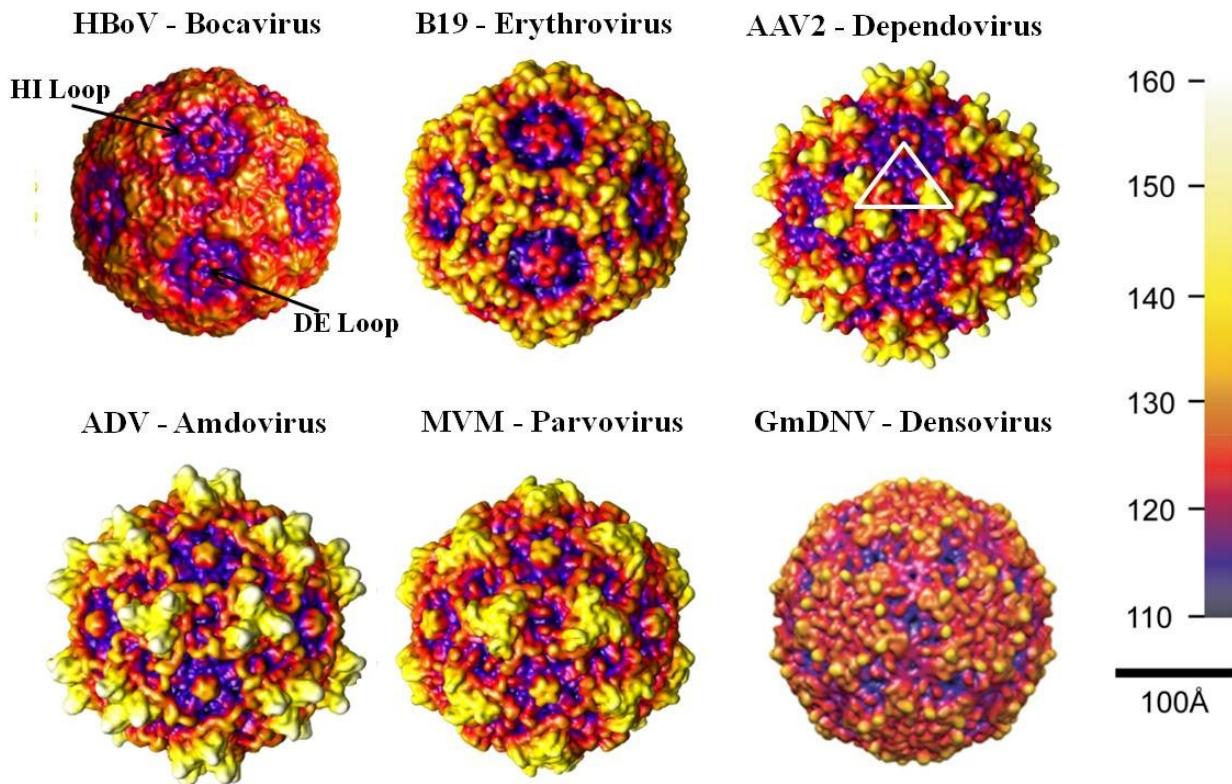


Figure 1-3. Capsid features of *Parvovirinae* subfamily members. Depth cued (blue-red-yellow-white) capsid surface representation of representative members of the five genera of *Parvovirinae* viruses, and one member from the *Densovirinae* is shown. The virus and the genus to which it belongs are labeled. A viral asymmetric unit (white triangle) is shown on the AAV2 image. A horizontal scale bar (100 $\text{\AA}$ ) for diameter measurement is shown on the right-hand-side and vertical color bar for radial distance ( $\text{\AA}$ ) from the center of the particle is also shown on the right-hand-side. These figures were generated using the UCSF-Chimera program (238). The coordinates used were obtained as described in the legend of Figure 1-2. This figure was adapted from (123).

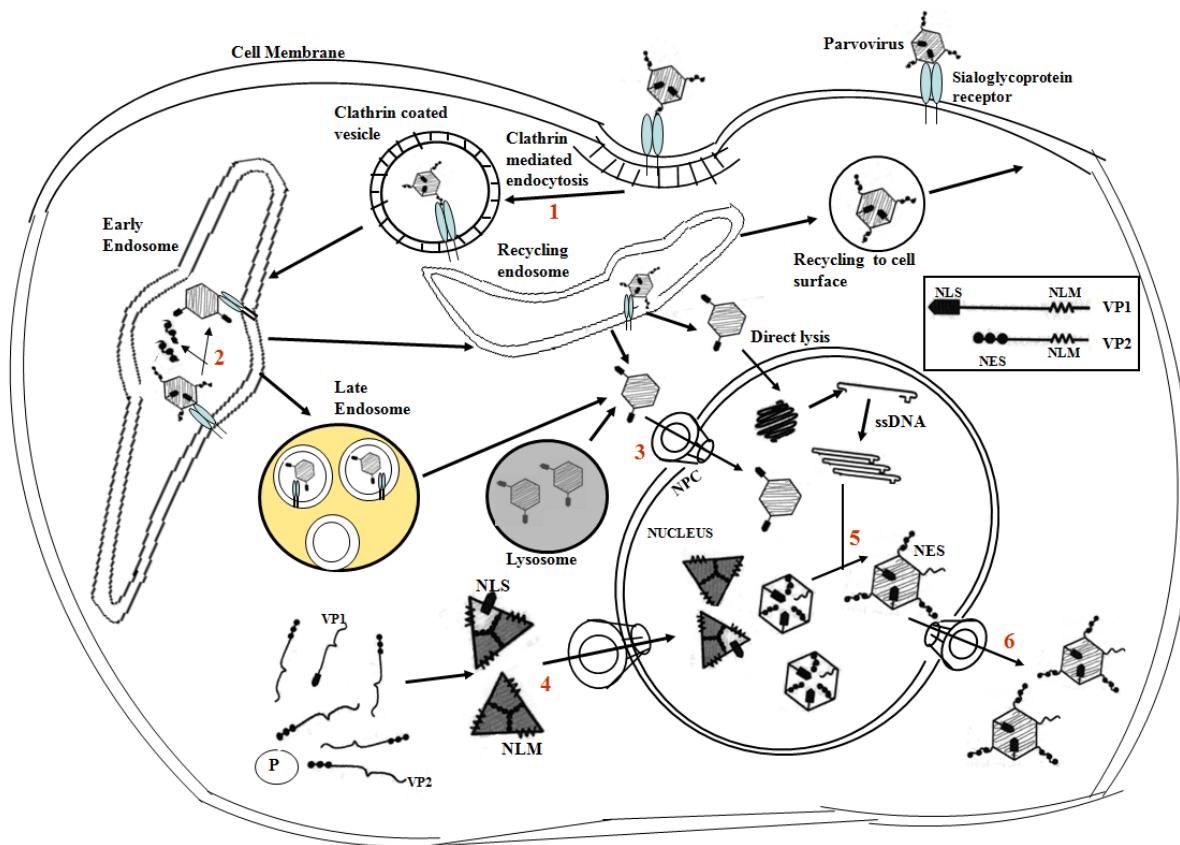


Figure 1-4. A schematic of the life cycle of parvovirus, MVM. (1) Virus binding to receptor followed by receptor mediated endocytosis, (2) NES cleavage and exposure of the NLS of VP1 during trafficking, (3) Nuclear entry, (4) Nuclear translocation of assembly intermediates, (5) Viral DNA encapsidated into preformed empty capsids, (6) Nuclear exit of newly formed virions. Figure inset shows the nuclear transport sequences identified in the VP proteins of MVM. NLS: Nuclear Localization Signal; NLM: Nuclear Localization Motif; NES: Nuclear Export Signal; NPC: Nuclear Pore Complex.

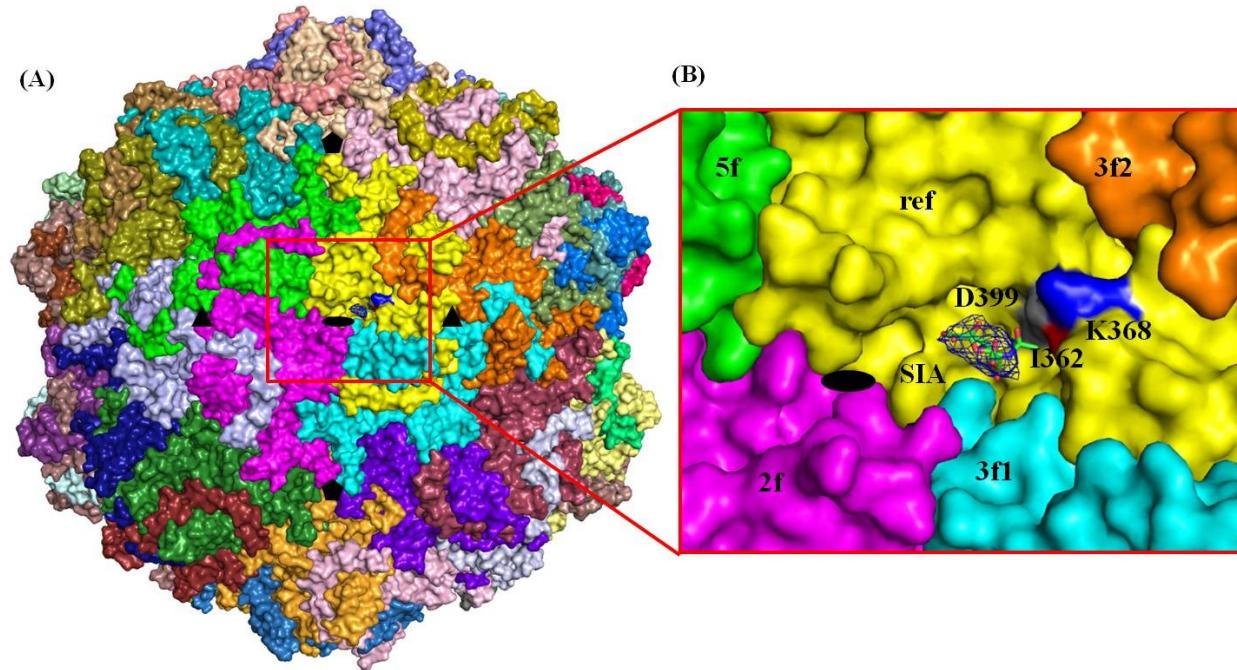


Figure 1-5. Sialic acid binding site in MVMP. (A) Surface representation of an assembled MVMP capsid showing the SIA binding site (region enclosed in red box) in context of the whole capsid with the 60 VP2 monomers in different colors. (B) Surface representation of the close-up of the enclosed region in (A) showing the depression at the icosahedral twofold axis of the MVMP capsid. The reference VP2 monomer (ref, in yellow), and icosahedrally related twofold (2f, in magenta), threefold (3f1 and 3f2, in cyan and orange), and fivefold (5f, in green) monomers are shown. The surface positions of residues I362, K368 and D399 are highlighted in red, blue and grey, respectively. The SIA model (colored according to atom type; carbon, nitrogen and oxygen in green, blue and red, respectively) is shown inside a  $2F_0 - F_c$  map (dark grey mesh) contoured at  $1.8\sigma$ . The approximate location of the icosahedral twofold axes is shown as a filled black oval. This figure was generated using the Pymol program (82)

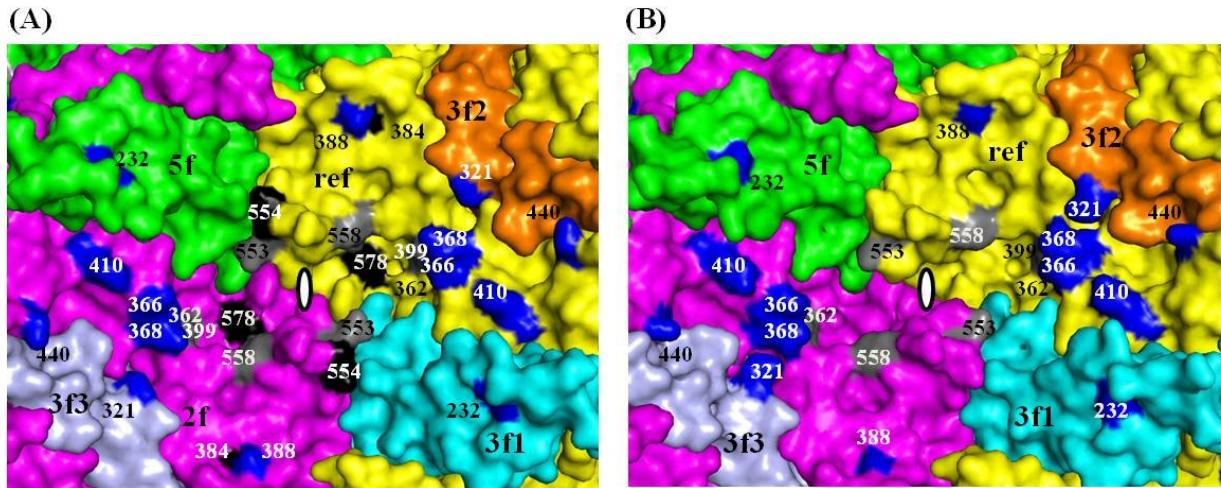


Figure 1-6. Structural clustering of MVMP/i amino acid differences. (A) Close-up view of the MVMP icosahedral twofold axes, with the positions of surface MVMP/i amino acid differences colored and labeled: blue for p/i differences, grey for forward fibrotropic mutations and black for host range switch mutations. (B) Close-up view of the MVMi icosahedral twofold axes, with the residues colored and labeled as in panel (A), except that the host range mutations are not shown. A viral asymmetric unit is shown in the panels (A) and (B). White ovals represent approximate icosahedral twofold axes. This figure was generated using the PyMol program (82).

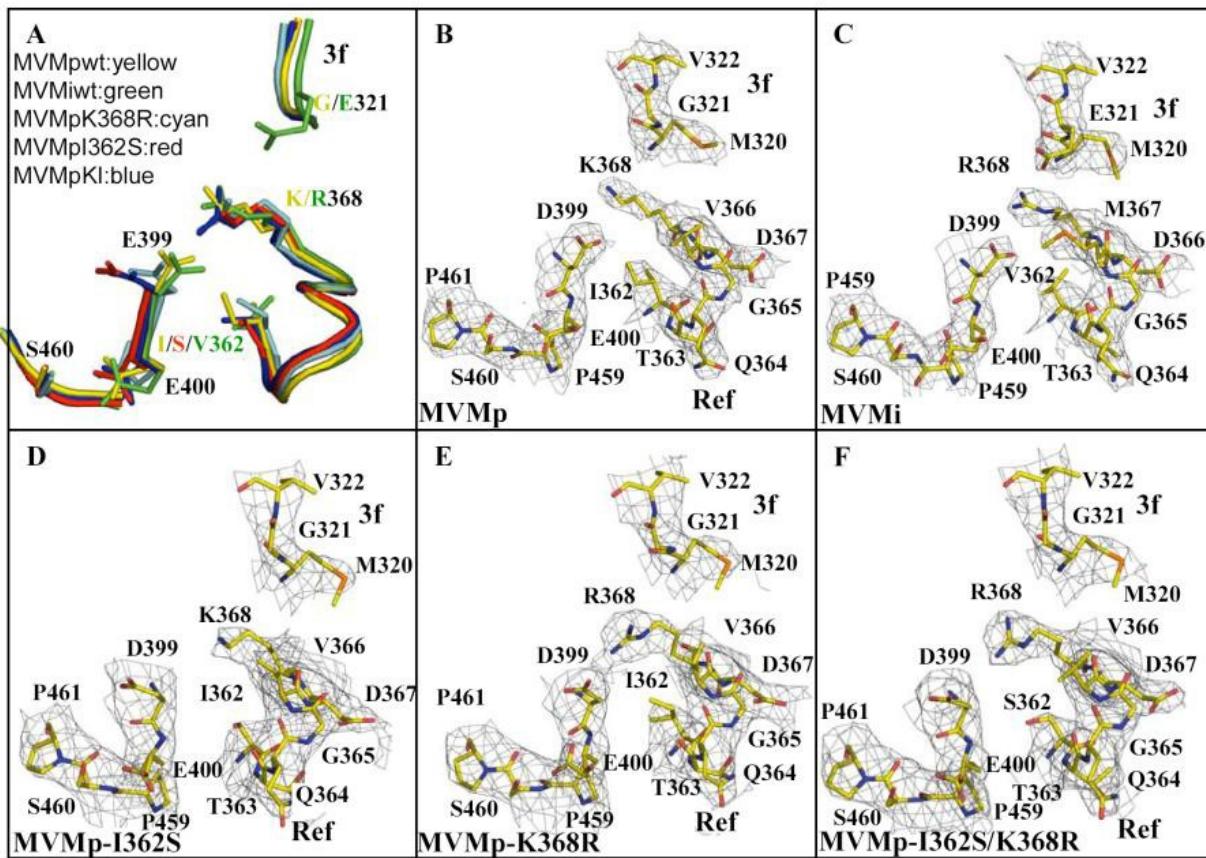


Figure 1-7. MVM capsid structures at  $\sim 3.5 \text{ \AA}$  resolution. (A) Ribbon diagram of VP2 of MVMp, MVMi, MVMp-K368R, MVMp-I362S, and MVMp-K/I showing the side chain variations at the icosahedral twofold axis. (B-F) The side chains of the differing amino acid residues are shown as sticks and colored according to atom type. The  $2F_o - F_c$  map (gray mesh) is contoured at  $2.0\sigma$ . This figure was generated using the PyMol program (82).

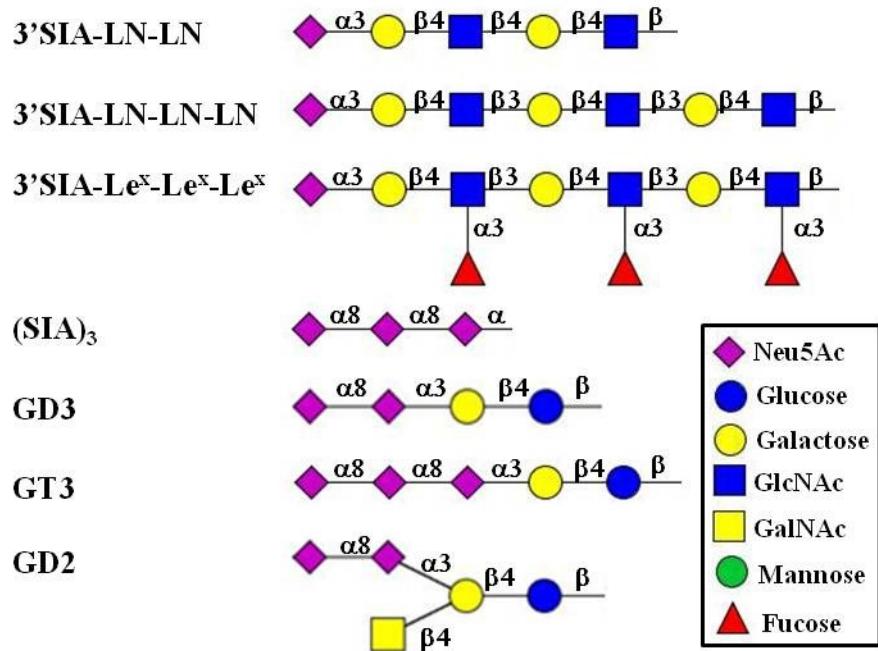


Figure 1-8. Glycans recognized by the MVM viruses. A schematic representation of the glycans recognized by the MVM viruses is shown. Figure inset depicts the glycan symbols.

CHAPTER 2  
ANALYSIS OF MINUTE VIRUS OF MICE (MVM) RECEPTOR COMPLEXES  
TOWARDS UNDERSTANDING THE MECHANISMS OF TISSUE TROPISM

**Background**

MVM serves as an ideal model for studying the capsid determinants of tissue tropism and pathogenicity dictated by receptor interactions due to several reasons (i) the small ssDNA genome size and the simple T=1 icosahedral capsid, formed from the assembly of three VPs, enables genetic manipulation and structural studies, (ii) it is known that MVM-SIA recognition is essential for infection and the depression at the twofold axis has been mapped as the SIA binding site on the MVMP capsid, (iii) two highly homologous strains of MVM, MVMP and MVMi have pronounced differences in tissue tropism and *in vivo* pathogenicity, (iv) the pronounced differences are associated with one or two VP amino acid differences localized at the SIA binding site and result in local structural variations that alter MVM-SIA receptor interactions and utilizations, (v) virulent MVMP variants with one or two VP changes associated with altered receptor specificity and affinity which confers a pathogenic phenotype have also been observed. In addition, capsid adaptations at the SIA binding site resulted in a host range variant, which also makes MVM a useful model for the study of emerging pathogens.

The crystal structure of wt empty MVMP and empty MVMi capsids, MVMP VLPs, MVMi and MVMP virions as well as for VLPs of virulent MVMP mutants, MVMP-I362S, MVMP-K368R, and MVMP-I362S/K368R, have been determined to ~3.5 Å resolution by X-ray crystallography ((7, 168), and unpublished data). The structures of the mutant viruses are similar to each other and to MVMP and MVMi, except for local main- and side-chain differences resulting in the loss or gain of intra-subunit interactions at or close to the mutated site (Figure 1-7). A chain of weak intra and intersubunit amino acid

interactions involving the MVMP*i* differing residues from the wall (E321) toward the floor (R368, D399) of the dimple are observed in MVM*i*, which are not possible in MVMP. A hydrogen bonding interaction between residues R368 and D399 was observed for MVMP-K368R and MVM*i*, but is not possible in MVMP. Significantly, mutation of D399 to G or A in MVM*i*, which abolishes this hydrogen bonding interaction, confers fibrotropism to MVM*i* (7, 79). It has also been proposed that MVM's tissue tropism and pathogenicity is likely controlled by an alteration of the surface charge in the vicinity of the twofold depression (168). These structural differences between MVM viruses colocalize with the SIA binding site of MVM as well as the tropism and pathogenicity determinants for other autonomous parvoviruses, such as CPV, FPV, PPV and ADV. Prior biochemical studies on CPV, FPV and MVM tissue tropism and pathogenicity differences suggested that a block in initial cell receptor attachment is not a restriction to infection, but interactions post entry with intracellular factors was a determinant. However, the colocalization of tropism/pathogenicity determinants that result in local structural variability with receptor recognition sites, close to the twofold depression and the shoulder of the threefold axis on the capsids, suggests a role for receptor interaction in these phenotypes.

The structures for the native capsids provides the 3D platform for comparison with resulting structures of capsid–glycan complexes for identification of the glycan binding site(s) as well as any capsid conformational changes resulting from the glycan interaction. Previous cell binding, glycan array screening and SPR studies provided information on the specificity and affinity of the sialylated glycans but did not define the role of the capsid surface amino acids in these interactions. Here, structural studies of

the MVM viruses complexed with the glycans identified in the glycan array screening were conducted to characterize the nature of the altered MVM-SIA receptor and other carbohydrate interactions that confer an infectious phenotype and the recognition of transformed tumor and neuronal cells. Cellular glycan profiling analysis on the three cell types differentially infected by the MVM strains: EL4 T lymphocytes, A9 fibroblasts and NB324K transformed fibroblast cell line, was conducted to correlate the MVM interacting glycans identified in the glycan array screening with their expression on the different cell types, and to investigate the possibility that differences in the glycan composition of these cells results in the differential tropism between MVMP and MVMI. Finally, to understand the potential for recognition of different modifications of sialic acid by MVM, three types of capsids, VLPs, wt empty particles (Empties) and DNA packaged virions (Fulls), were screened on a newly developed sialylated glycan microarray (SGM). The three types of capsids were analyzed on the SGM to investigate if they exhibit differences in SIA recognition and affinity and also to justify the use of VLPs in lieu of infectious virions to study MVM capsid-glycan receptor interactions.

## **Experimental Methods**

### **Cell Lines**

A9 cells are the ouabain-resistant derivative of the HGPRT defective cell line A9 and are permissive host cells for MVMP. EL4 T is an adherent variant of the T-cell lymphoma line EL4 and is permissive for MVMI. NB324K is a clone of SV40 transformed human newborn kidney fibroblast cells and is permissive for both MVMP and MVMI. For the glycomic profiling, all the three mammalian cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 5% heat-inactivated fetal calf serum and they were grown to a density of  $1 \times 10^7$  cells. Sf9 insect cells were

grown in suspension culture in Erlenmeyer flasks in Sf-900 II SFM media (Gibco/Invitrogen Corporation) supplemented with 1% Antibiotic-antimycotic (ABAM) at 27°C.

### **Generation of Full and Empty MVM capsids**

A9 ouabr11 cells were grown in spinner culture in DMEM containing 5% fetal bovine serum and antibiotics to a density of ~6x10<sup>5</sup> cells/ml. The cultures were infected with predetermined titers of transfection-derived parvovirus MVM (MVMP) stocks (GenBank accession number J02275) and expanded until cell counts indicated a progressive rise in numbers of dead cells. The cells were harvested by centrifugation, washed in phosphate-buffered saline without Ca<sup>2+</sup> or Mg<sup>2+</sup> (PBS; Invitrogen, Carlsbad, CA), pelleted, and resuspended in 10 ml of TE8.7 (50 mM Tris-HCl pH 8.7, 0.5 mM EDTA) per liter of infected cells. Following three cycles of freeze-thaw at 37°C the pellets underwent repeated centrifugation at 2,000 rpm (800g) to clarify extracts, and were stored at -20°C.

For purification, 6 ml aliquots were further clarified by centrifugation at 11,000 rpm in a Sorvall SS34 rotor at 4°C and then floated on top of a 6 ml iodixanol (OptiPrep; Axis-Shield, Oslo, Norway) step gradient (1 ml 55% and 2 ml 45% in TE8.7, followed by 2 ml 35% and 1 ml 15% in PBS plus 1 mM MgCl<sub>2</sub> and 2.5 mM KCl). Samples were centrifuged at 35,000 rpm for 18 h at 18°C in a Beckman SW41 rotor. Fractions were collected from the bottom of the gradient and full and empty particle concentrations were assessed by hemagglutination and also sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by staining with Coomassie Blue (Sigma). The capsid integrity was checked by negative-stain electron microscopy (EM) (Figure 2-1). For the EM, 5 µl of purified virus at an estimated concentration of 2.0 mg/ml was

spotted onto a 400 mesh carbon-coated copper grid (Ted Pella, Inc., Redding, CA, USA) for 1 min before blotting with filter paper (Whatman No.5). The sample was then negatively stained with NanoW for 1 min twice, blotted dry and viewed on a Hitachi 3000 electron microscope. MVMi was recovered in the same way as MVMP except that the transfection was done in NB324K cells. For SGM screening, all viruses were finally buffer-exchanged into 1XPBS.

### **Recombinant Virus Production and Purification**

Recombinant baculovirus constructs expressing the VP2 of wt MVMi, wt MVMP and the virulent MVMP mutants (MVMP-I362S, MVMP-K368R, and MVMP-I362S/K368R) that self-assemble into VLPs were constructed as described (188). Sf9 insect cells grown in suspension culture at 27°C were infected with a titered baculovirus construct at a multiplicity of infection (m.o.i) of 5.0 plaque-forming units per cell. Following incubation at 27°C for 72 h, the cells were spun down for 20 min at 1,500 rpm and then the pellet resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.2% Triton X-100, 10 mM MgCl<sub>2</sub>). The MVMP and MVMi VLPs were purified based on published procedures (132) with some modifications. The virus capsids were released from the cells by three cycles of rapid freeze-thaw with the addition of Benzonase (Merck KGaA, Germany) after the second cycle. The cellular debris was removed by low speed centrifugation in a JA-20 rotor (10,000 rpm, 15 min, 4°C). The supernatant was diluted with TNET buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.2% Triton X-100) and pelleted through a 20% (w/v) sucrose cushion by ultracentrifugation in 70Ti rotor at 45,000 rpm for 3 h at 4°C. The resulting pellet was resuspended overnight at 4°C in TNEM buffer (25 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.2% Triton X-100, and 2 mM MgCl<sub>2</sub>). The resuspended sample was then

subjected to a low-speed 2,000 rpm spin to remove particulate material and further purified by ultracentrifugation on a sucrose-step gradient (5-40% w/v in TNETM) at 35,000 rpm for 3 h at 4°C in a Beckman SW41Ti rotor. A visible blue fraction containing VLPs, sedimenting at ~20-25% sucrose, was extracted and dialyzed against TE buffer (25 mM Tris-HCl pH 7.5, 1 mM EDTA). Next, CsCl was added to the dialyzed sample to a final density of 1.40 g/cm<sup>3</sup>, and then subjected to equilibrium centrifugation in a Beckman SW41Ti rotor at 35,000 rpm for 24 h at 4°C. Visible virus fractions were extracted and extensively dialyzed into 1XPBS for SGM screening studies or into 10 mM Tris-HCl pH 7.5, 150 mM NaCl for crystallization screens. The concentration of the viruses were estimated from optical density measurements (assuming an extinction coefficient of 1.0 for MVM VLPs) and adjusted to 5-10 mg/ml using Ultrafree 100 kDa cut-off centrifugal filter units (Millipore, Billerica, MA). The purity and integrity of the virus capsids were monitored using SDS-PAGE and negative-stain EM, respectively (Figure 2-1).

### **Crystallization and Data Collection**

Crystallization drops were setup using the hanging drop vapor diffusion method (199) with VDX 24-well plates and siliconized cover slips (Hampton Research, Laguna Niguel, CA, USA). Crystals were grown under the same conditions previously used for MVMP VLPs (132) with the virus at a concentration of 5-10 mg/ml in 10 mM Tris-HCl pH 7.5, 150 mM NaCl buffer with polyethylene glycol (PEG) 8000 (1% w/v), 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O and 150 mM NaCl as precipitants. The glycans, 3'SIA(Le<sup>X</sup>)<sub>3</sub>, 3'SIA(LN)<sub>3</sub> and GT3 were provided by Core D of the Consortium for Functional Glycomics (CFG) (<http://www.functionalglycomics.org/>) and resuspended in 10 mM Tris-HCl pH 7.5 to give a concentration of 10 mg/ml. The co-crystallization drops (5 µl) contained 2 µl of

virus solution (10 mg/ml), 1  $\mu$ l of glycan (at a concentration to give capsid: glycan ratios of 1:180 or 1:600) and 2  $\mu$ l of reservoir solution equilibrated against 1 ml of reservoir solution at RT. Crystallization drops containing virus alone were also setup in preparation for glycan soaking experiments if the co-crystallization trials did not yield crystals amenable to diffraction. These drops (4  $\mu$ l) contained 2  $\mu$ l of virus solution (10 mg/ml) and 2  $\mu$ l of reservoir solution equilibrated against 1 ml of reservoir solution at RT. Certain virus-glycan co-crystallization trials yielded no useful crystals, and in these cases, VLP crystals were soaked with the respective glycan for 1-2 h prior to data collection. The virus-glycan complexes for which diffraction data was collected are summarized in Table 2-1

X-ray diffraction data was collected from crystals incubated in cryoprotectant solution containing 10 mM Tris-HCl pH 7.5, 150 mM NaCl and 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O with 10% PEG 8000 and 30% glycerol for 30 s and flash-cooled in liquid nitrogen vapor. Diffraction data for the various capsid-receptor complex crystals were collected at three beamline facilities, APS (Advanced Photon Source), BNL (Brookhaven National Lab) and CHESS (Cornell High Energy Synchrotron Source). Data was collected with a crystal-to-detector distance of 300 mm, an oscillation angle of 0.3° per image, and exposure times of 30-45 s per image. The reflections were indexed and integrated with the HKL2000 suite of programs (224), and scaled and merged with SCALEPACK (224) . The space group for the crystals was determined to be C2. The data collection and processing statistics are summarized in Table 2-1.

### **Structure Solution**

The diffraction intensity data sets were converted to structure factor amplitudes using the TRUNCATE program from CCP4 (Collaborative Computational Project,

Number 4) (62). The Matthew's coefficient ( $V_M$ ) for all the complex data sets was calculated to be  $\sim 3.5 \text{ \AA}^3 \text{ Da}^{-1}$ , corresponding to solvent content of 65% (195). The C2 unit cell contains two half particles with different orientations in the crystallographic asymmetric unit, as described for the previously solved MVMI and MVMP-VLP structures (7, 168, 184). The MVMP VP2 VLP structure coordinates (PDB accession no. 1Z14) (168) was used as the phasing model to initiate molecular replacement using the CNS program (46, 258). Iterative cycles of model refinement were performed using the simulated annealing, energy minimization and individual temperature factor (B-factor) refinement options in the CNS program while applying strict 60-fold non-crystallographic symmetry (NCS) operators (46). Five percent of the total data set was partitioned for monitoring of the refinement process (45). The refinement cycles were alternated with manual model building using the COOT program (97) into the sigma weighted averaged Fourier electron density maps ( $2F_o - F_c$  and  $F_o - F_c$  maps, where  $F_o$  represents the structure factors for the capsid-glycan complex data and  $F_c$  represents the calculated structure factors from the MVMP VLP model) that were generated in the CNS program using a molecular mask while applying strict 60-fold NCS operators (46). The coordinate files for the glycan molecules were obtained from the HIC-Up server (166) and the geometry restraints and dictionary files were generated using the subroutine phenix.elbow from PHENIX (2). The glycans models were docked into the  $F_o - F_c$  density using interactive rigid-body rotations and translations in COOT (97). The topology and dictionary files generated for the glycans in PHENIX (2) were then used for subsequent refinement in CNS (46). The refinement process was deemed to have converged when there was no further improvement in the agreement between the observed  $F_o$  and

calculated  $F_c$  structure factors ( $R_{\text{factor}}$ , where  $R_{\text{factor}} = (\sum|F_o|-|F_c|)/\sum|F_o|) \times 100$ ). The quality of the refined structures was analyzed using COOT (97) and MOLPROBITY (56). The refinement statistics are given in Table 2-1. The figures were generated using the program PYMOL (82).

### **Sialylated Glycan Microarray (SGM) Preparation and Virus Screening**

Terminally sialylated glycans with various sialic acid modifications were synthesized as previously described (277, 278). Briefly, a synthetic strategy that combines the bifunctional fluorescent tag 2-amino-(N-aminoethyl) benzamide (AEAB) and the one-pot-three-enzyme sialylation reaction was utilized for generating the sialylated glycans. The GAEAB (Glycan-AEAB conjugates) with terminal galactose residues were used as precursors and included AEAB conjugates of lactose, lacto-N-neo-tetraose (LNnT), lacto-N-tetraose (LNT) and asialo-galactosylated biantennary oligosaccharide (NA2), which were prepared from natural glycans and purified prior to the multi-enzyme sialylation step. Aliquots (50 µg to 1 mg) of the GAEABs were subjected to enzymatic sialylation in a combinatorial fashion. The addition of different modified ManNAc, ManNGc, or Man precursors in the reactions allowed formation of corresponding terminal Neu5Ac, Neu5Gc, or Kdn derivatives, respectively (these SIA derivatives were introduced in chapter 1). The products were purified by HPLC and the structure of each sialylated glycan was confirmed by MALDI-TOF and HPLC analysis. 77 sialylated structures were generated incorporating 16 different terminal sialic acids on 4 different underlying structures that were tagged with a fluorescent linker. Each glycan was quantified based on its fluorescence and printed onto N-hydroxysuccinimide (NHS)-activated glass slides in replicates of n=4 to generate the SGM. The array also

included three controls (LNnT, NA2, and Man5), corresponding to chart ID numbers 78-80.

The MVM samples (buffer-exchanged into 1XPBS) were screened on the SGM in collaboration with Core H of CFG. The printed slides were washed and blocked with 50 mM ethanolamine in 0.1 M Tris-HCl buffer (pH 9.0) for 1 h, and prior to screening the slides were rehydrated for 5 min in TSM buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM CaCl<sub>2</sub> and 2 mM MgCl<sub>2</sub>). The virus samples were diluted with Binding Buffer (TSM buffer plus 1%BSA and 0.05% Tween 20) to give a final volume of 70-100 µl (0.05 mg/ml) which was added to the slide and incubated at room temperature for 1 h. The slide was then washed with Wash Buffer (TSM buffer plus 0.05% Tween 20), and an anti-MVM capsid antibody (Tatt-2; polyclonal from rabbit) was added at a dilution of 1:5,000. The slide was then washed again with Wash buffer and incubated with Cy5-labeled-goat anti-rabbit IgG at 5 µg/ml. The slide was scanned with a Perkin Elmer ProScanarray microarray scanner and for Cy5 fluorescence, the wavelengths 649 nm (Excitation) and 670 nm (Emission) were used. The scanned images were analyzed with the ScanArray Express software to determine the average relative fluorescence units (RFU) and standard deviation (S.D) of the four replicates. To analyze the results, all glycans were ranked according to their signal-to-noise (S/N) ratio by dividing their mean RFU from four replicates by the mean background generated in the control wells lacking sialylated glycans. Variation within the 4 replicates was assessed as the coefficient of variation (%CV), which was calculated as 100 x S.D/Mean. Any value with a %CV of >30, was considered unacceptable as reported in other studies (278).

LNnT=Galβ1-4GlcNAcβ1-3Galβ1-4Glc; LNT=Galβ1-3GlcNAcβ1-3Galβ1-4Glc;

NA2=Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6(Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc; Man5=Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\beta$ 1-4GlcNAc $\beta$ 1-4 GlcNAc.

### Glycan Profiling of Permissive Cell Lines

A9 fibroblasts (infected by MVMP but not MVMi), EL4 lymphocytes (infected by MVMi but not MVMP) and NB324K transformed cells (infected by both MVMP and MVMi) were grown in suspension to a density of  $1\times10^7$  cells/ml. The cells were harvested by centrifugation at 250g for 5 min, washed with 1XPBS three times, and then pelleted. The cell pellets were resuspended in 1 ml of 1XPBS, transferred to a microfuge tube, and pelleted at 800g. The supernatant was aspirated, and the cell pellets were stored at -80°C. Further processing of the cell lines to derivatize, extract, purify, and analyze the glycans by mass spectrometry was carried out in collaboration with Core C of CFG as described in North *et al.* (216).

## Results

### Virus Purification and Crystallization

The purity and integrity of the MVM capsids used in this study (Full-MVMP, Empty-MVMP, MVMP-VLP, Full-MVMi, Empty-MVMi, MVMi-VLP, MVMP-I362S VLP, MVMP-K368R VLP and MVMP-I362S/K368R VLP) were verified by SDS-PAGE (Figure 2-1(1)) and negative-stain EM (Figure 2-1(2)), prior to crystallization and SGM screening. The SDS-PAGE analysis showed that the full MVMP and MVMi capsids contained VP1 (83 kDa), VP2 (64 kDa) and VP3 (61 kDa); empty MVMP and MVMi capsids contained VP1 and VP2; and all the VLPs contained only VP2. The negative stain EM analysis showed the full and empty capsids and VLPs to be intact. Co-crystals for the MVMP-3'SIA( $\text{Le}^\alpha$ )<sub>3</sub> complex and MVMP-K/I-3'SIA(LN)<sub>3</sub> complex and VLPs were obtained in approximately 3-4 weeks and they grew to dimensions of 0.3 x 0.2 x 0.1 mm.

## **Structural Analysis of MVM Capsid-Receptor Interactions**

Diffraction data sets were obtained for co-crystals of MVMP-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex and MVMP-K/I-3'SIA(LN)<sub>3</sub> complex. For the MVMP-3'SIA(LN)<sub>3</sub>, MVMI-3'SIA(LN)<sub>3</sub>, MVMI-3'SIA(Le<sup>X</sup>)<sub>3</sub>, and MVMI-GT3 complexes, the VLP crystals were soaked with the respective glycans and the diffraction data was collected (Table 2-1). The co-crystals of the MVM viruses with glycans, or MVM crystals soaked with glycans diffracted X-rays to ~3.3 to 3.9 Å resolution (Table 2-1). The crystals were isomorphous to those of wt MVMI and MVMP-VLP for which structures are available (7, 168). The completeness and quality of the complex data sets was consistent with those used to solve the crystal structures of MVMP and MVMI. Residues 39 to 587 of the VP2 were built into the MVM electron density maps. The averaged density maps were not interpretable beyond N-terminal residue 39 of VP2, as was previously reported for the MVMP-VLP structure (168). A comparison of all the MVM-receptor complex structures showed that all the glycan receptors, whether linked α2-3 or α2-8, utilized the same receptor binding pocket at the icosahedral twofold axis identified previously (188) on both the strains (Figure 2-2 and Figure 2-3). In all the complex structures, the glycans were modeled into the F<sub>o</sub>-F<sub>c</sub> difference density map at 3σ. Although longer sialylated glycans were utilized for these structural studies, all the carbohydrate components of a particular complexed glycan were not observed in the crystal structure. For the capsid-glycan complexes, there was variability in the size of the oligosaccharides that were observed even for the complexes where the similar glycan was used (Figure 2-2C and Figure 2-3C). The MVMP-3'SIA(Le<sup>X</sup>)<sub>3</sub> and MVMI-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex data allowed for the visualization of a longer ordered oligosaccharide (Neu5Aca2-3Galβ1-4(Fuca1-3)-GlcNAcβ: tetrasaccharide) in the pocket at the icosahedral twofold axis as compared to when only sialic acid was

soaked into the MVM capsids or the other complex data sets that were analyzed in this study (188) (Figure 2-2A and Figure 2-3A). For the MVMP-3'SIA(LN)<sub>3</sub>, MVMI-3'SIA(LN)<sub>3</sub> and MVMP-K/I-3'SIA(LN)<sub>3</sub> complexes, only 3'SIA-LN, terminal SIA, and 3'SIA-Gal were ordered, respectively (Figure 2-2C, Figure 2-3C, and Figure 2-2E). For the MVMI-GT3 complex, only densities for the terminal trisaccharides, (SIA)<sub>3</sub> was observed and densities for the other two sugars, Gal $\beta$ 1-4Glc, were not ordered (Figure 2-3E). In these complexes, the ring structure of the ordered carbohydrate molecules were placed with confidence in the difference density maps at 3 $\sigma$ , but the side groups were not covered at this sigma level. At lower sigma contour levels of 2 $\sigma$  to 1.5 $\sigma$ , the map extends over some of the side groups but their exact conformation could not be assigned. In the MVMP-3'SIA(Le $\text{x}$ )<sub>3</sub>, MVMI-3'SIA(Le $\text{x}$ )<sub>3</sub> and MVMP-3'SIA(LN)<sub>3</sub> complexes, where longer oligosaccharides were ordered, the conformation of the oligosaccharide reduced the degrees of freedom and allowed the modeling of some side groups. The side groups, such as N-acetyl, glycerol or carboxyl, that could not be assigned were modeled in the best possible conformation, i.e., to avoid clashes with other atoms and maximize hydrogen bonding and ionic interactions. Comparative analysis of the MVMP-VLP and MVMI virion structures (168) with the MVM capsid-glycan complex structures did not reveal any detectable conformational changes upon receptor binding.

In the MVMP-3'SIA(Le $\text{x}$ )<sub>3</sub> complex structure, the glycan extends from the pocket at the twofold depression towards the wall of the threefold protrusion that faces the pocket (Figure 2-2A and B). The glycan interacts (contact distance of 2.4 - 4.5 Å) with the residues M243, N323, K368, R375, E394, Y396, T397, W398, D399, T401, Y558, and V575. The glycan contact region contains residues that differ between MVMP and

MVMi, as well as the residues associated with *in vitro* tropism and pathogenicity, e.g., 362, 368, 399, 558, 575 etc. In the MVMP-3'SIA(LN)<sub>3</sub> complex, the glycan extends from the pocket to the wall of the threefold protrusion in a manner similar to that observed for the 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan in the MVMP-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex structure, but in this case the glycan follows the contour of the pocket wall and only 3'SIA-LN is observed in the structure (Figure 2-2C and D, and Figure 2-4A). For the MVMi-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex, the glycan lies further away from the twofold pocket or the wall of the pocket and interacts with the wall of the threefold protrusion that faces the pocket (Figure 2-3A and B). This glycan conformation was different from that observed in the MVMP-glycan complexes, described above (Figure 2-4C). Residues N323, R368, R375, Y378, E394, T397, Y558, K563, and V575 interact with the 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan on the MVMi capsid. The MVMi-3'SIA(LN)<sub>3</sub> complex data showed the terminal SIA binding deeper into the twofold pocket compared to the MVMP-3'SIA(Le<sup>X</sup>)<sub>3</sub>, MVMP-3'SIA(LN)<sub>3</sub>, and MVMi-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex structures, described above (Figure 2-3C and D). The residues involved in this interaction are E552, Y558 and I578. The GT3 binds along the wall of the twofold pocket interacting with residues E321, N323, M366, R368, R375, E394, Y396, T397, D399, Y558 and V575 on the MVMi capsid (Figure 2-3E and F). The binding pattern of 3'SIA(Le<sup>X</sup>)<sub>3</sub>, 3'SIA(LN)<sub>3</sub> and GT3 glycans on the MVMi capsid is different, wherein 3'SIA(Le<sup>X</sup>)<sub>3</sub> binds closer to the wall of the threefold protrusion that faces the pocket, the 3'SIA(LN)<sub>3</sub> glycan binds deep into the twofold pocket, and GT3 binds along the wall of the twofold pocket (Figure 2-4B). On the MVMP-K/I capsid, the 3'SIA(LN)<sub>3</sub> glycan binds further down in the twofold pocket as compared to the MVMP or MVMi glycan

complexes and interacts with the residues D218, K241, R349, R368 and D553 (Figure 2-2E and F).

### **Binding of MVM Viruses to SGM**

Three types of MVMP and MVMi capsids; VLPs, empty particles and virions, were screened on the SGM. For the particular strain, MVMP or MVMi, the binding profile of the VLPs was similar to that of the full and empty particles. The SGM had both  $\alpha$ 2-3 and  $\alpha$ 2-6 linked sialylated glycans and three asialoglycans as controls, but MVMP and MVMi specifically bound only to the  $\alpha$ 2-3 linked sialylated derivatives, which is consistent with the previous glycan array screening (Figure 2-5 and Figure 2-6) (209). MVMP and MVMi showed a preference for binding to SIA that was methylated at C-9 and not C-8. Out of a total of 4 glycans that were derivatized with Neu5Ac9-O-methyl SIA, MVMP and MVMi bound to only the  $\alpha$ 2-3 linked SGM 23 (Neu5Ac9Me $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4Glcitol) and SGM 55 (Neu5Ac9Me $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAcitol) glycans. The other two glycans, SGM 8 (Neu5Ac9Me $\alpha$ 2-6Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4Glcitol) and SGM 40 (Neu5Ac9Me $\alpha$ 2-6Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAcitol) are similar to SGM 23 and SGM 55, respectively, but are  $\alpha$ 2-6 linked and were not recognized by the MVM viruses. All the MVMP viruses also showed slight binding to glycans with the 9-O-acetylated (SGM 52: Neu5,9Ac $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAcitol) and the 9-O-lactoylated (SGM 60: Neu5Ac9L $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Neu5Ac9L $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAcitol) SIA derivative (Figure 2-6A, B and C). MVMi viruses showed the same profile as the MVMP viruses, except that they did not bind to 9-O-acetylated (SGM 52) and the 9-O-

lactoylated (SGM 60) SIA derivative (Figure 2-6D, E and F). The binding profile for the genome containing viruses, especially MVMp virions showed some non-specific binding with high %CV to other sialylated glycans, as compared to the empty capsids (Figure 2-6A). In addition, all the MVM viruses bound to the biantennary glycan, SGM 48 (Neu5Aca2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Neu5Aca2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAcitol) on the SGM array. The affinity ranking of the derivitized glycans recognized by MVMp was in the order 9-O-methylated-monosialylated (SGM 23) > 9-O-methylated-biantennary (SGM 55) > 9-O-acetylated-biantennary (SGM 52) ≈ 9-O-lactoylated-biantennary (SGM 60) ≈ SIA-biantennary (SGM 48) (Figure 2-6). The affinity ranking for the MVMi viruses was 9-O-methylated-monosialylated (SGM 23) > 9-O-methylated-biantennary (SGM 55) > SIA-biantennary (SGM 48) (Figure 2-6). The glycans that were recognized by the MVM viruses had different SIA modifications but they all had Type2 LacNAc (Gal $\beta$ 1-4GlcNAc) as the core structure (consistent with previous glycan array screening results) and were derived from either LNnT or NA2 precursor glycans. Also, consistent with the previous glycan array screening studies, none of the MVM viruses bound to the Neu5Gc or Kdn sialylated derivatives.

### Glycan Profiling of Cell Lines

Three cell lines, A9 fibroblasts (infected by MVMp), EL4 T lymphocytes (infected by MVMi) and NB324K transformed cells (infected by both MVMp and MVMi) were subjected to glycomic profiling by MALDI-MS analysis. The profiling provided information on the carbohydrate composition of the glycans present on these cells but not the linkages ( $\alpha$ 2-3 or  $\alpha$ 2-6) present in the glycans. However, SIA-SIA (multisialylated) can only be linked  $\alpha$ 2-8 or  $\alpha$ 2-9 to each other. The mass spectra

obtained from the peptide N-glycosidase F-released and -permethylated N-glycans from the three cell lines are shown in Figure 2-7. For the N-glycan (glycan that is covalently attached to protein at asparagine (Asn) residues by an N-glycosidic bond and consists of a common core sequence ( $\text{Man}\alpha 1\text{-}6(\text{Man}\alpha 1\text{-}3)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-Asn}$ ) fraction for all three cell lines, a full complement of oligomannose (Man residues are attached to the core) as well as complex type N-glycans (multi-branched glycopeptides containing outer chains of SIA, galactose (Gal) or N-acetylglucosamine (GlcNAc)) with bi-, tri-, or tetra-antennary (branches) structures were observed. Hybrid N-glycans (two branches from the core, one that terminates in Man and the other that terminates in a sugar of the complex type) were not observed. The N-glycans on these cell lines terminated in Man, Gal, GlcNAc, or SIA. Also, all the N-linked glycans have GlcNAc at their reducing end (which is attached to protein), which is the most common type of N-glycan linkage that has been reported. The presence of bisecting GlcNAc (the GlcNAc that bisects the  $\text{Man}\alpha 1\text{-}6(\text{Man}\alpha 1\text{-}3)\text{Man}\beta$  triad) in complex N-glycans is also observed for all three cell lines. The majority of the sialylated N-glycans observed are branched. For the low mass N-glycans, there are two (biantennary) to three (triantennary) branches, and as the molecular mass increases, there is an increase in the number of branches instead of an increase in the chain length (Figure 2-7). All the three cell lines also possessed core-fucosylated sialoglycans (i.e., fucose attached to GlcNAc at the reducing end) with the A9 cells showing the least expression of these glycans. Only in the NB324K cells, fucose is linked to GlcNAc in the antennae as part of the 3'SIA- $\text{Le}^x$  motif (Neu5Aca2-3Galβ1-4(Fuca1-3)GlcNAcβ) (e.g., m/z=3142.4, m/z=3952.3, etc.; enclosed in red box in Figure 2-7C and F). The glycan composition for the A9 and EL4

T cells were very similar for the low molecular mass (Figure 2-7A and B), but for the medium mass there were no similar glycans present (Figure 2-7D and E). The NB324K cell line expresses more glycans than A9 and EL4 T, especially in the medium mass fraction (Figure 2-7F). The low mass glycans present on the NB324K cells are similar to that expressed on the A9 and EL4 T cells (Figure 2-7A, B and C), but the medium mass fraction of the NB324K cells shares only one common glycan ( $m/z=3603.5$ ) with the A9 cells (enclosed in green box in Figure 2-7D and F), and one common glycan ( $m/z=3776.5$ ) with the EL4 T cells (enclosed in blue box in Figure 2-7E and F).

The profiles for the polar glycolipids and non-polar glycolipids are shown in Figure 2-8. For the ease of identification, the common names (if known) of the identified glycolipids are included in Figure 2-8. The glycolipids present in the spectra were; GM3 ( $m/z=855.5$ ), GM2 ( $m/z=1100.7$ ), GM1a ( $m/z=1304.9$ , major species), LST1d ( $m/z=1304.9$ , minor species), GT3 ( $m/z=1577.8$ ), GD1a ( $m/z=1665.9$ , major species), GD1c ( $m/z=1665.9$ , minor species), GT1c/GT1a ( $m/z=2027.0$ ) and GQc/b ( $m/z=2387.1$  (265)). A9 cells and NB324K cells have the same glycan profile for the non-polar glycolipid fraction although the relative abundance of the glycans in the NB324K cells is higher (Figure 2-8D and F). In the EL4 T cells, none of the known non-polar glycolipids could be assigned to the spectra observed (Figure 2-8E). A9 cells showed the expression of more types of polar glycolipids than EL4 T or NB324K cells (Figure 2-8A, B and C). All the glycolipids had SIA at the non-reducing end and glucose (Glc) at the reducing end in the polar and non-polar glycolipid fraction for all the cells. Also, there are three lower molecular weight sialylated glycans (GM3, GM2,  $m/z=1130.7$ ; enclosed in red box in Figure 2-8) that are present on both the polar and non-polar glycolipids

fraction. In all the three cell lines, there is an increased expression of the monosialylated glycolipids as compared to the multisialylated glycolipids, which is more evident in the polar glycolipids fractions. All the glycolipids in the polar and non-polar fraction for the three cell lines were identified to be gangliosides ( $\text{GalNAc}\beta\text{-}4\text{Gal}\beta\text{-}4\text{Glc}\beta\text{Ceramide}$  motif) except for LST1d (enclosed in black box in Figure 2-8A) which belongs to the neolactoseries ( $\text{GlcNAc}\beta\text{-}3\text{Gal}\beta\text{-}4\text{Glc}\beta\text{Ceramide}$  motif) (265). The multisialylated glycans (linked  $\alpha$ 2-8 or  $\alpha$ 2-9) are present in the polar glycolipid fraction for all the three cell lines. The ganglioside GT3 which was recognized by MVMi in the previous glycan array screening was present only on A9 cells at low abundance (enclosed in green box in Figure 2-8A) (209). However, the core motif in GT3 (Neu5Aca-Neu5Ac-Gal-Glc) is also present in GT1c that is expressed in EL4 T cells albeit at low abundance (enclosed in blue box in Figure 2-8A and B). Branching of glycans was also observed for the polar glycolipid fraction. Comparing the presence of the two derivatives of SIA, Neu5Ac is the major SIA in these cell lines, but Neu5Gc is present in trace amounts in the N-glycan fraction for A9 and EL4T cells, and in the glycolipid fraction for the A9 and NB324K cells (enclosed in dashed black box in Figure 2-7A, B and D; and Figure 2-8A, C, D and F).

The mass spectra of the permethylated O-glycans (an O-glycan is covalently attached via an N-acetylgalactosamine ( $\text{GalNAc}$ ) moiety to proteins at serine or threonine residues by an O-glycosidic bond) are shown in Figure 2-9. While the A9 and EL4 T cells are shown to express only one type of O-glycan, the NB324K cells express four types of O-glycans, of which one is similar to the one expressed on A9 and EL4 T cells. The glycan which is expressed in all the three cell lines at the highest abundance ( $m/z=534.5$ ; enclosed in red box in Figure 2-9) can be classified as T antigen and

belongs to either Core 1 O-glycan ( $\text{Gal}\beta 1\text{-}3\text{GalNAc}\alpha\text{Ser/Thr}$ ) or Core 8 O-glycan ( $\text{Gal}\alpha 1\text{-}3\text{GalNAc}\alpha\text{Ser/Thr}$ ), depending on whether the Gal is linked  $\alpha$  or  $\beta$  to the GalNAc (41). The only SIA containing O-glycan (enclosed in green box in Figure 2-9C) is expressed on the NB324K cells and belongs to Core 1. The remaining O-glycans expressed on the NB324K cells were identified to belong to Core 2 ( $\text{GlcNAc}\beta 1\text{-}6(\text{Gal}\beta 1\text{-}3)\text{GalNAc}\alpha\text{Ser/Thr}$ ) ( $m/z=779.4$ ,  $m/z=983.6$ ).

The N-glycan (low mass and medium mass), polar, and non-polar glycolipid composition for all the three cell lines showed the presence of terminal SIA glycans that were identified in the glycan array conducted previously and the SGM screening conducted in this study to recognize MVM capsids (209). In addition, glycans containing the SIA-LN motif ( $\text{Neu5Ac-Gal-GlcNAc}$ ) were present in the N-glycan fraction for all the cells and in the polar glycolipid fraction of A9 cells (for example, present in the glycans enclosed in the green or blue box in Figure 2-7D, E and F, or black box in Figure 2-8A), which was also consistent with the glycan array and SGM screening data. The biantennary glycan with the terminal SIA-LN motifs (i.e., SGM 48;  $\text{Neu5Ac}\alpha 2\text{-}3\text{Gal}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}2\text{Man}\alpha 1\text{-}3(\text{Neu5Ac}\alpha 2\text{-}3\text{Gal}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}2\text{Man}\alpha 1\text{-}6)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAcitol}$ ) was present in the low mass fraction of A9 cells and NB324K cells, but not in the EL4T cells ( $m/z=2792.2$ , enclosed in dashed red box in Figure 2-7A and C). Based on the linkage combinations possible for the glycan with  $m/z=4850.2$  (enclosed in a dashed green box in Figure 2-7F) in the medium mass N-glycan fraction of the NB324K cells, it possibly contains the 3'SIA-LN-LN motif which was recognized by the MVM viruses in the glycan array screening (209).

## Discussion

### Capsid-Receptor Interactions Dictate Infectious Outcome

The MVM capsid has been shown to play a crucial role in tropism and in the onset of infection, though the essential interactions between the capsid and host cell receptor in pathogenesis are poorly understood (15, 112, 189, 196). Structural characterization of the MVM capsid-glycan interactions were conducted in this study utilizing glycans, 3'SIA(Le<sup>X</sup>)<sub>3</sub> and 3'SIA(LN)<sub>3</sub> that were recognized by MVMP, MVMi and the MVMP virulent mutants, and GT3 that was specifically recognized by MVMi (209). Also, the recognition of 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan (a tumor cell marker) by MVM, and the recognition of GT3 (neuronal marker) by MVMi explained MVM's tropism for transformed cells and MVMi's tropism for neuronal cells (209).

In all the MVM-receptor complex structures determined in this study, the SIA glycans bound in the vicinity of the same site at the icosahedral twofold axis, that was identified previously when SIA alone was soaked into MVMP-VLP crystals (Figure 1-5 and Figure 2-2) (188). The MVM SIA binding pocket is shallow and surrounded by charged and hydrophobic residues (VP2 residues in MVMP: K241, M243, I362, K368, R375, Y396, W398, D399, D553, Y558, and T578) that would accommodate and stabilize a long chain of sugar molecules. Significantly, the residues determining *in vitro* tropism and *in vivo* pathogenicity (residues 317, 321, 362, 368, 399, 553 and 558) are localized in the vicinity of this SIA binding pocket (7, 15, 112, 196). Such a pocket profile is a common structural feature for other virus-SIA interactions (47, 90, 111, 338). In the capsid-glycan complexes, all the carbohydrate moieties that compose a particular oligosaccharide were not observed in the crystal structure. This observation suggests that the glycans at the reducing termini that were not observed in the crystal structure,

may not be forming tight interactions with the capsid surface, or that only the SIA conservatively interacts at all the sixty binding sites of the capsid, with the remaining glycans adopting different orientations. These two possible scenarios would be inconsistent with the icosahedral symmetry imposed during the structure determination, and would lead to lack of ordering of the density for these molecules. Due to these aforementioned reasons, the side groups for some of the carbohydrates structures were not visible in the density map but were modeled based on clashscore and bonding interactions.

In the MVMP-3'SIA( $\text{Le}^X$ )<sub>3</sub> complex, the glycan extends towards the wall of the threefold protrusions and it can be speculated that the carbohydrate molecules at the reducing end that are not ordered in the structure could interact with the V325 which is on the wall of the threefold protrusion (Figure 2-2B). Similar possibilities occur in the MVMi-3'SIA( $\text{Le}^X$ )<sub>3</sub> and MVMP-3'SIA(LN)<sub>3</sub> complexes (Figure 2-2D and Figure 2-3B). The MVMP-V325M is a virulent mutant and V325 has been shown to modulate SIA binding in a manner similar to I362 and K368 (188, 209, 259). The 3'SIA(LN)<sub>3</sub> and 3'SIA( $\text{Le}^X$ )<sub>3</sub> glycans bind along the wall of the pocket on the MVMP capsid and although the terminal SIA can be superposed, the rest of the molecule adopts a slightly different conformation as evident in Figure 2-4A and Figure 2-4D. The differences in the conformation could be attributed to the differences in glycan composition. The 3'SIA( $\text{Le}^X$ )<sub>3</sub> glycan contains additional fucose group as compared to the 3'SIA(LN)<sub>3</sub> glycan. The 3'SIA( $\text{Le}^X$ )<sub>3</sub> glycan contains the 3'SIA- $\text{Le}^X$  tumor cell marker that is suggested to be utilized by the MVM viruses to bind to cancer cells (139, 151, 152, 182). Differences in the glycan

composition might also explain the different binding pattern observed in the MVMi-3'SIA(Le<sup>X</sup>)<sub>3</sub> and MVMi-3'SIA(LN)<sub>3</sub> complexes (Figure 2-3 and 2-4).

Comparative analysis of all the complex structures shows that the binding pocket can accommodate the various glycans in different conformations (Figure 2-4C). Since the terminal SIA is recognized by all the MVM viruses, the SIA from all the capsid-glycan structures determined in this study were superimposed to understand the specificity of SIA binding to these viruses (Figure 2-4F). The SIA conformation is similar for the MVMP complexes (Figure 2-4D) but not for the MVMi complexes (Figure 2-4E), which makes it difficult to define the role of specific amino acid residues in dictating the glycan binding specificity. The heterogeneity in the terminal SIA binding observed for these structures could be due to the promiscuous nature of the binding pocket, and/or limitations of the crystal structure determination studies, such as the merging of separate crystal diffraction data sets to achieve more data completeness, probable twinning of the MVM crystals and icosahedral averaging. Although, the capsid-glycan interactions were not conserved, a few capsid surface residues made non-specific contacts with the glycans on all the MVM capsids. Such as the residue 558, which is a cell tropism determinant and makes contacts with all the glycans whether on MVMP or MVMi capsid. Interaction with the residue 368 which is also a cell tropism determinant is stronger for the MVMP-glycan complexes (2.5 - 3.0 Å) as compared to the MVMi complexes (greater than 4.0 Å). Also, R368 is involved in hydrogen bonding interactions with D399 and E321 in MVMi; and with D399 in MVMP-K368R, which are not present in MVMP (which has K368, D399 and G321) and MVMP-K/I (D399 points away from R368) (Figure 1-7) (7, 168). Residue 399 interacts with the glycans on the MVMP capsid

but not on the MVMi capsid, except in the case of α2-8 linked GT3 glycan. In the previously solved MVMP-SIA glycan complex structure, residue D399 was the only residue within interaction distance (3.2 Å) (Figure 1-5). These observations suggest that D399 is not necessary for binding to α2-3 linked SIA glycans on the MVMi capsid, but is required for binding to α2-8 linked glycans. It is known that a forward second site mutation that confers fibrotropism to MVMi occurs at D399 (D399 mutated to A or G results in loss of hydrogen bonding interaction with R368) when either 317 or 321 are mutated; and D399 is essential for infection in lymphocytes (7, 63, 79). Also, mutation at D399 (D399A or D399G) in MVMi accompanied with a second non-coding change in the NS gene (that results in the accumulation of NS2) confers fibrotropism (7, 63, 79). Towards dissecting the role of D399 in tissue tropism, the MVMi-D399A mutant was screened on the CFG Mammalian Printed Array Version 4.1 alongside wt MVMP and MVMi (data not shown). The data showed that this mutant binds to α2-3 linked SIA but not to α2-8 linked SIA, confirming the suggestion that D399 is not necessary for binding of α2-3 linked SIA on MVMi capsid, but it is important for α2-8 linked SIA recognition. Residue 368 might have a direct effect on the SIA binding or an indirect effect through interaction with D399. The D399-R368 interaction present in MVMi and MVMP-K368R mutant must configure a capsid surface topology required for binding to glycans with α2-8-linked SIA. Apart from residue 368 (conserved between MVMi and MVMP-K368R which bind to GT3) and 399, there are other residues that are involved in unique interactions with GT3 such as M366 and E321 (also cell tropism determinants) that might be neurotropic determinants. The interaction between D399 and R368 is present in the MVMP-K368R mutant, but not the R368 and E321 interaction (G321 in MVMP),

which might explain its lower affinity to GT3 than MVMi, and its fibrotropic phenotype *in vitro* (which is different from MVMi's lymphotropic phenotype) (188, 209). Apart from Y558, there are other conserved residues that make contacts with the glycans on both the viruses, such as R375, E394, Y396, T397 and 575 (a leukopenia associated mutation).

The binding affinity for the 3'SIA(LN)<sub>3</sub> was MVMP=MVMi>>MVMP-K368R>MVMP-I362S=MVMP-K/I (209). Also, the MVMP-K/I mutant showed the most reduced affinity to the 3'SIA(LN)<sub>3</sub> glycan than MVMP, MVMi or MVMP-K368R, but was the most pathogenic *in vivo* (188, 209). In the MVMP-K/I mutant, the mutation at residue 362 that changes Ile to Ser (instead of Valine as in MVMP-K368R and MVMi) and the different conformation adopted by D399 in the MVMP-K/I mutant compared to the MVMi or MVMP-K368R mutants (although R368 is common to these three viruses) seems to be involved in its non-recognition of GT3. MVMP-I362S and MVMP-K/I have the same reduced affinity for the 3'SIA(LN)<sub>3</sub> glycan which implicates residue 362 in regulating the binding specificity to the α2-3 linked SIA glycans, while residue 368 regulates the binding specificity for both the α2-3 linked and α2-8 linked SIA glycans. Although, no binding interactions with residue 362 were observed in the complex structures analyzed in this study, or in the SIA soaked MVMP-VLP structure solved previously, this residue lies in the vicinity of the pocket (188). These structure determination studies structurally verified the role of certain residues such as, 368, 399 and 558 in glycan binding but could not identify the determinants of SIA binding specificity. The SIA receptor binding pocket is flexible and colocalizes with the tropism and pathogenicity determinants and also allows for differential engagement of the SIA receptor. The differences in the

capsid-receptor interactions suggest a role in dictating pathogenicity but could not explain the differences in cell tropism for the MVM strains and mutants.

### **Recognition of Sialic Acid Derivatives by MVM Viruses**

Previous glycan array screening studies focused on identifying the SIA motif(s) recognized by MVM among a large number of different glycans (209), while screening on the SGM array enabled the exploration of the potential role of SIA derivatives in dictating differential MVM strain recognition. Three types of MVM capsids, VLPs (no genome and assembled from VP2 alone), Empties (no genome and assembled from VP1 and VP2), Fulls (contains wt genome and assembled from VP1, VP2 and VP3) were screened on the SGM array to investigate if their different composition results in differences in glycan binding properties. The diversity of SIA modifications on this array allowed for the detection of novel MVM glycan interactions. On this array, the glycans are sorted for (a) discrimination of  $\alpha$ 2-3 and  $\alpha$ 2-6 linkages, (b) derivatives of Neu5Ac, Neu5Gc or Kdn (defined in chapter 1), and (c) underlying structures derived from lactose, LNnT, NA2 or LNT (defined in the experimental methods in this chapter).

Depending on the strain, MVMP or MVMi, the glycan binding profile for VLPs, empty and full capsids was similar (Figure 2-5 and Figure 2-6). The data for the full MVMP viruses, however, exhibited high background noise as evident by the high %CV (Figure 2-6A). The recognition of similar glycans by full and empty capsids of the particular strain indicates that the encapsidation of DNA by the full capsids followed by its maturation i.e cleavage of VP2 to give VP3 does not affect its receptor binding ability compared to empty capsids, and this agrees with previous observations (181, 282). Also, the similar recognition profile for empty capsids and VLPs validates previous claims that they are antigenically and structurally equivalent (132).

Complex N-glycans are reported to have two types of LacNAc (Gal-GlcNAc) repeats; Gal $\beta$ 1-3GlcNAc (Type1) and Gal $\beta$ 1-4GlcNAc (Type2), of which Type2 repeats are more common. All the MVM viruses bound to SIA in only  $\alpha$ 2-3 linkages to Type2 LacNAc (Gal $\beta$ 1-4GlcNAc) of LNnT and NA2 precursor, which is consistent with the previous glycan array screening data (209) (Figure 2-5 and Figure 2-6). This suggests that the type of Gal-GlcNAc linkage plays a role in the receptor-capsid interaction. Preliminary modeling of the Type1 LacNAc glycan (such as Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-3GlcNAc) in the MVM SIA binding site (guided by the capsid-glycan complex structures determined in this study and the MVMP-VLP-SIA complex structure solved previously (188) indicates that the kink associated with the Gal $\beta$ 1-3GlcNAc linkage creates steric hindrance in the binding site.

The MVMP and MVMi viruses showed a strong preference for binding to SIA that is methylated at position C-9 of Neu5Ac (SGM 23 and SGM 55), compared to the other derivatives (acetylated or lactoylated Neu5Ac, such as SGM 52 and SGM60, respectively, or Neu5Gc) and also non-derivatized Neu5Ac glycan (SGM 48) (Figure 2-6). The hydrophobic methyl group specifically at position C-9 and not at C-8 of Neu5Ac is preferred by both MVMP and MVMi. Among the SIA glycans methylated at position C-9, binding affinity of SGM 23 (methylated, monosialylated but single antennae) is higher than SGM 55 (methylated, monosialylated but biantennary). The specific recognition of Neu5Ac9-O-methyl derivative of sialylated glycans by both MVMP and MVMi is interesting but since this is a synthetic derivative that has not yet been isolated in nature, no further conclusions can be made. Also, preliminary modeling studies of SIA glycans methylated at C-9 or C-8 into the MVMP glycan binding site did not provide any

clues to this differential specificity. The C-8 or C-9 of Neu5Ac modeled in the complex structures (i.e., structures determined in this study and the MVMP-VLP-SIA complex structure solved previously (188)) is not in close proximity to any capsid surface amino acid residues to have any effect on the glycan binding interaction. Even though the binding for the biantennary glycan SGM 48 by MVMP and MVMi viruses is low as compared to the Neu5Ac methylated derivatives, it is still the highest among the Neu5Ac non-derivatized glycans which suggests a strong preference of MVMP and MVMi for this glycan.

MVMP viruses showed binding to the 9-O-acetylated Neu5Ac derivative (SGM 52) (Figure 2-5, and Figure 2-6A, B and C). Binding to glycans containing Neu5,9Ac is significant, because it has been seen that SIA 9-O-acetylation is upregulated (e.g., 9-O-acetylated GD3) in melanoma cells in humans (249). The enveloped viruses, *Bovine Coronavirus* (BCoV) and Influenza C virus also require 9-O-acetylated SIA containing receptor for a successful infection. These viruses possess 9-O-acetyl esterase activity that promotes escape of virion progeny (266). None of the MVMP or MVMi viruses bound to Neu5Gc or its derivatives even though it exists in mammals, and other parvoviruses such as AAV5 and CPV have been reported to bind it ((19, 303), and unpublished data). So, an N-glycolyl modification at position C-5 of Neu (as in Neu5Gc) is not preferred for binding as compared to an N-acetyl modification at the same position (as in Neu5Ac). Also, the Kdn derivative in which the amino group at C-5 of Neu5Ac is removed is not recognized. The MVMP viruses also bound to the 9-O-lactoylated Neu5Ac derivative (SGM 60), which are present in serum glycoproteins (165). This implies that the binding of MVM viruses to the sialylated derivatives is highly

dependent on the type of substitutions at position C-5 and C-9 of the sialic acid. The MVMi viruses did not bind to 9-O-acetylated (SGM 52) and the 9-O-lactoyl (SGM 60) SIA derivatives (Figure 2-5 and Figure 2-6). In the case of MVMP viruses, the SIA modifications that are recognized are: N-acetylation at C-5 (Neu5Ac) alone or with O-acetylation at C-9, or O-lactylation at C-9, or O-methylation at C-9. For MVMi, the SIA modifications that are recognized are: N-acetylation at C-5 (Neu5Ac) alone or with O-methylation at C-9. These observations suggest that the MVMP capsid allows for binding to more derivatives than the MVMi capsid, but preliminary modeling studies of these SIA derivatives on the MVMP or MVMi SIA binding site showed that the glycerol chain containing C-8 or C-9 is not involved in any capsid interactions that could explain these differences. SGM screening for the first time identifies the SIA derivatives specifically recognized by the MVM viruses, and also shows recognition for the biantennary glycans with terminal 3'SIA-LN motifs, that are present on the cell types permissive for MVM viruses (results of the glycan profiling experiment).

### **Cellular Glycan Profiling Validates Glycan Screening Data**

The glycan profiling data from the three cell lines, A9 fibroblasts, EL4 T lymphocytes and NB324K transformed cells, correlates with the previous glycan array screening and current SGM array results since the glycan motifs that were recognized by MVM viruses in these arrays are also expressed on these cell types (Figure 2-7 and Figure 2-8). The abundance and variability of SIA containing glycans on these cell surfaces, makes them an obvious choice for receptors, as exemplified by the huge number of viruses that utilize sialylated receptors. The glycan arrays attempt to mimic the diversity of glycans present on a cell surface, and if a virus specifically recognizes a particular type of glycan on these arrays, it is suggested that the particular virus

probably recognizes similar glycans in its natural environment. As shown by the glycan profiling of the three cell lines infected by MVM viruses, there are many glycans with carbohydrates other than SIA at the non-reducing end, but MVM viruses specifically recognized the SIA glycans, which is consistent with the glycan array screening and SGM data (209). In addition, glycans with the SIA derivative, Neu5Gc (recognized by AAV5 and CPV) are also present in all the three cell lines, as in the glycan array and SGM, but MVM viruses did not recognize these derivatives in either array ((19, 209, 303), and unpublished data). Humans (but not mice) lack the hydroxylase required to produce Neu5Gc from Neu5Ac but we do observe a Neu5Gc containing glycan (enclosed in dashed black box in Figure 2-8C and F) in the glycolipids fraction of NB324K cells which are human transformed cells (309). This might be explained by the ability of human cells to incorporate Neu5Gc from mammalian foods, particularly red meat and milk (309). Also, cancer cells have been found to express higher amounts of Neu5Gc on their surface (309). Contamination due to fetal bovine serum glycoproteins might also explain this observation.

Glycans containing the SIA-LN (Neu5Ac-Gal-GlcNAc) motif (for example, present in the glycans enclosed in the green or blue box in Figure 2-7D, E and F, or black box in Figure 2-8A) were recognized by both MVMP and MVMi in glycan array and SGM screening, and their expression in all the three cell lines would explain the common recognition of these glycans by both MVMP and MVMi (Figure 2-5) (209). Also, the biantennary complex N-glycan, SGM 48, which was recognized by both MVMP and MVMi viruses in the SGM array, is present in the low mass fraction of A9 cells and NB324K cells ( $m/z=2792.2$ , enclosed in dashed red box in Figure 2-7A and C). The

glycan SIA-LN-LN that was recognized in previous glycan array screening is possibly (based on the linkage combinations allowed) present in the NB324K medium mass N-glycan fraction ( $m/z=4850.2$ ) (enclosed in a dashed green box in Figure 2-7F).

The gangliosides, although broadly distributed, are predominantly expressed in the brain, whereas the neolacto-series glycolipids are common on certain hematopoietic cells such as leukocytes (265). The presence of the three common glycan structures (GM3, GM2, and a Neu5Gc containing glycan; enclosed in red box in Figure 2-8) in both the polar and non-polar glycolipids fraction in all the three cell lines suggests that these glycans do not differentiate between the polar and non-polar glycolipids present on these cell surfaces. The GT3 ganglioside present in the A9 fibroblasts (enclosed in green box in Figure 2-8A) was recognized by MVMi in the previous glycan array screening (209). There was no binding to monosialylated or GM gangliosides in the previous glycan array screening (209), even though these are the most highly expressed gangliosides in the glycolipids fraction for all the three cell lines (Figure 2-8). All the  $\alpha$ 2-8 multisialylated glycans that were recognized by MVMi viruses in the glycan array screening are present only in the polar glycolipid fractions in these cell lines. However, previous experiments have shown that a SIA containing glycoprotein (and not glycolipid) is utilized as a receptor for MVM infection (78, 188). Although the presence of multisialylated glycans on glycolipids has been well documented in the past, several recent studies have shown that  $\alpha$ 2-8 multisialylated glycans, especially GT3 and GD3 are also present on glycoproteins in mouse brain (138, 262, 263), which would be consistent with MVMi viruses utilizing an  $\alpha$ 2-8 multisialylated glycoprotein for infection in neuronal cells. Multisialylated glycoproteins with  $\alpha$ 2-8 linkages, such as Neural Cell

Adhesion Molecule (NCAM), and CD166 (or Activated Leukocyte Cell Adhesion Molecule (ALCAM)), are commonly found expressed on the neuronal cells and may serve as receptors for MVMi on these cells (263, 271).

The O-glycans have GalNAc at their reducing end and are as such classified as the mucin type (Figure 2-9) (41). The sialylated T-antigen ( $m/z$  895.5) is annotated only in the NB324K cell line which is consistent with the observation of this motif on the cell surfaces of many leukemia and tumor cells (41). Results from previous biochemical and screening assays could not conclude whether the sialylated glycans recognized by MVM were N-linked or O-linked (209). But the absence of sialylated O-linked glycans and the abundance of N-linked sialylated glycans on these cell surfaces would imply that MVM binds to N-linked sialylated glycoproteins on these cell surfaces.

The NB324K cell line is a transformed cell line that expresses sialylated N-glycans with the SIA-Le<sup>X</sup> motif (e.g.,  $m/z$ =3142.4,  $m/z$ =3952.3, etc.; enclosed in red box in Figure 2-7C and F) that was recognized previously by the MVM viruses on the glycan array (209), and is a known carbohydrate marker for cancer cells (139, 151, 152, 182). The expression of this motif has also been specifically observed in the N-glycan profile of other transformed cell lines such as THP-1 (monocytic leukemia cell line), HL-60 (promyelocytic leukemia cells), and K562 (erythromyeloblastoid leukemia cell line) and this data is available at the CFG database (<http://www.functionalglycomics.org/glycomics/publicdata/glycoprofiling-new.jsp>). This validates the suggestion that the SIA-Le<sup>X</sup> motif on the tumor cell lines is utilized by the oncotropic MVM virus to bind to these cells. NB324K cells can be infected by both MVMP and MVMi, while A9 fibroblasts can be infected only by MVMP and EL4 T

lymphocytes can be infected only by MVMi. Comparing the SIA glycan profile (because it has been shown that MVM viruses only bind to SIA containing glycans) of these three cell lines, two glycans were identified to be present in A9 cells and NB324K cells but not in the EL4 T cells ( $m/z=2792.2$ , enclosed in dashed red box in Figure 2-7A and C; and  $m/z=3603.5$ , enclosed in green box in Figure 2-7D and F). The glycan with  $m/z=2792.2$  (enclosed in dashed red box in Figure 2-7A and C) was also recognized by both MVMP and MVMi in the SGM screening (SGM 48), which suggests that EL4 T cells might restrict infection by MVMP due to the absence of this glycan on their cell surface. Since this glycan was recognized by MVMi and is present on A9 cells, it still unclear why MVMi cannot bind and infect these cells. Both A9 and EL4 T cells have a similar glycan profile which fails to explain the differences in the cell tropism between these viruses based on receptor interaction. However, the glycan profiling data does confirm the role of sialylated glycans in MVM infection.

## Summary

Structural studies of the homologous MVM strains and mutants in complex with the respective receptors provided an understanding of how a single or few amino acids substitutions affect receptor binding specificity. A common receptor binding pocket at the twofold axis was utilized by all the MVM viruses but the glycan conformation and the capsid-receptor interactions were different. In the MVMP-3'SIA( $\text{Le}^{\text{x}}$ )<sub>3</sub>, MVMP-3'SIA(LN)<sub>3</sub>, MVMP-3'SIA( $\text{Le}^{\text{x}}$ )<sub>3</sub> and MVMi-GT3 complexes, longer oligosaccharides were visualized. In the complexes where longer oligosaccharides were observed, interaction with the residue V325 (virulence determinant) was possible. Residue D399 is not required for binding of  $\alpha$ 2-3 linked SIA on MVMi capsid but is a determinant for binding to  $\alpha$ 2-8 linked SIA glycans. This study also suggested that residues 558, 375, 394 could dictate

SIA binding, and residues 321 and 366 might dictate the neurotropic properties of MVMi virus. The binding profile for VLPs, empty and full particles was similar on the SGM which validates the use of VLPs in lieu of infectious virions for structural and biochemical studies that examine receptor interactions. MVMP and MVMi specifically bound only to the  $\alpha$ 2-3 linked sialylated derivatives, which is consistent with the previous glycan array screening. Both MVMP and MVMi showed a preference for binding to sialic acid that was methylated at C-9 compared to the other derivatives. All MVMP viruses also bound to 9-O-acetylated SIA and 9-O-lactoylated SIA derivatives, but MVMi did not bind to these glycans. In addition, all the MVM viruses bound to a biantennary SIA glycan with 3'SIA-LN motif that is also present in the A9 and NB324K cells. To summarize the SGM data, MVMP can tolerate SIA modifications such as: N-acetylation at C-5 (Neu5Ac) alone or with O-acetylation at C-9, or O-lactylation at C-9, or O-methylation at C-9; while for MVMi only N-acetylation at C-5 (Neu5Ac) alone or with O-methylation at C-9 is recognized. Glycomic profiling of A9 fibroblasts (infected by MVMP but not MVMi), EL4 T lymphocytes (infected by MVMi but not MVMP) and NB324K cells (infected by both MVMP and MVMi) validated the presence of the glycans that were recognized by the MVM viruses in the previous glycan array and the current SGM screening, such as glycans with SIA-LN motif and SIA-Le<sup>x</sup> motif, the  $\alpha$ 2,8 multisialylated glycans, and the biantennary glycans. The expression of sialylated N-glycans with the SIA-Le<sup>x</sup> motif only on the NB324K transformed cell line implies that the MVM viruses utilize the SIA-Le<sup>x</sup> motif on the tumor cell lines to bind to these cells. The biantennary glycan that was recognized by MVM viruses in the SGM screening is

present in A9 and NB324K cells but not in EL4 T cells. This suggests that EL4 T cells might restrict infection by MVMP due to the absence of this glycan on their cell surface.

A common SIA binding pocket is utilized by both the MVM strains and the capsid-SIA receptor interactions involved residues known to play a role in dictating differences in cell tropism and pathogenicity which suggests that the flexible receptor binding pocket does play a role in cell recognition, especially cancer cell recognition. The differences in the capsid-SIA receptor interactions might dictate the differences in pathogenicity but could not explain the differences in cell tropism between the MVM strains. The differences in cell tropism and pathogenicity have been mapped to the VP2 at the twofold depression in the vicinity of the receptor binding pocket (15, 112, 196). Previous studies have shown that both MVM strains bind and enter the same cells (282) and the glycan profiling data also showed that similar glycans are expressed on both A9 and EL4 T cells. Also, hybrid cells from fibroblasts and lymphocytes can propagate both strains of MVM (294), suggesting that the block is due to the lack of a differentiation-dependent cellular factor in restrictive cells. This would suggest that the block in infection is post entry but regulated by the capsid twofold pocket. It is possible that post cellular entry, the virus utilizes the twofold pocket to interact with cellular host proteins once the SIA receptor is recycled and the absence of such a cellular protein in the non-permissive cell line restricts virus infection. Another possibility is that differential utilization of the SIA receptor by MVMP and MVMi could trigger different cell signaling pathways in the permissive and non-permissive cell lines resulting in different outcomes of virus infection. Also, SIA receptor binding and/or the low pH encountered during endocytic trafficking might trigger structural transitions in the capsid that facilitate

interactions with cellular proteins that lead to successful infection. There are a number of other viral systems in which tropism appears to be controlled by the coat protein, yet restriction appears to act at an intracellular stage, such as *Hamster polyomavirus* (81), and *Human Immunodeficiency Virus* (HIV) for which the tropism determinant maps to the *env* coding sequence but appears to act after receptor binding and prior to reverse transcription (291).

Table 2-1. Data processing and refinement statistics

Parameter	MVMp-s(LN) <sub>3</sub> Soak	MVMp-s(Le <sup>X</sup> ) <sub>3</sub> Co-crystal	MVMI-s(LN) <sub>3</sub> Soak	MVMI-s(Le <sup>X</sup> ) <sub>3</sub> Soak	MVMI-GT3 Soak	MVMp-K/I- s(LN) <sub>3</sub> Co-crystal
Wavelength ( $\lambda$ , Å)	1.000	1.000	0.918	0.918	0.918	0.918
Space Group	C2	C2	C2	C2	C2	C2
Unit cell parameters (Å, °)	a=448.7, b=416.5, c=306.1, $\beta=95.7$	a=441.9, b=410.4, c=301.2, $\beta=95.9$	a=440.8, b=409.7, c=301.1, $\beta=95.7$	a=440.7, b=408.7, c=300.3, $\beta=95.7$	a=444.1, b=411.9, c=302.7, $\beta=95.7$	a=441.9, b=409.9, c=301.2, $\beta=95.9$
Resolution (Å)	40-3.7 (3.8-3.7) <sup>a</sup>	40-3.3 (3.4-3.3) <sup>a</sup>	50-3.9 (4.0-3.9)	50-3.5 (3.6-3.5)	40-3.6 (3.7-3.6)	40-3.5 (3.6-3.5)
Completeness (%)	73.4	76.1	71.1	78.4	71.4	66.7
Redundancy	1.7	2.3	2.0	2.0	2.4	2.0
$R_{\text{sym}}^{\text{b}}$ (%)	13.7	12.9	18.6	10.3	15.6	10.0
$R_{\text{factor}}^{\text{c}}/R_{\text{free}}^{\text{d}}$ (%)	34.6/34.8	31.3/31.6	39.6/39.5	39.2/39.4	35.0/35.1	34.6/35.3

<sup>a</sup>Values in parenthesis are for highest resolution shell.

<sup>b</sup> $R_{\text{sym}}=(\sum|I-\langle I \rangle|/\sum\langle I \rangle) \times 100$ , where I is the intensity of a reflection with indices h, k, l and  $\langle I \rangle$  is the average intensity of all symmetry equivalent measurements of that reflection.

<sup>c</sup> $R_{\text{factor}}=(\sum|F_o|-|F_c|/\sum|F_o|) \times 100$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factor amplitudes, respectively.

<sup>d</sup> $R_{\text{free}}$  is calculated the same as  $R_{\text{factor}}$ , except it uses 5% of reflection data omitted from refinement.

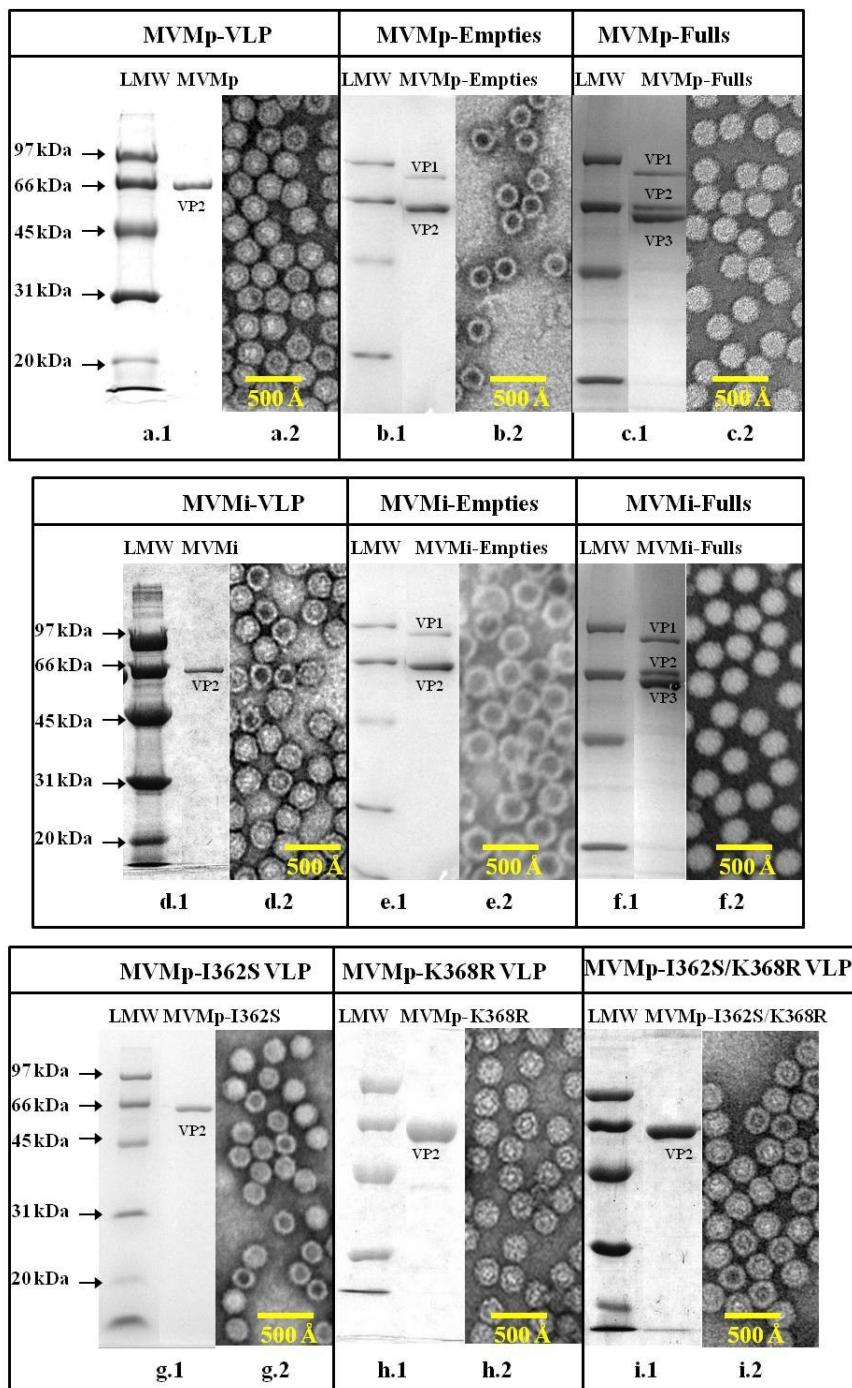


Figure 2-1. Virus purification. Coomassie stained SDS-PAGE showing the VP2 (1) for (a) MVMP-VLP, (b) Empty-MVMP, (c) Full-MVMP, (d) MVMi-VLP, (e) Empty-MVMi, (f) Full-MVMi, (g) MVMP-I362S VLP, (h) MVMP-K368R VLP, and (i) MVMP-I362S/K368R VLP; and negative stain EM showing intact capsids (2) imaged at magnification of 100,000 times for (a), (d), (i), (e), and (f); 60,000 times for (b), (c), and (g); and 40,000 times for (h).

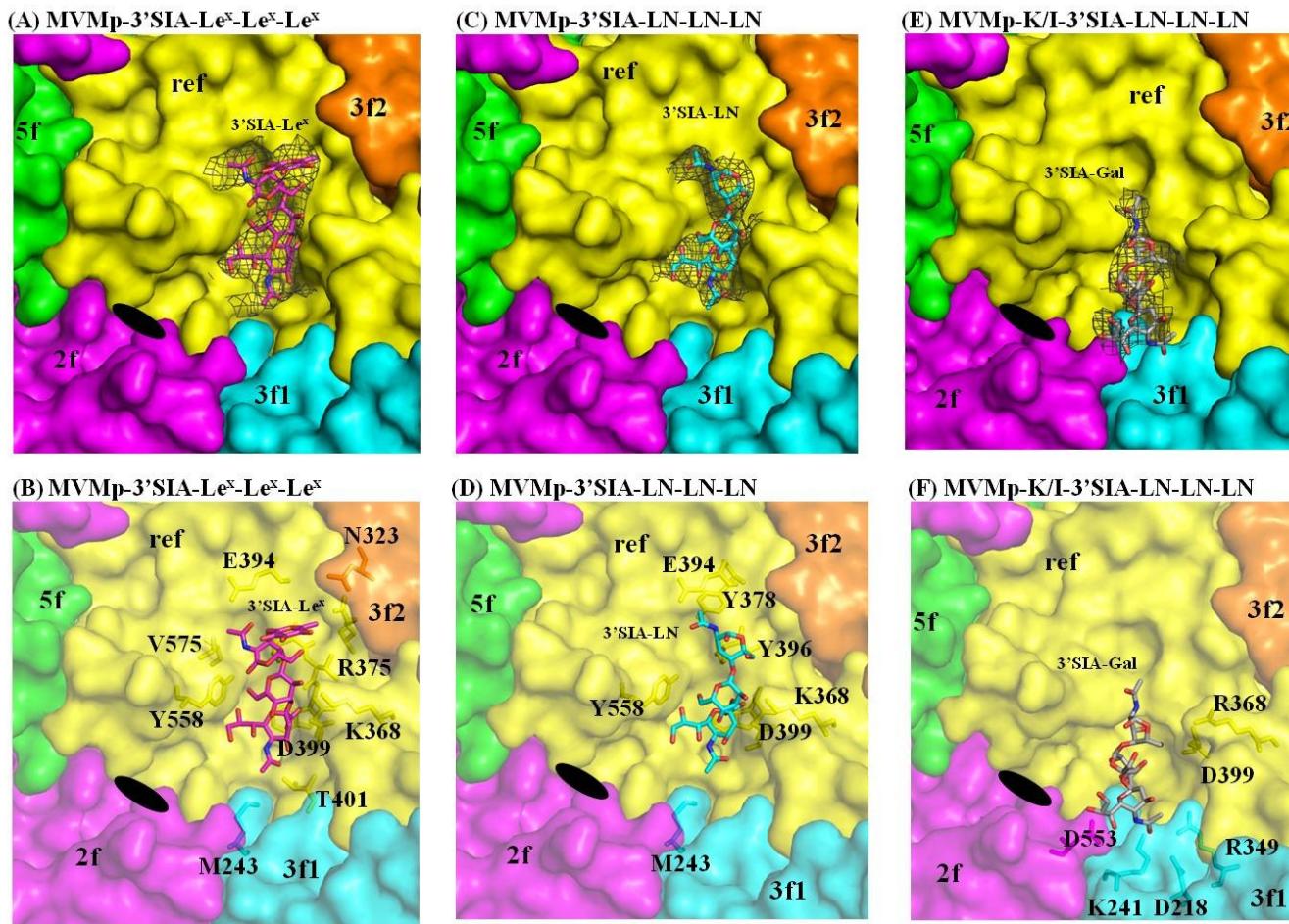


Figure 2-2. Glycan binding sites on the MVMp and MVMP-K/I capsids. (A to F) Surface representation of the MVM capsids showing the glycan binding site. The glycans are shown in the stick form inside a  $2F_o - F_c$  density map (gray mesh) in (A, C and E) and colored according to atom type (the carbon atoms are colored in magenta, cyan and grey for the MVMP-3'SIA( $\text{Le}^x$ )<sub>3</sub>, MVMP-3'SIA(LN)<sub>3</sub>, and MVMP-K/I-3'SIA(LN)<sub>3</sub> complexes, respectively; nitrogen in blue and oxygen in red). The symmetry related monomers are colored differently and labeled. The amino acid residues interacting with the glycan and/or tropism determinants are labeled. The filled oval denotes the twofold axis.

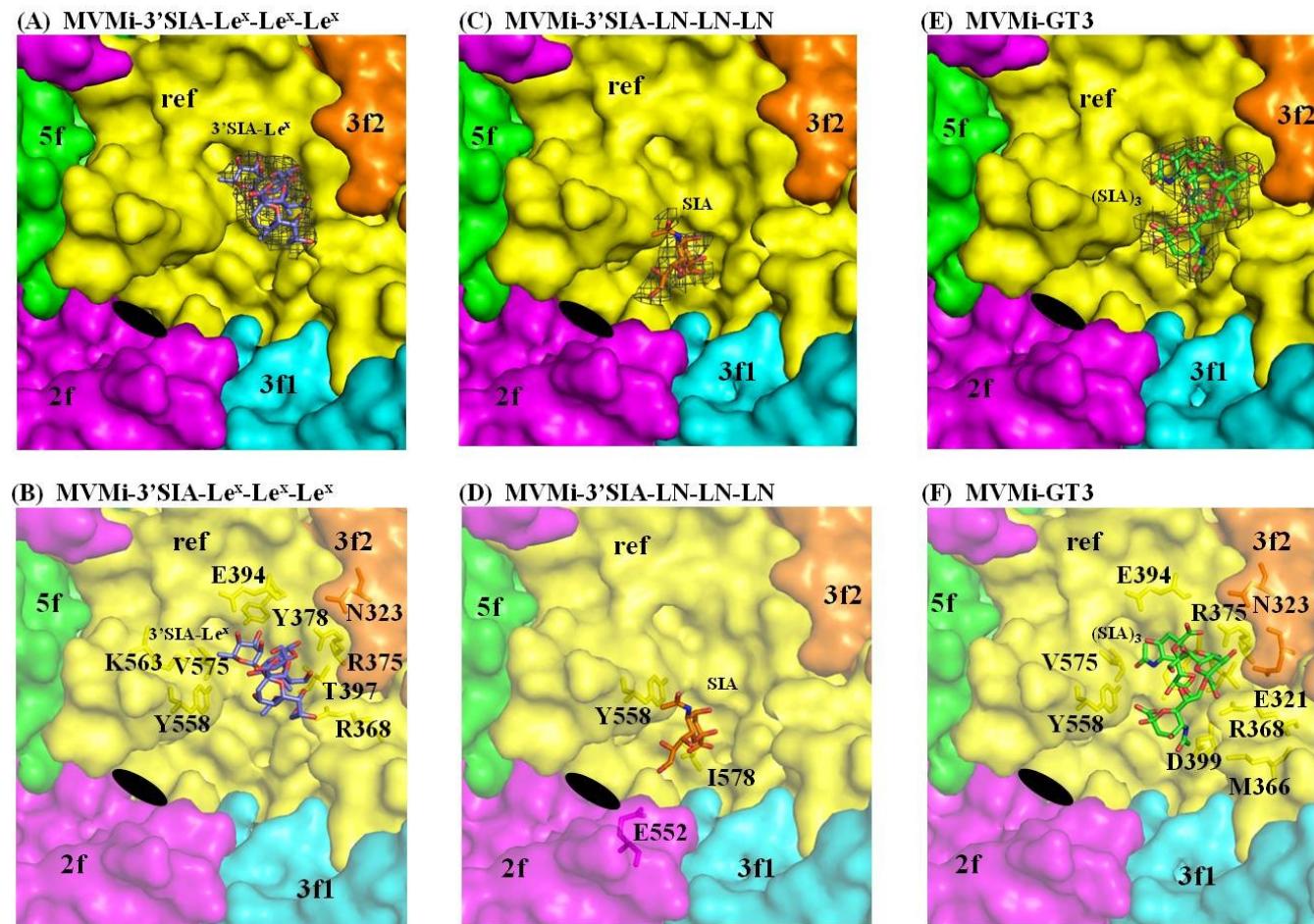


Figure 2-3. Glycan binding sites on the MVMi capsids. (A to F). Surface representation of the MVMi capsids showing the glycan binding site. The glycans are shown in the stick form inside a  $2F_o - F_c$  density map (gray mesh) in (A, C and E) and colored according to atom type (the carbon atoms are colored in blue, orange and green for the MVMi-3'SIA( $\text{Le}^x$ )<sub>3</sub>, MVMi-3'SIA(LN)<sub>3</sub> and MVMi-GT3 complexes, respectively; nitrogen in blue and oxygen in red). The symmetry related monomers are colored differently and labeled. The amino acid residues interacting with the glycan and/or tropism determinants are labeled. The filled oval denotes the twofold axis.

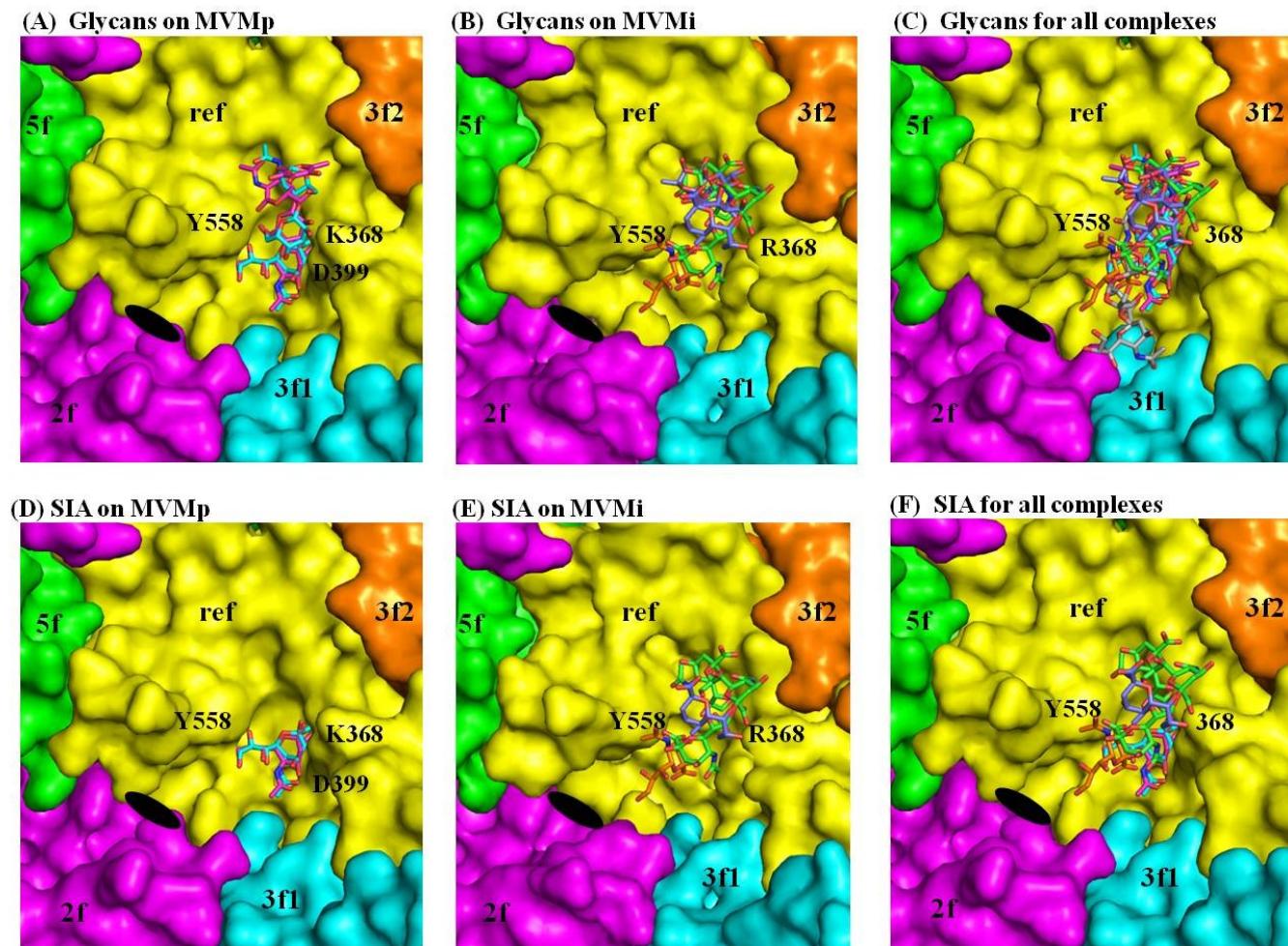


Figure 2-4. Superposition of glycans on MVM capsid. (A to F) Surface representation of the MVM capsids with the bound glycans are shown as in Figure 2-2 and Figure 2-3. (A) The 3'SIA( $\text{Le}^x$ )<sub>3</sub> and 3'SIA(LN)<sub>3</sub> glycan are superimposed on MVMP capsid, (B) The 3'SIA( $\text{Le}^x$ )<sub>3</sub>, 3'SIA(LN)<sub>3</sub> and GT3 glycan are superimposed on MVMI capsid, (C) all the glycans are superimposed, (D) SIA of the 3'SIA( $\text{Le}^x$ )<sub>3</sub> and 3'SIA(LN)<sub>3</sub> are superimposed on MVMP capsid, (E) SIA of the 3'SIA( $\text{Le}^x$ )<sub>3</sub>, 3'SIA(LN)<sub>3</sub> and GT3 are superimposed on MVMI capsid, (F) all the SIA are superimposed.

Glycan#	Glycan structures (cartoon)	Full-MVMP	Empty-MVMP	VLP-MVMP	Full-MVMi	Empty-MVMi	VLP-MVMi
23		✓	✓	✓	✓	✓	✓
55		✓	✓	✓	✓	✓	✓
48		✓	✓	✓	✓	✓	✓
60		✓	✓	✓	X	X	X
52		✓	✓	✓	X	X	X

◆ Neu5Ac    ● Glucose    ● Galactose    ■ GlcNAc    ■ GalNAc    ● Mannose    ▲ Fucose

Figure 2-5. Sialylated derivatives recognized by the MVM viruses. The glycan number and the schematic representation is shown.

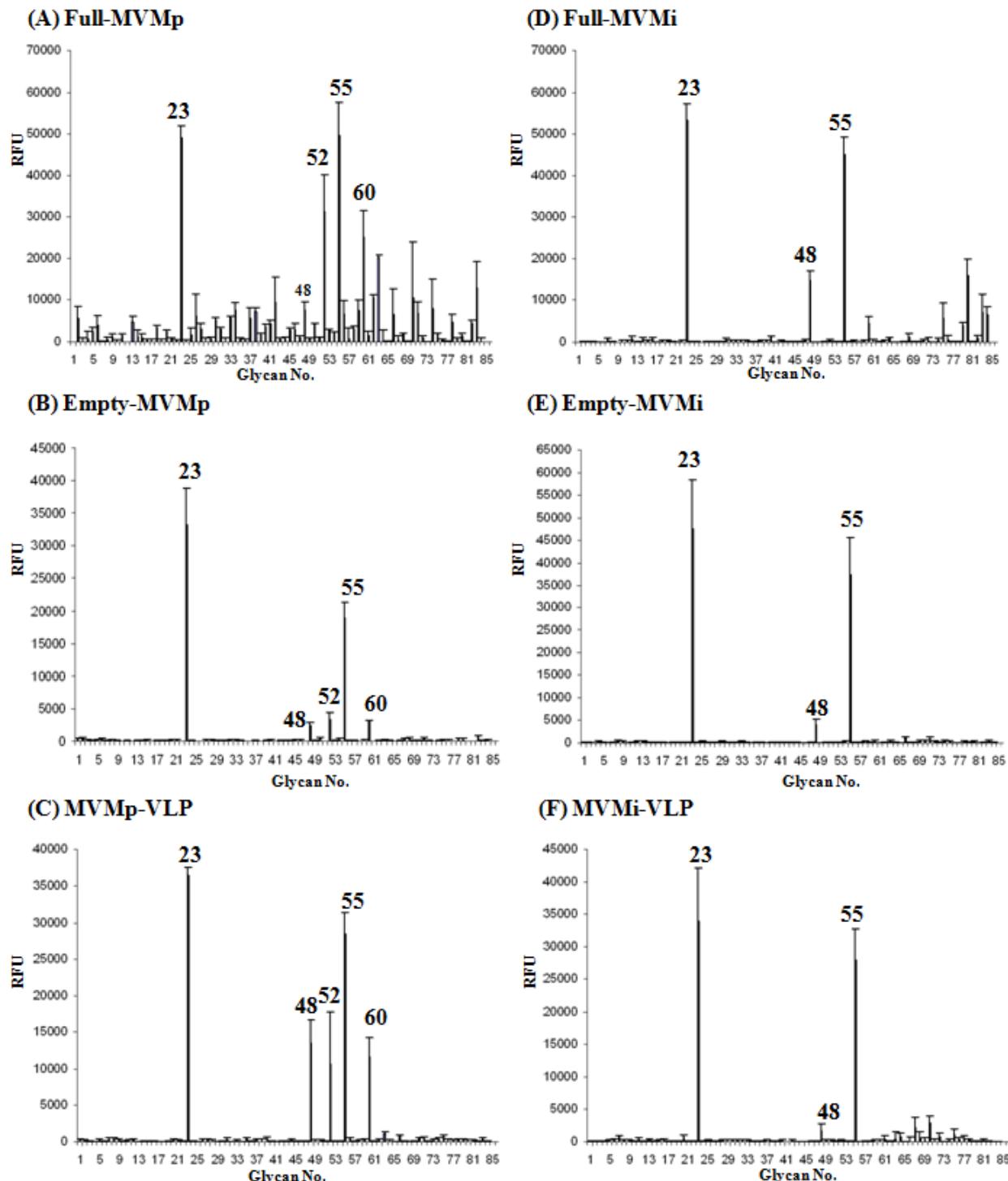


Figure 2-6. Data from the SGM array. (A) Full-MVMP, (B) Empty-MVMP, (C) MVMP-VLPs, (D) Full-MVMi , (E) Empty-MVMi-empties, (F) MVMi-VLPs. The bars represent relative fluorescence for a given glycan. The glycans showing specificity to the MVM viruses are labeled.

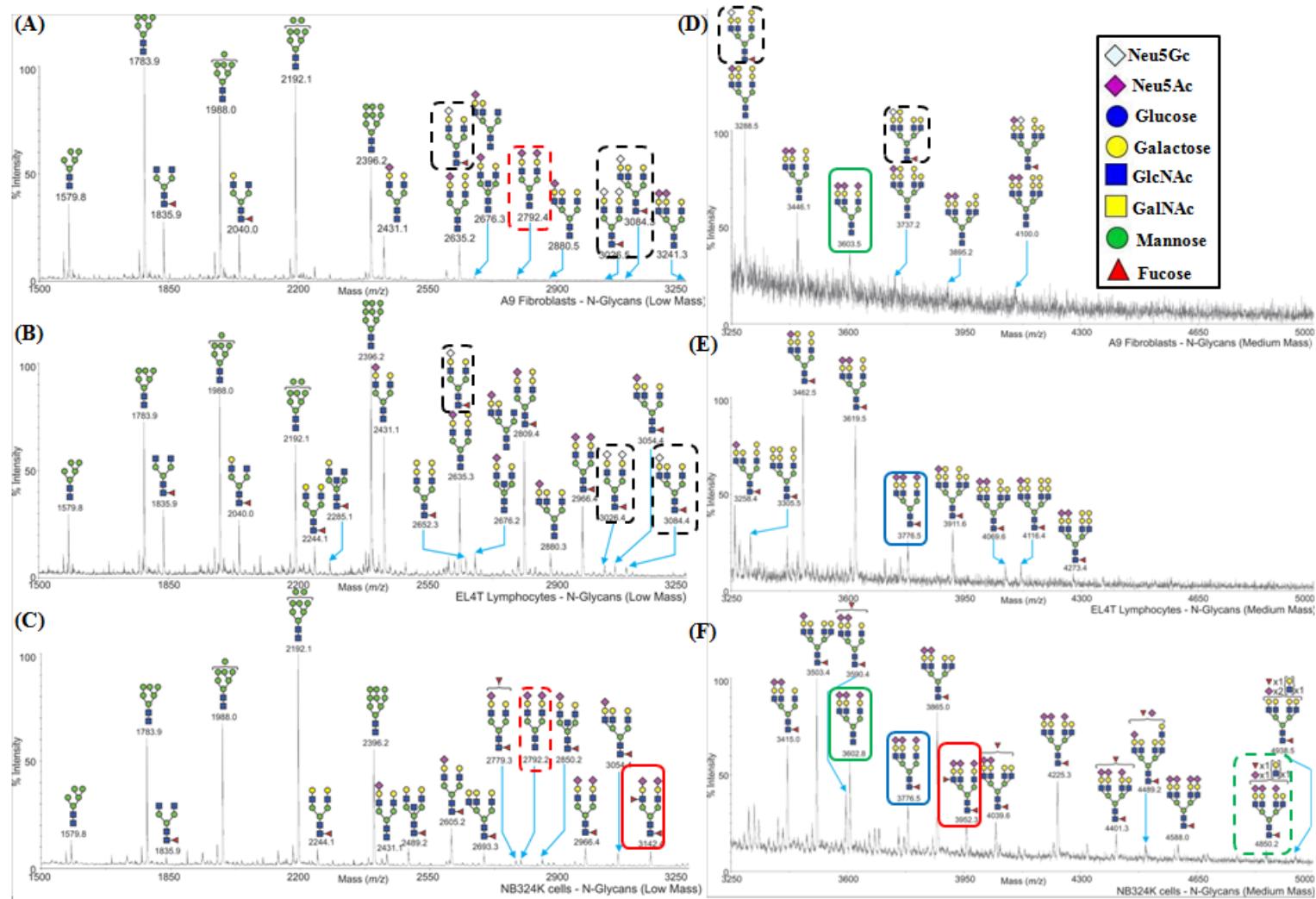


Figure 2-7. N-glycan expression on A9, EL4 T and NB324K cell lines. Low mass (A, B and C) and medium mass (D, E and F) N-linked glycans are shown. The glycans are annotated in the cartoon form. Figure inset depicts the cartoon legend. The y-axis represents relative glycan abundance.

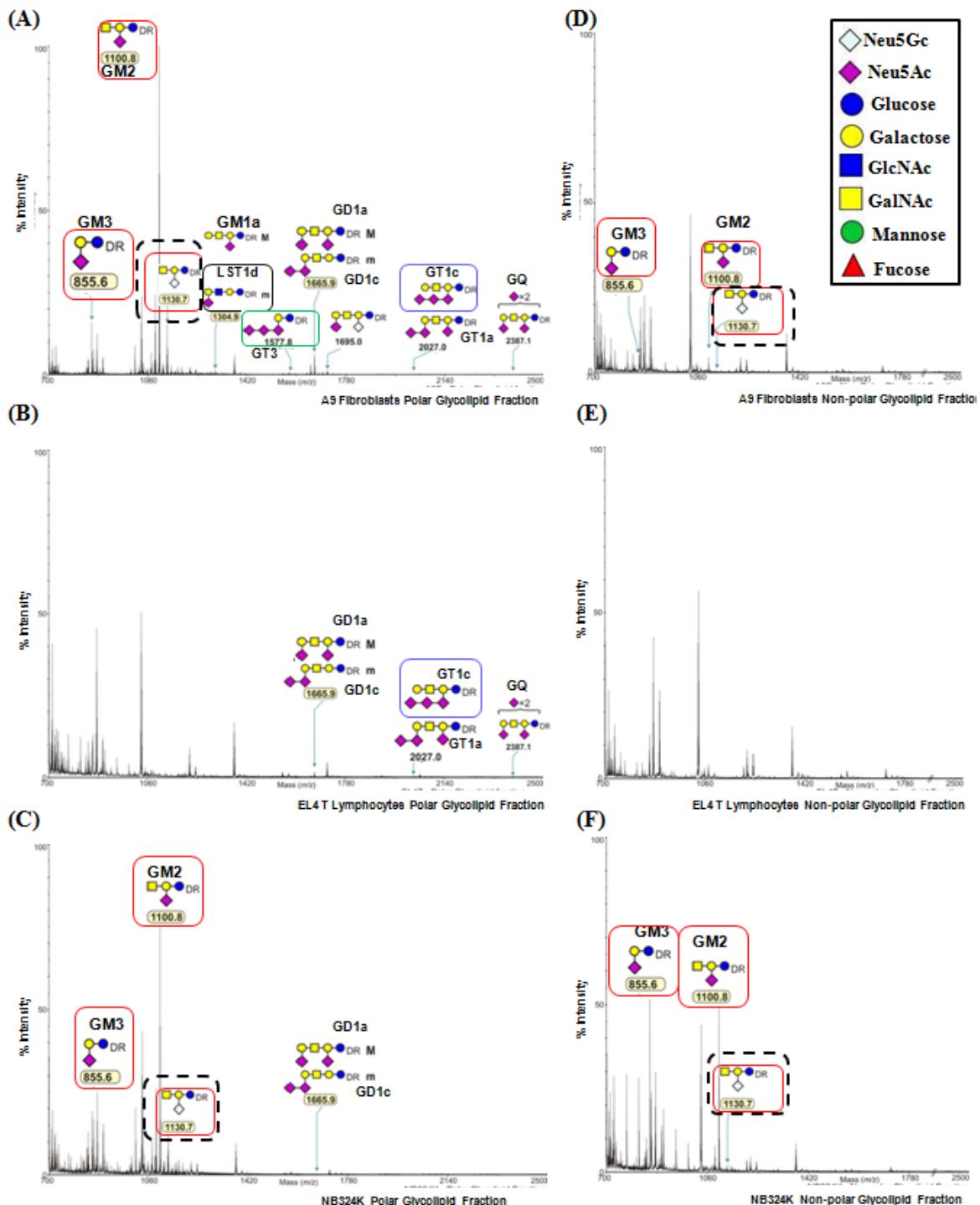


Figure 2-8. Polar and non-polar glycolipid expression on A9, EL4 T and NB324K cell lines. Polar glycolipids (A, B and C) and non-polar glycolipids (D, E and F) derived from A9 (A and D), EL4 T (B and E) and NB324K cells (C and F) are shown. The glycans are annotated in the cartoon form. Derivatives (DR) were made to differentiate between symmetrical molecules. M and m refer to major and minor species, respectively. Figure inset depicts the cartoon legend.

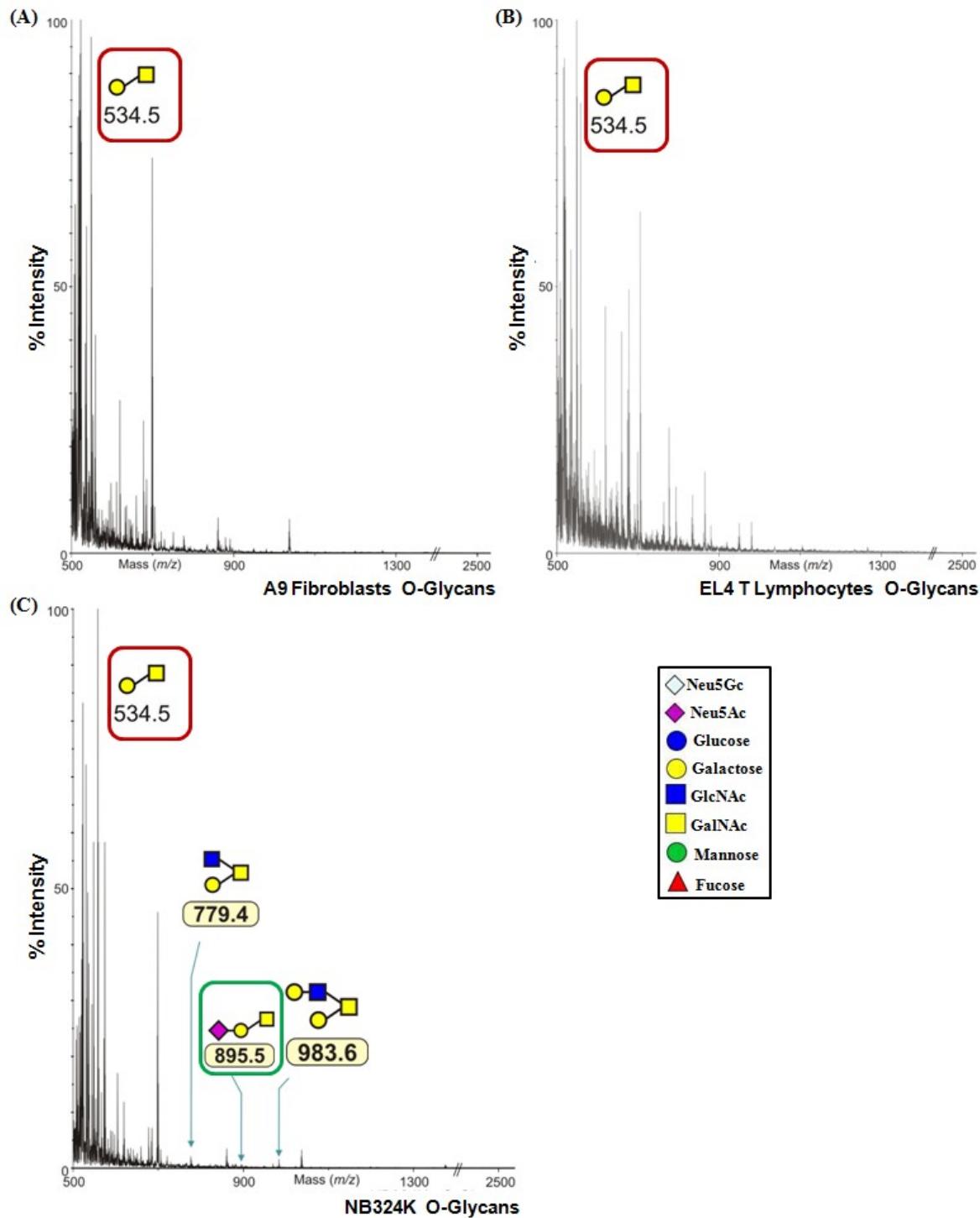


Figure 2-9. O-glycan expression on A9, EL4 T and NB324K cell lines. MALDI-TOF MS profiles of the O-glycans derived from (A) A9, (B) EL4 T and (C) NB324K cells are shown. The glycans are annotated in the cartoon form. Figure inset depicts the cartoon legend. The y-axis represents relative glycan abundance.

CHAPTER 3  
STRUCTURAL CHARACTERIZATION OF H-1PV: INSIGHTS INTO CAPSID-GENOME  
INTERACTIONS AND RECEPTOR BINDING

**Background**

H-1 parvovirus (H-1PV) is a member of the rodent subgroup of the parvovirus genus of the *Parvoviridae* and shares 66% VP2 sequence identity with both MVM<sub>p</sub> and MVM<sub>i</sub> that also belong to the same subgroup, and for which 3D structures are known (26). H-1PV was isolated from rats transplanted with HEP-1 (human liver adenocarcinoma cell ), and from aborted human fetuses (301). H-1PV, like MVM and Lulli (member of the rodent subgroup), has oncolytic properties and is selectively cytotoxic for transformed or tumor-derived cells of various species including human (discussed in chapter 1). However, it is non-pathogenic in humans and does not integrate its genome into cellular chromosomes, which makes it an excellent choice as a vector for gene therapy. The first clinical study of H-1PV was conducted in 1965 with the injection of wt virus into two osteosarcoma patients (302). Although this treatment did not completely ablate tumor development, neutralizing antibodies against the virus were found. In a recent preclinical study, H-1PV was assessed for the killing of human neuroblastoma cells and hepatoma cells, and demonstrated tumor-selective lytic effects and low toxicity for non-transformed cells (173, 201). Local or systemic treatment of advanced rat and human gliomas in rat models with H-1PV has also been observed to induced regression (114). These studies have set the stage for the first phase I/IIa clinical trial using H-1PV vectors in subjects suffering from glioblastoma multiforme (113). Most of these recombinant parvoviral vectors (based on, H-1PV, MVM, etc) targeted for tumor therapy, utilize a double-edged strategy that takes advantage of their inherent oncotropicism and selective cytotoxicity plus their ability to deliver therapeutic

genes that code for toxins such as thymidine kinase or cytokines/chemokines that enhance host immune response against the tumor cells (98, 124, 202, 221, 320). The cytotoxicity has been attributed, in part, to NS1 (39). It has been shown that interaction with host cellular factors expressed during S-phase, or interferon response mounted by the tumor cells dictates successful infection by H-1PV or MVM. However, the mechanisms underlying tumor cell recognition which is the first key step towards a successful infection, and regulates tissue tropism and pathogenesis are not yet fully understood (110, 120, 237). Glycan array screening studies of MVM viruses have suggested that their recognition and binding to cancer cells is likely mediated via interactions with glycans containing the 3'SIA-Le<sup>X</sup> motif which is a known tumor cell marker (139, 151, 152, 182, 209). Similar to MVM, H-1PV also utilizes the SIA component of a glycoprotein receptor for cell surface attachment (10, 77) and it is likely (based on its high VP2 sequence identity with MVM) that H-1PV also utilizes 3'SIA-Le<sup>X</sup> containing glycans to infect tumor cells . However, despite their common oncotropic properties, NB324K transformed cells are more permissive for infection by H-1PV than MVM (64). In a recent study (226) it was shown that the autonomous parvoviruses H-1PV, LullI, and MVM exhibit variation in tumor tropism and the oncolytic activity of LullI was mapped to the capsid VP2 gene using viral chimeras generated between LullI and MVM. The capsid was shown to assist a post-entry and pre-DNA amplification step in the viral life cycle that lead to successful infection and transformed cell killing (226). The H-1PV capsid has also been suggested to play a role in cellular tropism in another study in which different tumor cell lines were transduced with H-1PV or H-1PV/AAV hybrid vectors (172). These viruses exhibited variable tropism between tumor types. It has also

been demonstrated that the autonomous parvoviruses, such as MVM and H-1PV are less tolerant to replacement of wild-type genome sequence with a foreign sequence for vector production, as compared to AAVs (161, 174). It has been suggested that the specificity of the capsid-genome interactions, which was observed to be different for MVM and AAVs, might affect packaging of recombinant genomes and vector infectivity (7, 119, 177, 210, 214). The need to understand the role of H-1PV capsid in cell tropism, efficient therapeutic gene packaging, and neutralization by antibodies, has thus generated a need for the structural characterization of H-1PV and the mapping of capsid regions involved in these steps in the infectious life cycle. In this study, the specific carbohydrate motifs recognized by H-1PV were identified using glycan array analysis and structural studies of H-1PV alone and in complex with the identified glycans were conducted to examine the structural interactions involved in receptor attachment and possibly oncotropism. The capsid gene has been mapped as the oncotropism determinant, so comparison of H-1PV and MVM capsid structures would identify structurally variable regions that may be responsible for the differences in their oncotropism. Structure comparison of H-1PV and MVM capsids might aid in understanding the mechanism of host range switching by the MVMP virus F1 that adapted in tissue culture to infect the non-permissive rat fibroblasts (H-1PV also infects rat fibroblasts) and acquired mutations at the twofold depression on the MVMP capsid (99). In addition, the structures of wt H-1PV virions and wt H-1PV empty capsids were determined and compared in this study to investigate the capsid-genomic DNA interactions. Information on the capsid dependent mechanism of receptor recognition, genome packaging and host immune response would aid in the engineering of MVM or

H-1PV based gene therapy vectors with specific tissue targeting capability, improved therapeutic gene packaging efficiency and low immunogenicity.

## **Experimental Methods**

### **Cell Lines**

NB324K cells (simian virus-40 transformed human newborn kidney fibroblast cells) were grown in Eagle's minimal essential medium (GIBCO-BRL) with 5% fetal calf serum, glutamine, and antibiotics.

### **Virus Production and Purification**

NB324K cells grown as monolayers to 50% confluence (cell density of  $4 \times 10^6$ /10-cm plate) were infected with wt H-1PV virus for 1 h at 37°C with m.o.i of 0.1 PFU per cell and followed by occasional rocking of the plates. The cells were incubated for an additional 5-7 days till cytopathic effect was observed (~80% cell lysis). Cells were harvested by scraping and pelleted by low speed centrifugation. The cell pellet was resuspended in TE buffer (50 mM Tris-HCl pH 8.7, 0.5 mM EDTA) and stored at -20°C. The virus capsids were released from the cells by three cycles of rapid freeze-thawing. The cellular debris was removed by low speed centrifugation in a JA-20 rotor (10,000 rpm, 15 min, 4°C). The supernatant was then treated with micrococcal nuclease. Next, CsCl was added to the solution to a final density of 1.40 g/cm<sup>3</sup>, and the sample subjected to equilibrium centrifugation at 35,000 rpm in a Beckman SW41Ti rotor for 24 h at 4°C. The bands corresponding to empty capsids (1.32 g/cm<sup>3</sup>) and full virions (1.41-1.46 g/cm<sup>3</sup>) were removed from the tubes by side puncture. The fractions were subjected to another round of CsCl equilibrium centrifugation as in the previous step, bands corresponding to empty and full capsids were extracted and the fractions dialyzed into Tris-HCl Buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 8 mM

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ ). The dialyzed fraction were subjected to a third round of CsCl equilibrium centrifugation as described previously. The virus bands were extracted and finally dialyzed into Tris-HCl Buffer. The particle concentrations were determined by hemagglutination of sheep erythrocytes as well as absorbance measurements (assuming an extinction coefficient of 1.0 and 7.0 for calculations in mg/ml for empty and full particles, respectively) and adjusted to 10 mg/ml using Ultrafree centrifugal filter units (Millipore, Billerica, MA). The purity and integrity of the viral capsids was monitored using SDS-PAGE with Coomassie blue staining and negative-stain EM, respectively (Figure 3-1A, B and C). For the EM visualization, 5  $\mu\text{l}$  of purified virus sample at an estimated concentration of 2.0 mg/ml was spotted onto a 400 mesh carbon-coated copper grid (Ted Pella, Inc., Redding, CA, USA) for 1 min before blotting with filter paper (Whatman No.5). The sample was then negatively stained with 5  $\mu\text{l}$  NanoW for 1 min twice, blotted dry, and viewed on a HITACHI 3000 electron microscope.

### Glycan Array Analysis

The H-1PV empty capsids were screened to determine their glycan recognition specificity on the Mammalian Printed Array Version 3.1 containing 377 natural and synthetic glycans in collaboration with the Core H of CFG. The glycans present on this array are listed on the CFG website at

<http://www.functionalglycomics.org/static/consortium/resources/resourcecoreh11.shtml>.

The experiment was conducted as described previously (209) but with several modifications as mentioned below. To generate the printed array, a library of natural and synthetic mammalian glycans with amino linkers was printed onto N-hydroxysuccinimide (NHS)-activated glass microscope slides (SCHOTT Nexterion) in replicates of n=6. H-1PV was diluted with Binding Buffer (10 mM Tris-HCl pH 7.5, 150

mM NaCl, 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O, 1% BSA, 0.05% Tween 20), 70 µl of this diluted sample was applied to the surface of the printed array slide, a cover slip was used to cover the slide, and then the slide was incubated at RT in a dark, humidified chamber for 1 h. The cover slip was removed, the slide was then rinsed 4 times in Wash Buffer (Binding Buffer minus BSA) and 4 times in Wash Buffer without Tween 20. Seventy microliters of anti-H-1PV mouse antibody in hybridoma media (diluted 1:10 in 1XPBS) was added to the slide and it was covered by a cover slip for incubation at room temperaturefor 1 h. Following incubation, the cover slip was removed and the slide was washed as above. Finally, the slide was incubated with 70 µl of goat anti-mouse IgG-Alexa488 in 1XPBS at a concentration of 5 µg/ml for 1 h. The slide was washed as above, followed by 4 washes in distilled water. The fluorescence intensity was detected using a ScanArray 5000 confocal scanner (Perkin-Elmer, Waltham, MA). The image obtained was analyzed using the IMAGENE image analysis software (BioDiscovery, El Segundo, CA). The mean fluorescence intensity for each glycan in the printed array was obtained by removal of the the highest and lowest values from each set of six replicates followed by averaging of the four remaining readings, and the standard deviation (S.D) was also calculated. To analyze the results, all the glycans were ranked according to their signal-to-noise (S/N) ratio by dividing their mean relative fluorescence units (RFU) from the four replicates by the mean background generated in the control wells lacking glycosides. Variation within the 4 replicates was assessed as the coefficient of variation (%CV), calculated as 100 x S.D/Mean and was considered low if it was less than 30, as reported in other studies (278). The selection of glycans with specific binding affinity was based on two independent criteria: high overall total binding as measured by RFU

and low variation among the 4 replicates as assessed by %CV. Three sets of data with varying concentrations of H-1PV (200 µg/ml, 400 µg/ml and 900 µg/ml) were collected.

### **Crystallization**

Crystals of empty (no DNA) and full (DNA containing) H-1PV particles were grown using the hanging-drop vapor-diffusion method (199) with VDX 24-well plates and siliconized cover slips (Hampton Research, Laguna Niguel, CA, USA). The reservoir solution contained 1-3% (w/v) PEG 8000, 150 mM NaCl, and 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O as precipitants in 10 mM Tris-HCl, pH 7.5. The drops were prepared by mixing 2 µl of virus solution (10 mg/ml) in Tris-HCl Buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O) with 2 µl of reservoir solution and then equilibrated against 1 ml reservoir solution at RT. Crystals of empty H-1PV capsids complexed with the glycans were obtained by co-crystallization using the same crystallization conditions as above. The 3'SIA(LN)<sub>2</sub> and 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycans provided by Core D of the CFG were resuspended in 10 mM Tris, pH 7.5 to give a stock solution at 10 mg/ml. The co-crystallization drops (5 µl) were prepared by mixing 2 µl of virus solution (10 mg/ml) with 1 µl of glycan (at a concentration that results in capsid:glycan ratio of 1:180 or 1:600) and 2 µl of reservoir solution. The drop was equilibrated against 1 ml of reservoir solution at RT. Crystals obtained were soaked for 30 s in cryoprotectant solution containing the precipitant solution (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O) with 10% PEG 8000 and 30% glycerol and flash cooled in liquid nitrogen vapor prior to X-ray diffraction data collection.

### **Diffraction Data Collection and Processing**

Diffraction data were collected for crystals of full H-1PV particles at the F1 beamline at the Cornell High Energy Synchrotron Source (CHESS) on an ADSC

Quantum 270 CCD detector and empty H-1PV particles at the X29 beamline at Brookhaven National Laboratory (BNL) on an ADSC Quantum 315 CCD detector. For the H-1PV fulls, a total of 392 usable images were collected from 2 crystals with a crystal-to-detector distance of 230 and 300 mm, an oscillation angle of 0.3° per image, and an exposure time of 30 s per image at a wavelength of  $\lambda = 0.9186 \text{ \AA}$ . A total of 232 usable images were collected from a single crystal of empty H-1PV particles at a wavelength of  $\lambda = 1.0895 \text{ \AA}$  with a crystal-to-detector distance of 400 mm, an oscillation angle of 0.3° per image, and an exposure time of 20 s per image. The crystals diffracted X-rays to beyond 2.7 Å and 3.2 Å resolution for the full and empty H-1PV particles, respectively (Figure 3-1E). The measured diffraction intensities were indexed and integrated with the HKL2000 suite of programs (224), and scaled and merged with SCALEPACK (224). The crystal system was determined to be primitive monoclinic. Inspection of the 0k0 class of reflections (for k=2n) showed systematic absences for the odd reflections indicating the presence of a  $2_1$  screw axis, and thus the crystals belong to the space group P2<sub>1</sub>. The full and empty H-1PV data sets scaled with an  $R_{\text{sym}}$  of 13.2% (93.1% completeness) and 11.4% (61.3% completeness), respectively. The details of the data collection strategy and processing statistics are summarized in Table 3-1.

For the virus-glycan receptor studies, two complex structures, H-1PV-3'SIA(LN)<sub>2</sub> and H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> were determined. X-ray diffraction images were collected for empty H-1PV-3'SIA(LN)<sub>2</sub> co-crystals at the CHESS F1 beamline on an ADSC Quantum 270 CCD detector with the crystal-to-detector distance set at 350 and 400 mm. A total of 303 images were collected from two empty H-1PV-3'SIA(LN)<sub>2</sub> co-crystals with an

oscillation angle of 0.3° and an exposure time of 60 s per image at a wavelength of  $\lambda = 0.917 \text{ \AA}$ . The crystals diffracted X-rays to 2.7 Å resolution, and the  $R_{\text{sym}}$  and completeness for this dataset were 12.8% and 75.8%, respectively. The data processing statistics are given in Table 3-1. X-ray diffraction data on the empty H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> co-crystals was also collected at the CHESS F1 beamline at a wavelength of  $\lambda = 0.918 \text{ \AA}$ , oscillation angle of 0.3°, distance of 300 mm, and exposure time of 45 s per image. The total number of usable images was 234. The crystals diffracted X-rays to 2.9 Å resolution, and the  $R_{\text{sym}}$  and completeness for this dataset were 12.1% and 68.7%, respectively. The data processing statistics are given in Table 3-1.

### Structure Determination

The X-ray diffraction intensity data set collected was converted to structure factor amplitudes using the TRUNCATE program from CCP4 (Collaborative Computational Project, Number 4) (62) for the structure determination process. The Matthew's coefficient ( $V_M$ ) was calculated to be  $3.07 \text{ \AA}^3 \text{ Da}^{-1}$  for the full H-1PV data and  $2.95 \text{ \AA}^3 \text{ Da}^{-1}$  for empty H-1PV data, corresponding to solvent contents of 64% and 58%, respectively, with two particles per unit cell related by a  $2_1$  screw axis, with one particle occupying a crystallographic asymmetric unit (195). The orientations of the two H-1PV virus particles in the crystal unit cell were determined with a self-rotation function using the General Lock Rotation Function (GLRF) program (299) computed with 10% of the observed data between 10.0 and 5.0 Å resolution as large terms at  $\kappa=72^\circ$ ,  $120^\circ$  and  $180^\circ$  to search for the twofold, threefold and fivefold icosahedral symmetry axes, respectively. The radius of integration was set to 120 Å. The rotation function for  $\kappa=180^\circ$  established the orientation of the icosahedral symmetry axes relative to the crystal axes

and confirmed that the two particles in the unit cell are related by the crystallographic twofold screw axis (b axis).

The structure of the full H-1PV particle was determined by molecular replacement (258). A Ca model of the MVMP VP2 crystal structure (PDB accession no. 1Z14) (168) was generated by the MOLEMAN program (167) and expanded to 60 subunits (one capsid) using the Oligomer Generator subroutine available at the VIPER database (50) for use as a molecular replacement phasing model. Structure factors in the 10.0-5.0 Å resolution range were calculated for the phasing model using the SFALL program from CCP4 (Collaborative Computational Project, Number 4) (62) and used to calculate cross-rotation and translation functions, using the AMoRe program (212). The highest peak obtained for the cross-rotation function was used in the translation function calculation which was performed in both primitive monoclinic space groups, P2 and P2<sub>1</sub>, using the structure factor correlation coefficient (CC) as a key parameter for determining the correct solution. A clear solution to the translation search using the oriented MVMP model was found for space group P2<sub>1</sub> which gave the highest correlation coefficient (CC) of 35.4% and lowest  $R_{\text{factor}}$  ( $(\sum|F_o|-|F_c|/\sum|F_o|) \times 100$ ) where F<sub>o</sub> and F<sub>c</sub> are the observed and calculated structure factors, respectively) of 46.5%. A similar approach was used to phase the empty H-1PV data set which gave similar molecular replacement solutions to those for the full H-1PV data . Rigid-body refinement (FITTING in AMoRe) improved the CC and  $R_{\text{factor}}$  to 40.3% and 44.9%, respectively (212). The MVMP particle was then rotated and translated into the full H-1PV data set unit cell according to the final molecular replacement solutions obtained, with an orientation, in Eulerian angles, of  $\alpha=22.31^\circ$ ,  $\beta=70.69^\circ$ , and  $\gamma=28.97^\circ$  and fractional coordinate position of 0.2510,

0.0000, and 0.2489. The MVMP all atom model (PDB accession no. 1Z14) was then superimposed individually on the 60 subunits of the oriented and positioned capsid, to get an all-atom phasing model to be used in the subsequent refinement steps using the CNS program (46). For the full H-1PV data, initial phases for the oriented and positioned all-atom MVMP model were calculated to 2.7 Å resolution and further improved by refinement in the CNS program using simulated annealing, energy minimization, conventional positional, and individual temperature factor (B-factor) refinement (46). This procedure was followed by real-space electron density averaging using a molecular mask. The refinement and averaging procedures were conducted while applying strict 60-fold noncrystallographic symmetry (NCS) in CNS program (46). Five percent of the total data set was partitioned for monitoring of the refinement process with an  $R_{\text{free}}$  calculation  $((\sum|F_o|-|F_c|/\sum|F_o|) \times 100$ , calculated with a 5% randomly selected fraction of the reflection data not included in the refinement (45). Electron density that would be interpreted as residues 38 to 593 (last C-terminal residue, VP2 numbering) of the H-1PV VP2 was built into the averaged sigma-weighted  $2F_o - F_c$  electron density map by interactive substitution, insertion, and deletion of amino acids relative to the MVMP model using the COOT program (97). Following model building, new phases were calculated and improved by several alternating cycles of refinement, real-space electron density averaging, and rebuilding. To improve the quality of the maps, density map modification was carried out using the Density Modification subroutine in the CNS program which performed solvent flattening and NCS averaging (46).

Following the building of VP2 amino acids 38 to 593 into the averaged density maps, ordered unassigned density in the full H-1PV capsid interior, located in a pocket created by the reference VP monomer and an icosahedral twofold related VP monomer, was modeled as nine ssDNA nucleotides (NTs) using the ssDNA modeled in the previously solved structures of MVMi and CPV as templates (PDB accession no. 1Z1C, 4DPV, and 1P5W) (118, 168, 329). Density interpreted as two Mg<sup>2+</sup> metal ions which coordinate the phosphate backbone of the ssDNA in MVMi and CPV were also ordered in the H-1PV pocket and modeled. In addition to the stretch of nine nucleotides, three more regions of unassigned density were observed in the F<sub>o</sub>-F<sub>c</sub> density map inside the capsid and were modeled as cytosine base, with one of these located in a conserved AAV DNA binding pocket (119, 177, 210, 214). Finally, 98 water molecules and two molecules of ethylene glycol were built into the rest of the unassigned averaged positive F<sub>o</sub>-F<sub>c</sub> density (at density threshold of  $\geq 3.0\sigma$  or above) constrained by hydrogen bonding distance geometry. Significantly large peaks (F<sub>o</sub>-F<sub>c</sub> density sigma threshold  $> 4\sigma$ ), which refined with very low temperature factors (Bfactors), were re-assigned as metal ions, with a test of B-factors for Na<sup>+</sup>, Ca<sup>2+</sup>, or Mg<sup>2+</sup>. Several cycles of refinement and averaging procedures interspersed by visual inspection of the 2F<sub>o</sub>-F<sub>c</sub> and F<sub>o</sub>-F<sub>c</sub> electron density maps and model building in COOT (97) were performed till there was no longer any improvement in the agreement between the observed (F<sub>o</sub>) and calculated (F<sub>c</sub>) structure factors as given by  $R_{\text{factor}}$  (Table 3-1). The CONTACT subroutine in the program CNS was used to analyze the ssDNA-VP2 interaction within the binding pocket (46).

For the empty H-1PV data set, the refined full H-1PV VP2 structure, without the modeled nucleotides (NTs), solvent molecules, metal ions, and with the residues interacting with the NTs mutated to alanine, was used as a starting model for refinement since the molecular replacement solutions were identical. The refinement procedure was carried out as described above for the full H-1PV data. For the H-1PV co-crystal data sets, the refined full H-1PV VP2 structure alone was used for molecular replacement, and the structure refinement followed the procedure as outlined above for the full H-1PV capsid structure determination. The  $F_o$ - $F_c$  electron density maps at  $2.5\sigma$  revealed well-defined electron densities at the twofold axis for the SIA component of the complexed glycans. The coordinate files for SIA were obtained from the HIC-Up server (166) and the geometry restraints and dictionary files were generated using the subroutine phenix.elbow in the PHENIX program (2). The SIA was docked into  $F_o$ - $F_c$  density using interactive rigid-body rotations and translations in COOT. The topology and dictionary files generated for SIA in the PHENIX program were then used for subsequent refinement in the CNS program (46). The quality of the refined structures were analyzed using COOT and MOLPROBITY (56, 97). The values of root mean square deviations (rmsd) from ideal bond lengths and angles were obtained from the CNS program (46) and average B-factor for VP2 models, solvent molecules, NTs and ions were calculated using MOLEMAN (167). The refinement statistics are given in Table 3-1. The figures were generated using the PYMOL program (82).

### **Structural Alignment Comparison**

The VP2 coordinates of MVM, CPV, FPV, and PPV (PDB accession no. 1Z1C, 1P5W, 1C8F and 1K3V, respectively) were aligned with the refined H-1PV structure using the secondary structure matching (SSM) program available in PDBefold (169).

This program superimposes C $\alpha$  positions and gives the overall rmsd for the C $\alpha$  positions aligned, and also a list of the atomic distances (in Å) between each aligned C $\alpha$  position. It also provides information on amino acid residues that are structurally equivalent and gap regions. The output from this program was used to identify variable regions that contain two or more residues with C $\alpha$  atoms that are  $\geq 1.0$  Å apart between the superimposed structures, as previously defined (168).

## Results

### Virus Purification and Crystallization

The analysis of the purified empty and full H-1PV particles by SDS-PAGE showed the expected two bands corresponding to VP1 and VP2; and three bands corresponding to VP1, VP2, and VP3, respectively (Figure 3-1A). The abundance of VP2 is approximately same as VP3 in the H-1PV full virions indicating that in these particles half of the VP2 proteins have undergone cleavage to VP3, although the process may not be complete. Examination by negative stain EM showed intact particles (stain penetrated and stain excluded for empty and fulls, respectively) for both samples (Figure 3-1B &C). Thin plate shaped and thin rod shaped crystals grew in ~4-6 weeks from the 10 mM Tris-HCl pH7.5, 150 mM NaCl, 8 mM CaCl<sub>2</sub>.2H<sub>2</sub>O and 3% PEG 8000 condition (Figure 3-1D). The approximate crystal dimensions were 0.3 x 0.15 x 0.005 mm for the thin plate shaped crystals and 0.15 x 0.01 x 0.005 mm for the thin rod shaped crystals. Diffraction data was collected on the thin rod shaped crystals.

### Glycans Recognized by H-1PV

The empty H-1PV particles were screened at three different concentrations of 200 µg/ml, 400 µg/ml and 900 µg/ml on the glycan array. Depending on the concentration of H-1PV analyzed, the RFU signal and the %CV for each glycan was different and taking

into account these variables the highest affinity glycans were identified (Figure 3-2). Of the 377 glycans present on the array, H-1PV recognized only three  $\alpha$ 2-3 sialylated glycans linked to a common Gal $\beta$ 1-4GlcNAc motif: Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc $\beta$ 1-3Gal $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc $\beta$  (3'SIA-Le $\chi$ -Le $\chi$ -Le $\chi$  or 3'SIA(Le $\chi$ )<sub>3</sub>; glycan 227), Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc $\beta$  (3'SIA-LN-LN or 3'SIA(LN)<sub>2</sub>; glycan 236), and a biantennary glycan Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Neu5Ac $\alpha$ 2-6Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$  (glycan 316).

### **H-1PV Capsid Structure**

The structure of the full and empty H-1PV particles were determined to 2.7 and 3.2 Å resolution, respectively, by X-ray crystallography. The refinement values are within the range reported for other virus structures determined at a comparable resolution, as calculated by the Polygon subroutine (307) in the program PHENIX (2). The similarity of  $R_{\text{factor}}$  and  $R_{\text{free}}$  for virus structures is a result of the high noncrystallographic icosahedral symmetry of the capsid. The map quality and refinement statistics show that the refined structure is not biased by the MVMP phasing model (Figure 3-3A).

While the empty H-1PV particles were assembled from VP1 and VP2, and full H-1PV virions were assembled from VP1, VP2 and VP3, only the C-terminal 554 amino acid (aa) residues, corresponding to residues 38 to 593 (VP2 numbering) of the VP2 (and VP3 in virions) common sequence were traceable in the icosahedrally averaged electron density map (Figure 3-3B). This lack of N-terminal ordering is also reported for all other parvovirus structures determined to date (exception B19V (155)). In the full H-1PV virion structure, low signal threshold (less than  $0.5\sigma$ ) density is also observed under the icosahedral fivefold axis, but it is not connected to the first interpretable N-

terminal residue (aa 38) and was thus not modeled. In addition to N-terminal residues 1-37, two amino acids at the top of the DE loop (the loop between  $\beta$ -strands D and E, see below), which together with four five-fold symmetry related DE loops assemble the fivefold channel, were not ordered in electron density maps for both full and empty H1-PV structures.

The VP2 structure of the full and empty H-1PV particles conserved the parvovirus capsid VP topology. It contains the core eight stranded jelly roll motif, consisting of two  $\beta$ -sheets, BIDG and CHEF, common in most virus structures, with long loops between the strands, and an  $\alpha$ -helix ( $\alpha$ A; residues 128 to 138, VP2 numbering) located close to the icosahedral twofold axis (Figure 3-3B). Small regions of antiparallel  $\beta$ -strands are observed in the loops between the strands as previously described for other parvoviruses (reviewed in (53)). The  $\beta$ -barrel motif represents ~17% of the ordered VP2 amino acid sequence, and forms the contiguous shell. The loops clustered from icosahedral symmetry related VP2 monomers form the capsid surface. Two of the small stretches of strand structure, between  $\beta$ D and  $\beta$ E, form a  $\beta$ -ribbon (DE loop) which clusters with the  $\beta$ -ribbons from the fivefold symmetry related monomers to form the conserved cylindrical channel at the fivefold axes (Figure 3-3C and D). Depressions are observed at and surrounding the twofold axis and surrounding the fivefold channel (Figure 3-3D). The floor of the depression around the channel is lined by the HI loop (between strands  $\beta$ H and  $\beta$ I) that resembles a flower petal extending from the adjacent fivefold symmetry related monomer (Figure 3-3D and E). The intertwining of six large surface loops, two from each threefold-symmetry related monomer, forms the mound-like protrusion at the icosahedral threefold axes (Figure 3-3 D). These loops are

between  $\beta$ E and  $\beta$ F and between  $\beta$ G and  $\beta$ H (Figure 3-3B). The shoulder of the protrusions is formed by loops between  $\beta$ B and  $\beta$ C and between  $\beta$ G and  $\beta$ H.

In the full H-1PV structure, ordered density for nine ssDNA nucleotides was observed in the capsid interior at the twofold axis in a conserved DNA binding pocket previously reported for other members of the parvovirus genus (Figure 3-4A) (7, 306). The chain direction of nucleotides was evident from the positive difference density in the  $F_o$ - $F_c$  map for the phosphate groups, and the relative position of the sugars, phosphates and bases also aided in the model building. The good quality of the icosahedrally averaged electron density maps for the H-1PV data set enabled the clear distinction between purine and pyrimidine bases, and for some nucleotides also between A or G, and C or T. The DNA sequence 5'-TGCCTTCAA-3' was built into the well-ordered density (NTs numbered from 2-10 to correspond to the NT numbering in the CPV structure) (Figure 3-4A). The overall conformation adopted by the ordered ssDNA stretch in the H-1PV virions is similar to that found in MVMi and CPV, with the bases pointing outward to interact with the capsid protein while the phosphate-deoxyribose backbone is on the inside of the loop with the negatively charged phosphates held together by  $Mg^{2+}$  ions (Figure 3-4A and Figure 3-5) (7, 306). The 5' end of the DNA strand is close to the fivefold axis. Capsid protein-DNA interactions include non-specific (with backbone and deoxyribose sugar) and specific (nucleotide base atoms) van der Waals and polar interactions. The binding site contains several residues that are within acceptable distances (2.4 Å to 4.0 Å) for hydrogen bonding interactions, such as L146, Q148, L184, S186, N187, I189, T272, Y275, I276, D480, N497, N498, P500 (from the reference monomer), D57, T285, P545, L547, H590 (from the threefold related

monomer), and T51, Y52, K53 and F54 (from the fivefold related monomer) which are highly conserved among the autonomous parvoviruses (Table 3-3) and (Figure 3-6A). The DNA strand is stabilized by base-base stacking interactions, and there are two distinct stacks: one consisting of T6, T7, C8, and A10, and the other consisting of G3 and A9). The protein-DNA complex structure is stabilized by extended DNA-DNA base stacking as well as NT-protein stacking with aromatic amino acids (Figure 3-6A). For example, C5 base stacks with H590. Apart from the ssDNA sequence, three cytosine molecules were also modeled into the full H-1PV structure (Figure 3-7). The capsid amino acid residues interacting with the cytosines are H483, R486 (Cyt1), W59, S140, N142, P281, K540 (Cyt2), and W59, K540 (Cyt3).

The VP2 structures for the full and empty H-1PV particles structures superimposed with an overall rmsd of 0.47 Å for all 554 residues. In the empty H-1PV crystal structure, there was no density observed for the nucleotides, which is consistent with the data for other empty autonomous parvovirus structures (168, 276, 322). The full and empty H-1PV structures were almost identical except for side chain conformation differences observed at the DNA binding pocket (Figure 3-8A). In full H-1PV structure, H590 exists in two different conformations, interacting with the C5 nucleotide in one conformation. However, in the empty H-1PV structure it exists in only one conformation that faces slightly away from either of the two conformations adopted in the full H-1PV structure. The other three detectable residue level changes were at K53, G56, and D57. Four cis-peptide bonds (3 proline: A423-P424, I427-P428, Y469-P470, and one non-proline cis-peptide bond, H348-D349) were observed in the H-1PV virion structure (Figure 3-9). The cis-peptide bonds in the full H-1PV structure were identified based on

the backbone geometry and better fit into the electron density map. Residue P470 is located at the twofold axis of symmetry just above the conserved  $\alpha$ A helix . The other cis-peptide bonds (2 proline: A423-P424; I427-P428 and one non-proline cis-peptide bond: H348-D349) are at the threefold and interact with loops from the symmetry related molecules. Two of the cis-peptide bonds (423-424 and 427-428) are present in the same loop.

### **Structural Studies of H-1PV and Glycan Complexes**

Two of the glycans (3'SIA(LN)<sub>2</sub> and 3'SIA(Le<sup>X</sup>)<sub>3</sub>), recognized by empty H-1PV capsids when screened on the glycan array were co-crystallized with empty H-1PV particles, and X-ray diffraction data was collected on the co-crystals for structure determination. The structure of H-1PV-3'SIA(LN)<sub>2</sub> and H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> co-crystal structures were determined to 2.7 Å and 2.9 Å resolution, respectively. The data refinement statistics are summarized in Table 3-1. For the H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex data, positive density was observed in the F<sub>o</sub>-F<sub>c</sub> difference maps at 2.5 $\sigma$  at the depression at the icosahedral twofold axis of symmetry. This site was also identified as the SIA binding site for MVM in previous studies (188) (Figure 3-10A, B and D), and was modeled as a single molecule of SIA. Residues, L248, N249, R355, D367, I368, D405, A406, and S409 interact (2.4 - 4.0 Å) with the SIA molecules (Figure 3-10C). I368 and D405 in H-1PV are homologous to the MVMP residues, I362 and D399 which are known to be involved in tissue tropism and pathogenicity determination for MVMP (188, 259). For the modeled SIA, the N-acetyl group interacts with D367, I368, D405 and A406; the glycerol group interacts with S409, L248, and N249; and the carboxyl group interacts with R355. For the H-1PV-3'SIA(LN)<sub>2</sub> complex data set, the F<sub>o</sub>-F<sub>c</sub> difference map showed density that could be modeled as a SIA in a pocket adjacent to the

previously identified MVMP SIA binding site (Figure 3-10C) (188). This binding site on the H-1PV capsid is narrower than the MVMP glycan binding site. The modeled SIA in this site interacts with charged and hydrophobic residues, such as Q326, N327, D373, H374, T415, A423, P424 and R437. Residues N327 and H374 are homologous to MVMP residues G321 and K368, respectively, which are known tissue tropism and pathogenicity determinants for MVM (15, 112, 188, 196, 259). The glycerol chain of SIA sits in a pocket lined with D373, T415, R437 (D367, T409, R431 in MVMP), the ring structure interacts with H374, A423 and Q424 (K368, A417, P418 in MVMP), and the N-acetyl group interacts with residues Q326 and N327. There were no detectable changes observed in the capsid upon receptor attachment.

## Discussion

### Recognition of Sialylated Glycans by H-1PV

Empty H-1PV particles were screened on a glycan array containing 377 different glycan motifs and specifically recognized only three structures with terminal sialic acid linked α2-3 to a common Galβ1-4GlcNAc motif; Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ (3'SIA-Le<sup>X</sup>-Le<sup>X</sup>-Le<sup>X</sup> or 3'SIA(Le<sup>X</sup>)<sub>3</sub>; glycan 227), Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ (3'SIA-LN-LN or 3'SIA(LN)<sub>2</sub>; glycan 236), and a biantennary glycan Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ (glycan 316) (Figure 3-2). Complex N-glycans are reported to have two types of LacNAc (Gal-GlcNAc) repeats; Galβ1-3GlcNAc (Type1) and Galβ1-4GlcNAc (Type2) of which Type2 repeats are more common. Specifically, H-1PV did not recognize any sialylated glycans linked to Type1 LacNAc motif, and only bound to glycans with the common motif Neu5Acα2-3Galβ1-4GlcNAc (3'SIA-LN), similar to observations with

MVM (209). This suggests that the type of Gal-GlcNAc linkage plays a role in the receptor-capsid interaction for H-1PV and MVM, and preliminary modeling studies with Neu5Aca2-3Gal $\beta$ 1-3GlcNAc (Type1) in the SIA binding site of H-1PV and MVM indicates that the kink associated with the Gal $\beta$ 1-3GlcNAc linkage creates steric hindrance in the binding site. H-1PV bound with high specificity to 3'SIA(Le $^X$ )<sub>3</sub> (3'SIA-Le $^X$ -Le $^X$ -Le $^X$ ; glycan 227) which was also recognized by MVM in a previous glycan array screening (209) and is a known tumor cell marker. Another oncotropic rodent parvovirus, Lulll also showed recognition for the 3'SIA(Le $^X$ )<sub>3</sub> in glycan array screening (unpublished data). This suggests a common theme of utilization of the 3'SIA-Le $^X$  motif by the oncotropic rodent parvoviruses, MVM, H-1PV, and Lulll to bind and infect transformed cells. The LN-LN linkage (GlcNAc $\beta$ 1-3Gal) in glycan 236 that was recognized by H-1PV is different from that in the 3'SIA-LN-LN glycan (Neu5Aca2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-4Gal $\beta$ 1-4GlcNAc) recognized by MVM in previous glycan array studies, but is unfortunately not present in the current array (209). The biantennary glycan (glycan 316) has two terminal sialic acids; the Neu5Aca2-3 is on the 3-branch of the trimannosyl core (Man $\alpha$ 1-3) and the Neu5Aca2-6 is on the 6-branch of the trimannosyl core (Man $\alpha$ 1-6). Switching the positions of Neu5Aca2-3 and Neu5Aca2-6, and also when both the sialic acids are  $\alpha$ 2-3 or both  $\alpha$ 2-6, binding to H-1PV is abolished. This suggests that in the context of a biantennary glycan, H-1PV recognizes  $\alpha$ 2-3 SIA only on the 3-branch and  $\alpha$ 2-6 SIA only on the 6-branch, while it won't bind to a single chain glycan with  $\alpha$ 2-6 linked SIA. However, MVM viruses screened on the SGM (data reported in chapter 2) identified a biantennary glycan with terminal SIA linked only  $\alpha$ 2-3 to both the 3-branch and 6-branch. H-1PV and MVM have some

common properties, including their recognition of similar glycans, but there are some differences in their properties which may be capsid-mediated.

### **Structure of H-1PV**

The H-1PV VP2 structure has the conserved parvovirus capsid viral protein topology, consisting of a core eight stranded anti-parallel β-barrel motif with large loop insertions between the strands (53) (Figure 3-3B). The VP1u and the VP2 N-terminal residues (1-37) were not observed in the crystal structure. Low copy numbers of VP1 and VP2 in the mature virions or different conformations adopted by the VP1/2/3 N-termini, which is incompatible with the 60-fold icosahedral averaging applied during structural determination, are postulated to result in the lack of N-terminal VP ordering (53). In the full virions of other autonomous parvoviruses, such as MVM, CPV and also for the dependovirus AAV8, low sigma threshold electron density was observed inside the capsid, under the twofold axis, running between fivefold related VP monomers, and also within the fivefold channel (4, 7, 210, 306). In MVM and CPV, this density was modeled as the conserved glycine-rich N-terminal region of one VP2 subunit spanning residues 28 to 38, to represent the externalized VP2. In MVM, additional density corresponding to residues 36 to 38 from the remaining four fivefold related VP was modeled as extending back into the particle interior at the base of the fivefold cylinder (7). The lack of ordering of the residues that form the tip of the DE loop in full and empty H-1PV particles might be due to the reported dynamic nature of the fivefold cylinder (168), as is required for its reported function as a portal for VP2 and VP1u externalization and genome encapsidation (75, 100-102). The structural transitions that occur during these processes might not be compatible with icosahedral averaging carried out during structure determination process.

## **Structural Comparison to Other Autonomous Parvoviruses**

A structural superposition of the ordered VP2 region of H-1PV with the analogous VP2 structures of other autonomous parvoviruses MVM, CPV, FPV and PPV shows that the  $\beta$ -barrel core and the  $\alpha$ A are highly conserved and they have very similar loop topologies (Figure 3-11). While the VP2 sequence identity between these parvovirus members ranges from 52% to 68%, the structural homology ranges from 75% to 90% (Table 3-2) with an overall rmsd between the C $\alpha$  positions of H-1PV VP2 and these viruses ranging from 0.63 Å - 0.94 Å (Table 3-2). The high structural homology is due to the high sequence identity for the  $\beta$ -barrel core, while the surface loops are the most varied in sequence and show structural variations in between the members. In Kontou *et al.* (168), eight surface loop regions that showed structure variability (C $\alpha$  differences of as much as 5 Å) between these viruses (except for H-1PV) were identified and numbered 1 to 8. These regions also differ between the highly homologous parvovirus strains, such as MVMi and MVMP, CPV and FPV, and CPV and its host range mutants. These variable loop regions are clustered on the assembled 3D capsid at and around the icosahedral twofold axis, on the shoulder of the threefold protrusion, and at the fivefold axis. Prominent differences between H-1PV and the other parvovirus structures are also observed at the threefold protrusions, twofold depression and at the fivefold axis (Figure 3-11A and B). The majority of these differences are located in the previously defined variable regions, including VR1 (152-176), VR2 (231-253), VR3 (297-311), VR4a (314-340), VR4b (421-433), VR5 (361-380), VR6 (388-398), VR7 (508-525), and VR8 (553-572). In addition, H-1PV differs from MVM in a surface loop consisting of residues 88-100 located at the tip of the threefold protrusion due to a four amino acid insertion in H-1PV. This VR was not identified in the previous study by Kontou *et al.*

(168) and is labeled as VRX. Variable region 2, VR4a and VR4b contribute to the threefold protrusions, and have been mapped as antigenic sites on the MVM and CPV capsids (125, 157, 321). Differences in VR8 (between H-1PV and MVM), and VR5 (for all parvoviruses compared), contribute to capsid surface variation at the floor of the twofold depression, and at the wall of the twofold depression, respectively. These capsid regions serve as the SIA receptor binding sites in MVM and CPV, in addition to being involved in tissue tropism and pathogenicity determination for these and other parvoviruses (5, 19, 188). In a recent study (226) it was shown that the rodent parvoviruses H-1PV, LullI, and MVM exhibit variation in tumor tropism and oncolytic activity of LullI was mapped to the capsid VP2 gene. These observations suggest that the VP2 structural variations observed between MVM and H-1PV, might dictate the differences in tumor tropism. In an effort to correlate the tropism of H-1PV with that of the MVMP-F1 hot switch mutants that adapted to infect rat fibroblasts, the conformation of the residues in H-1PV that are homologous to the residues mutated in the MVMP-F1 virus (334, 384, 554, 578 in MVMP correspond to 340, 390, 560 and 584 in H-1PV) were analyzed. Out of the four host range switch mutations *in vitro* in MVMP that conferred the ability to grow in rat fibroblasts (similar to H-1PV's cell tropism), only two of those (residues 560 and 584 in H-1PV) are structurally variable between MVMP and H-1PV and are located at the twofold depression and might play a role in the virus adaptation to a new host. A comparison of the available dependovirus capsid structures also identified similar variable regions that are involved in various functions such as transduction, receptor recognition and antibody binding (119). Comparative analysis of the parvovirus structures provides information on the capsid interactions at the

intermonomer interfaces that are required for efficient assembly, and interdigitation of the loops which enables identification of capsid regions that can tolerate loop insertions and deletions. Recent developments in gene therapy vector design include engineering of mosaic/chimeric vectors by shuffling capsid genes from related viruses in an effort to design vectors with specific tissue tropism and reduced immunogenicity. Structure guided vector design identifies the capsid regions of parvoviruses that are not conducive to peptide insertions because they would either create steric hindrance at the symmetry axes due to mismatched loop sizes or lack residues involved in stabilizing inter-monomer interfaces.

### **DNA-Binding Pocket**

The ordered density for ssDNA in the full H-1PV structure was observed in the capsid interior in a cavity present at the interface of three VP monomers (reference, threefold related and fivefold related monomer). Ordered density for ssDNA has also been observed in an analogous region of the capsid structures of other autonomous parvoviruses such as CPV and MVM, thus providing additional proof of the importance and existence of this conserved DNA binding pocket (Figure 3-5) (7, 306). However, as mentioned in previous reports, the ordering of the DNA density is unusual because the structure determination procedure assumes icosahedral symmetry (or 60 equivalent positions). Since the capsid packages only one copy of the genome and it has been seen that the genome sequence observed in such viruses has a low degree of repetition within the genome, it is not possible that the exactly same sequence is present in 60 equivalent sites in the capsid (329). Electron density for a single nucleotide has been observed in the capsid structures of AAV4, AAV6 and AAV8 at the icosahedral threefold axis (119, 177, 210, 214). While in the autonomous parvoviruses, ~12 to 25% of the

genome is ordered, in the dependoviruses only ~1% is ordered. This is suggested to be due to the packaging of predominantly the negative sense strand (or a single polarity) of the genome by the autonomous parvoviruses (except for Lulli) and packaging of both sense strands by the dependoviruses inside different capsids. Therefore, the AAV crystals would contain different populations of capsids, which would result in the heterogeneity of DNA sequence in the crystallized virus, thus decreasing the probability of visualizing a longer DNA sequence in the crystal structure after icosahedral and crystal averaging applied during structure determination. The amino acid residues in the nucleotide binding pocket of the parvovirus genus members and dependoviruses differs, so specific capsid-DNA interactions could also dictate the amount of ordered DNA. Structure determination of Lulli, a member of the parvovirus genus that packages either sense strand in different particles, would aid in determining if the structural ordering of ssDNA is dictated by genome sequence and packaging constraints and/or a conserved DNA binding pocket in the capsid. Similar to the autonomous parvoviruses, icosahedrally ordered ssDNA was also observed in the *Bacteriophage φX174* (197). Icosahedrally ordered RNA has been found in Flock House Virus where non-specific interactions were observed between the sugar-phosphate backbone and the capsid protein (104), and in bacteriophage MS2 where the capsid protein is involved in extensive contacts with nucleotide bases (308).

The nucleotides in H-1PV capsid are involved in specific interactions with capsid surface amino acid residues that are highly conserved among the autonomous parvoviruses (Table 3-3). H590 exists in a dual conformation in full H-1PV particles, and in one conformation base stacks with C5. T6, T7, C8 and A10 form a stacking ladder, so

for C5 to flip out of the stabilizing stacking ladder to interact with H590, indicates the high specificity of this interaction and an important role of H590 in genome packaging (Figure 3-6A). One of the modeled cytosines (labeled Cyt1 in Figure 3-7B) is located in the conserved AAV nucleotide binding pocket under the icosahedral threefold axis and interacts with the conserved residue, R486 (Table 3-4). Although the DNA binding pocket in parvovirus genus members is highly conserved, comparison of the ssDNA sequence observed in the autonomous parvoviruses, H-1PV, MVMi and CPV, shows that there are less nucleotides (NTs) ordered in the H-1PV structure (9 NTs; Chain C) compared to the MVMi (21 NTs; Chain B and C) and CPV (11 NTs; Chain B) structures (Figure 3-4) (7, 306). There were also differences in the ssDNA sequence observed in H-1PV, MVM and CPV structures (Table 3-5). T2 (base followed by nucleotide number) interacts with highly conserved residues (L146, Q148, p500, T272, Y275), and Y275 interacts with the methyl group, so T is observed in this site for H-1PV, CPV and MVMi. Y275 and N497 interact with the amino group of G3 (guanine 3) and I276 interacts with the carboxyl group, so only a G or C can be accommodated at this site. Since I276 differs between these viruses at this site, it allows the accommodation of A in CPV and C in MVM. C4 is present in the three viruses as it interacts only with the highly conserved N497. At nucleotide position 5, only a pyrimidine is allowed because the amino acid residues (T285, P545, I547, H590) interact with the carboxyl group and the carbon at position 2 and 4 of the pyrimidine ring. Also, since the residues at this site differ between the three viruses, either a C or T could be accommodated. P545 interacts with methyl group, T51 with the carboxyl group, Y52 with the amine and F54 with the carboxyl group at position 4 of T6. Due to the recognition of methyl group by

P545, only a T can fit into the H-1PV pocket at this site. For the other viruses, this site contains variable residues. K53 interacts with the carboxyl group at position 2 of T7, so only a C or T can be accommodated at this site (C in CPV). D57 and K53 interact only with the amino group in C8, so a lot of variation is allowed. D57 interacts with the amino group of A10, so even at this site either A, G or C can be accommodated. An analysis of the nucleotides and binding pocket of these capsids that are ~50% identical at the nucleotide and amino acid level, shows that the common interactions allow for DNA packaging and stabilization while the differences between the structures dictate the packaging of a specific genome sequence. It has been demonstrated using a MVM and Lulli chimeric virus that secondary structure elements, including stem-loops and guanidine rich regions, can interrupt packaging (72, 73). Further analysis of the DNA binding pocket reveals that the amino acids interacting with the DNA are associated with the antiparallel  $\beta$  strands, similar to that observed for other non-viral proteins which contain oligonucleotide binding motif (OB fold) (298).

The conservation of the DNA binding pocket suggests that this DNA-protein interaction may play an important role in the virus life cycle, such as in capsid assembly or stabilization. An analysis of the mechanical properties of the empty and full MVM capsids suggested that the bound genome reinforces the stiffness of the particle, especially at the twofold symmetry axis (49). In the study by Reguera *et al.* (247), alanine mutants of three of the ssDNA interacting residues in MVM (Y270, D273, D474 corresponding to F266, D269, D475 in CPV, and Y275, N278 and D480 in H-1PV, respectively) displayed drastic reduction in infectivity, with the Y270A defective in capsid assembly and D273A unable to encapsidate genome, while there were other mutants

(D58, W60, N183, T267, or K471) that were less stable. These results further propose the role of capsid-DNA non-covalent interactions in maintaining capsid stability. The same study showed that at least one acidic residue at each DNA-binding pocket is required for genome packaging.

### **Conformational Variation Between Full and Empty Capsids**

To examine possible conformational modifications imposed on the capsid proteins upon ssDNA binding, the full and empty H-1PV capsid structures were superimposed for comparison. The structures superimposed within overall rmsd of 0.47 Å for all 554 residues. The structures were almost identical except for side chain conformation differences observed at the DNA binding pocket at H590, K53, G56 and D57 (Figure 3-8A). Comparisons of full and empty structures of CPV and MVM show variation at the top of the βDE loop that forms the opening of the fivefold channel, wherein the loop rearrangement increases the diameter of the opening in the empty capsid structure (Figure 3-8B and C) (7, 168, 322). Due to the lack of ordered density for this loop in both the empty and full capsids of H-1PV, no conclusions can be made. Large main chain and side chain displacements were observed at the ssDNA binding site in full and empty CPV particles, compared to small displacements in the full and empty MVM structures. No ordered density for DNA was observed in the empty H-1PV capsids which is consistent with the structures of empty particles of other members of this genus (168, 276, 322). However, AAV VLPs package small amounts of cellular DNA and in crystal structure of AAV VLP, ordered density for a single nucleotide is observed, which suggests that in the case of the dependoviruses, capsid-DNA interactions might play a role in capsid assembly (168, 276, 322).

## Cis-Peptide Bonds

An analysis of the protein structures in the PDB database conducted in Weiss *et al.* (318) and Jabs *et al.* (140) found only 0.03% of all X-Xnp, and 5.2% of all X-Pro peptide bonds to occur in the cis conformation (where X=any residue, Xnp=any non-Pro residue). A Pro-Pro bond has the highest frequency to be in the cis form, followed by Tyr-Pro. The cis-peptide bonds, especially non-Pro cis-peptides are rare due to the steric strain but when present they are usually located near functional sites and have been suggested to be important to a protein's function and may serve as a kind of energy reservoir that drives conformational changes (40). As an example, an Ala-Phe cis-peptide bond at the binding site for the antiviral assembly inhibitor drug CAP-1 in the HIV capsid protein was suggested to facilitate capsid assembly (159).

Residue P470 was also shown to exist in the cis conformation in the CPV capsid structure (322). It is conserved among the members of the parvovirus genus, but was not refined as a cis-peptide in any of the other structures and is located at the twofold axis of symmetry just above the conserved  $\alpha$ A helix (Figure 3-9). The other cis-peptide bonds (2 proline: A423-P424; I427-P428 and one non-proline cis-peptide bond: H348-D349) are at the threefold and interact with loops from the symmetry related molecules (Figure 3-9). Two of the cis-peptide bonds (423-424 and 427-428) are present in the same loop. The loops that are involved in the interactions at the threefold axis of symmetry are very closely interdigitated and the presence of the cis-peptide bonds could impart flexibility to the loop. As suggested in Riolobos *et al.* (248), if the MVM capsid assembles via trimeric intermediates then these cis-peptides could play a major role in folding/unfolding of the loops to allow efficient capsid assembly.

## Glycan Binding Site on H-1PV Capsid

Preliminary structural studies of the empty H-1PV capsids complexed with 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan showed that the SIA receptor binding site is the same as identified for MVMP in previous studies and in the complex structures determined in chapter 2 (Figure 3-10A) (188). Only the SIA component of the glycan was observed in the crystal structure and modeled into the F<sub>o</sub>-F<sub>c</sub> difference density map at 2.5σ. The carbohydrate moieties at the reducing end may not be forming tight interactions with the capsid surface, or do not bind to all the sites with an equal occupancy, or they are very flexible and dynamic, or adopt different conformations. These possible scenarios would be inconsistent with the icosahedral symmetry imposed during structure determination, and would lead to lack of ordering of the density for these molecules. Although structurally homologous residues (D405, I368 in H-1PV correspond to D399 and I362 in MVMP, respectively) were involved in the SIA recognition, the surface of the pocket and the side chain conformation of the interacting residues is different from that observed in MVMP. The pocket in H-1PV is wider and shallower in comparison to MVMP (Figure 3-10). Comparative analysis of the structures of 3'SIA(Le<sup>X</sup>)<sub>3</sub> complexed to MVM (data presented in chapter 2) and H-1PV, shows that although interaction with I368 and D405 (I362 and D399 in MVMP) is conserved, the glycan binds deeper into the H-1PV pocket and interacts with residues, D367, R355, A406, L248, N249 and S409 (correspond to residues D361, R349, E400, M243, N244 and F403 in MVMP) that it does not interact with on the surface of MVMP or MVMi capsid. It has been shown that H-1PV and MVM vary in their oncotropic properties, and this could be regulated by the differences in their binding interactions with 3'SIA(Le<sup>X</sup>)<sub>3</sub>, the tumor cell marker (226). MVMP-K368R and MVMP-I362S virulent mutants have been shown to have a reduced affinity for the sialic

acid receptor, and a recent study reported that analogous substitutions introduced into the H-1PV capsid (H374R and I368S) resulted in an even more drastic reduction in cell binding as compared to the MVMP mutants (10, 188). This suggests that SIA is an important determinant of H-1PV and MVM's infectivity, and common residues regulate their SIA binding, but there are differences in receptor affinity (demonstrated by differences in sensitivity to neuraminidase treatment) that results in differences in their *in vivo* tropism. Also, interactions with other amino acid residues in the pocket may affect the receptor binding affinity. For the H-1PV-3'SIA(LN)<sub>2</sub> complex, the shape and composition of the SIA binding site was different from that identified for MVMP or the H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex (Figure 3-10D). This binding site is narrower and occluded. In this site, the glycan interacts with residues homologous to MVMP, such as H374 and Q326 (corresponds to residues K368 and G321 in MVM), that are known to dictate tissue tropism and pathogenicity in MVMP. The refinement of the H-1PV co-crystal structures needs to be completed to be able to fully define the role of the specific amino acids in dictating the mechanism of cell recognition.

## **Summary**

Glycan array screening of empty H-1PV particles showed that it recognized only α2-3 sialylated glycans linked to a common Type2 LacNAc (Galβ1-4GlcNAc) motif, similar to observations with MVM (209). The recognition of 3'SIA-Le<sup>X</sup>, a tumor cell marker, indicates that H-1PV like MVM also utilizes 3'SIA-Le<sup>X</sup> motif on cancer cells to recognize and bind to these cells. The overall topology of VP of H-1PV contains a highly conserved eight-stranded anti-parallel β-barrel core with large loop insertions between the strands that forms the capsid surface, as is previously reported for other parvovirus capsid structures. The H-1PV full and empty capsid structures are identical except for

slight conformational variations in side-chain amino acids at a conserved nucleotide binding pocket previously reported in other autonomous parvoviruses. Electron density interpretable as nine nucleotides was observed at the icosahedral twofold axis in a conserved DNA binding pocket previously reported for other members of the parvovirus genus. The capsid surface amino acid residues are involved in specific interactions with the nucleotide. Four cis-peptide bonds (3 proline and one non-proline cis-peptide bonds) were observed in the H-1PV virion structure that may play a role in capsid assembly. A comparison to other autonomous parvoviruses identified the most significant structural differences on the capsid surface at the loop regions surrounding the icosahedral two-, three-, and fivefold axes at residues reported to control receptor binding, tissue tropism, pathogenicity, DNA packaging, and antibody recognition. For the H-1PV-3'SIA(Le<sup>X</sup>)<sub>3</sub> complex, SIA bound at the same site as SIA binding site on MVMp, however, there were differences in the shape and composition of the pocket which might dictate the differences in the oncotropism between these two viruses. For the H-1PV-3'SIA(LN)<sub>2</sub> complex, SIA was modeled into a pocket adjacent to the previously identified MVMp SIA binding site. The SIA binding on H-1PV capsid involves interactions with residues I368, H374, Q326 (corresponds to residues I362, K368 and G321 in MVM), that are known tissue tropism and pathogenicity determinants for MVM.

Table 3-1. Data processing and refinement statistics

Parameter	Empty H-1PV	Full H-1PV	H-1PV- 3'SIA(Le <sup>X</sup> ) <sub>3</sub>	H-1PV- 3'SIA(LN) <sub>2</sub>
Wavelength ( $\lambda$ , Å)	1.0895	0.9186	0.918	0.917
Space Group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Unit cell parameters (Å, °)	a=255.2, b=350.1, c=272.1, $\beta$ =90.12	a=255.4, b=350.4, c=271.6, $\beta$ =90.34	a=258.0, b=348.0, c=272.1, $\beta$ =90.088	a=258.4, b=347.9, c=272.4, $\beta$ =90.37
Resolution (Å)	40-3.2 (3.3-3.2) <sup>a</sup>	50-2.7 (2.8-2.7)	40-2.9 (2.9-2.8)	50-2.7 (2.8-2.7)
No. of unique reflections	479,478 (45,513)	1,210,268 (123,000)	725,031 (63004)	993,289 (102,849)
Completeness (%)	61.3 (58.3)	93.1 (94.7)	68.7 (59.8)	75.8 (78.7)
I/ $\sigma$	4.9 (1.3)	6.9 (2.4)	6.5	7.2 (2.9)
R <sub>sym</sub> <sup>b</sup> (%)	11.4 (40.0)	13.2 (37.8)	12.1	12.8 (33.9)
No. of atoms (protein/H <sub>2</sub> O/DNA/ion/other solvent)	4371/69/0/0/0	4375/98/205/4/8	4371/0/0/0/21	4360/0/0/2/21
Average Bfactor (Å <sup>2</sup> )(protein/H <sub>2</sub> O/DNA/ion/other solvent)	62/57/-/-	30/30/73/46/44	38.9/-/-/-	27.6/-/-/30/-
R <sub>factor</sub> <sup>c</sup> /R <sub>free</sub> <sup>d</sup> (%)	25.0/25.4	21.5/21.6	27.5/27.7	20.1/20.2
Rmsd bonds (Å) and angles (°)	0.007, 1.35	0.006, 1.36	0.006, 1.38	0.006, 1.38
Ramachandran Statistics (%)	91.8/	96.5	94.7	96.5/
Most favoured/Allowed/Outliers	8	3.5	2	3.5

<sup>a</sup>Values in parenthesis are for highest resolution shell.<sup>b</sup>R<sub>sym</sub>=( $\sum|I - \langle I \rangle| / \sum \langle I \rangle$ ) $\times 100$ , where I is the intensity of a reflection with indices h, k, l and  $\langle I \rangle$  is the average intensity of all symmetry equivalent measurements of that reflection.<sup>c</sup>R<sub>factor</sub>=( $\sum |F_o - |F_c|| / \sum |F_o|$ ) $\times 100$ , where F<sub>o</sub> and F<sub>c</sub> are the observed and calculated structure factor amplitudes, respectively.<sup>d</sup>R<sub>free</sub> is calculated the same as R<sub>factor</sub>, except it uses 5% of reflection data omitted from refinement.

Table 3-2. RMSD in C alpha positions for the parvovirus structures

Autonomous Parvovirus	% Identity	Maximum RMSD in Ca position between H-1PV and the other autonomous parvoviruses for VR regions(Å)										
		Overall	VRX	VR1	VR2	VR3	VR4a	VR4b	VR5	VR6	VR7	VR8
MVMp	68 <sup>a</sup> (89) <sup>b</sup>	0.63	1.7 +gap	1.5 +gap	3.3 +gap	1.7	1.2 +gap	1.8	1.5	1.7	2.2	1.6 +gap
MVMi	68 <sup>a</sup> (78) <sup>b</sup>	0.68	2.2 +gap	2.5 +gap	3.9 +gap	2.8	1.4 +gap	1.6	1.0	1.2	2.4	2.6 +gap
CPV	53 <sup>a</sup> (89) <sup>b</sup>	0.94	1.2 +gap	2.2 +gap	2.7 +gap	2.9 +gap	2.7 +gap	1.7	3.6 +gap	2.2	2.8 +gap	3.3 +gap
FPV	52 <sup>a</sup> (81) <sup>b</sup>	0.87	0.9 +gap	4.3	2.8 +gap	3.0 +gap	2.6 +gap	2.2	3.7 +gap	2.2 +gap	1.9 +gap	3.5 +gap
PPV	53 <sup>a</sup> (74) <sup>b</sup>	0.94	2.0 +gap	2.3 +gap	3.3 +gap	3.5 +gap	1.8 +gap	1.7	1.9	3.9 +gap	3.9 +gap	3.2 +gap

<sup>a</sup>Values are for VP2 sequence identity.

<sup>b</sup>Values in parenthesis are for VP2 structural identity.

Table 3-3. Amino acid sequence comparison for autonomous parvoviruses at nucleotide binding pocket

Virus	51	52	53	54	57	146	148	184	186	187	189	272	275	276	285	480	497	498	500	545	547	590
H-1PV	T	Y	K	F	D	L	Q	L	S	N	I	T	Y	I	T	D	N	N	P	P	L	H
MVM	H	Y	R	F	D	L	Q	V	S	N	I	T	Y	Y	T	D	N	N	P	A	T	R
CPV	E	F	K	F	N	F	Q	L	S	N	T	T	F	F	T	D	N	N	P	A	H	R
FPV	E	F	K	F	N	F	Q	L	S	N	T	T	F	F	T	D	N	N	P	A	H	R
PPV	E	F	Q	Y	E	F	Q	L	T	N	T	T	Y	H	T	D	N	N	P	S	N	R

\*Identical residues are shown in red, similar in green and residues that are different in black.

Table 3-4. Amino acid sequence comparison in cytosine binding pocket

Virus	59	140	142	281	540	483	486
H-1PV	W	S	N	P	R	H	R
MVM	W	Q	N	P	R	H	R
CPV	W	E	H	P	K	L	R
FPV	W	E	H	P	K	L	R
PPV	L	E	N	S	T	L	R
AAV2						G	H
AAV4						G	H
AAV8						G	H

\*Similar residues are shown in green and residues that differ between these viruses are shown in black.

Table 3-5. The ssDNA sequence present in H-1PV, CPV and MVMi

Parvovirus	Nucleotide sequence
H-1PV (9 ntds)	TGCCTTCAA
CPV (11 ntds)	ATACCTCTTGC
MVMi (21 ntds)	ATCCTCTATCAC - chain B ACACCAAAA - chain C

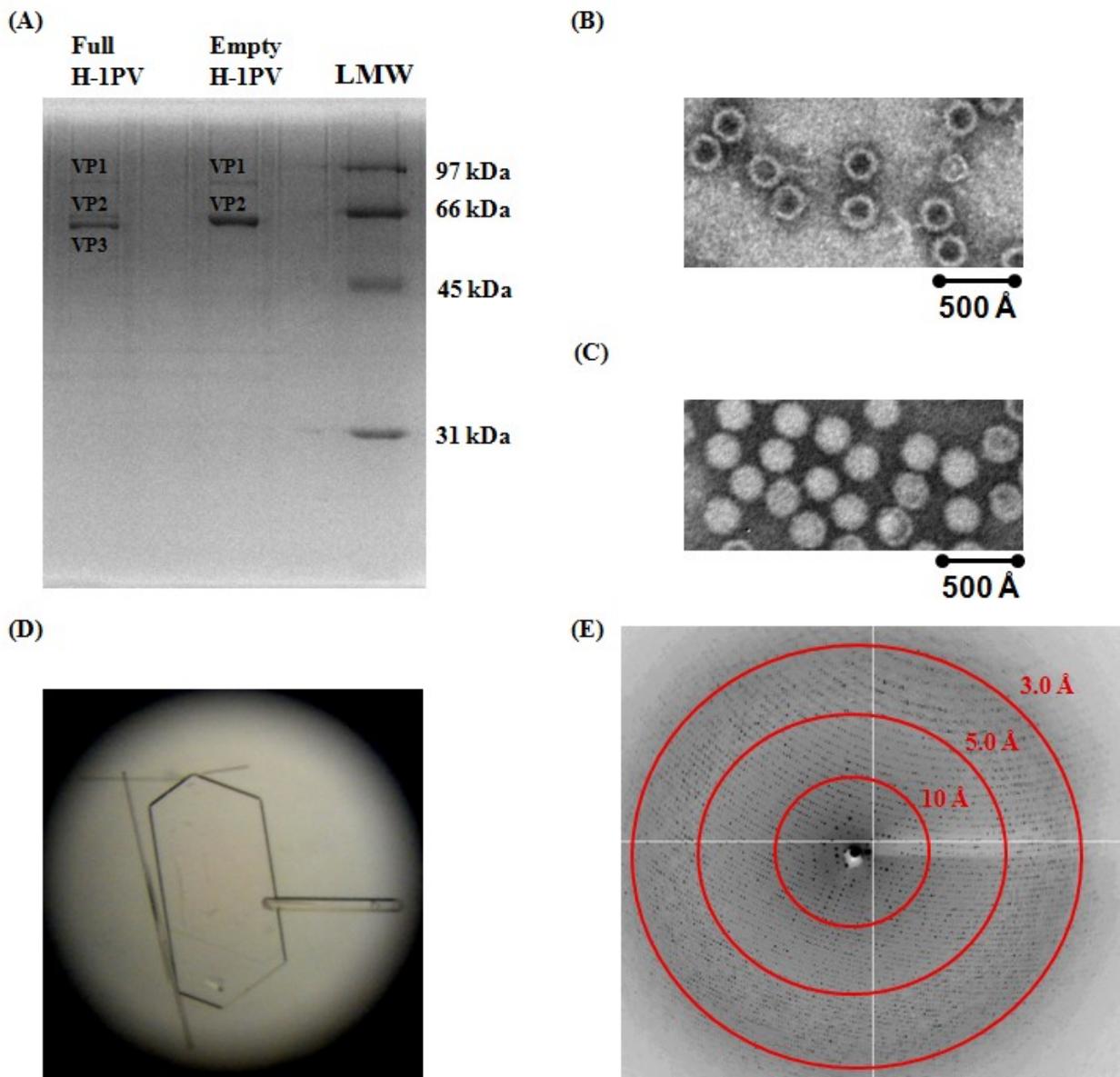


Figure 3-1. Virus purification, crystallization and diffraction. (A) An SDS-PAGE of the purified empty H-1PV and full H-1PV particles. Electron microscopy images of (B) empty H-1PV viewed at 60,000X magnification and (C) full H-1PV viewed at 100,000X magnification. (D) Optical photograph of H-1PV crystals showing the rod-shaped and plate-shaped crystals at 4X magnification. (E) A diffraction image of the full H-1PV crystal. The concentric rings indicate the 10.0 Å, 5.0 Å and 3.0 Å resolution shells.

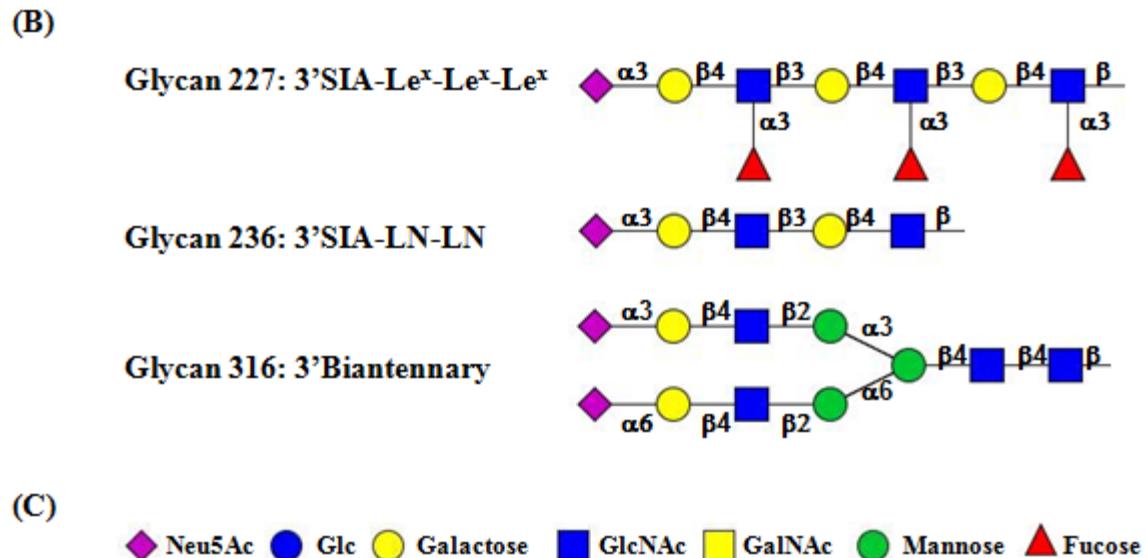
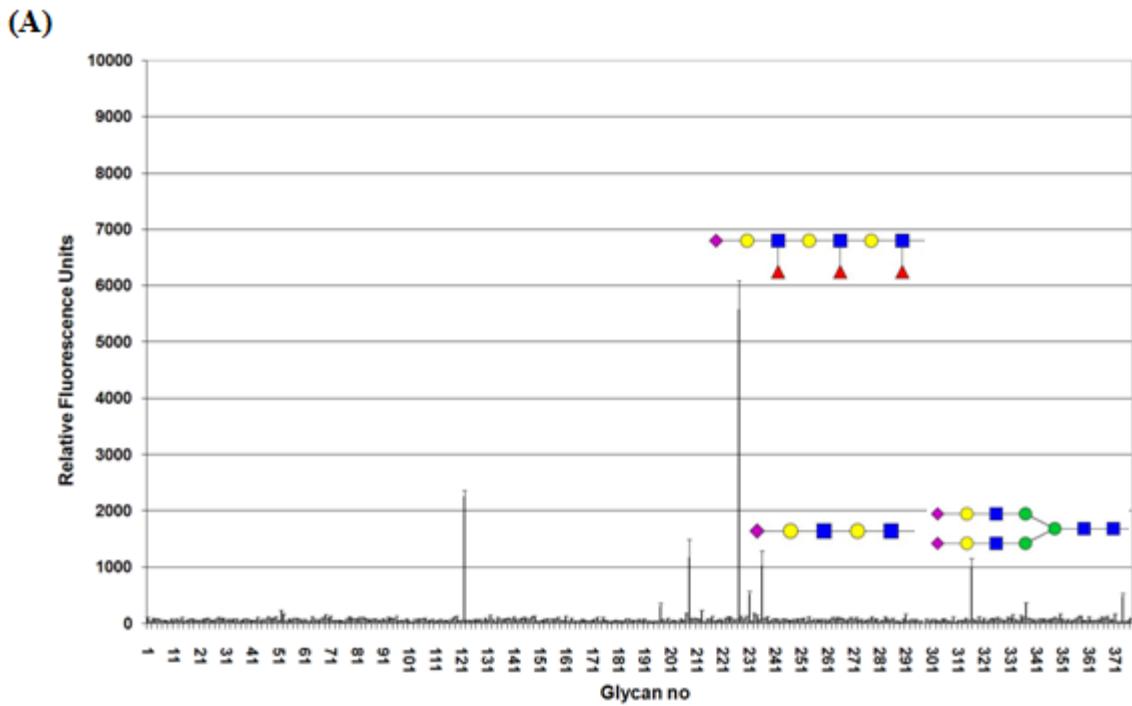


Figure 3-2. Glycan array data for H-1PV. (A) The glycan array data for H-1PV is shown. The bars represent relative fluorescence for a given glycan. The y-axis represents relative fluorescence units (RFU) and the x-axis represents the glycan numbers on the array (1-377). The glycans showing specificity to H-1PV are annotated in the cartoon form. (B) Cartoon representation of the glycans recognized by H-1PV. (C) Glycan symbols.

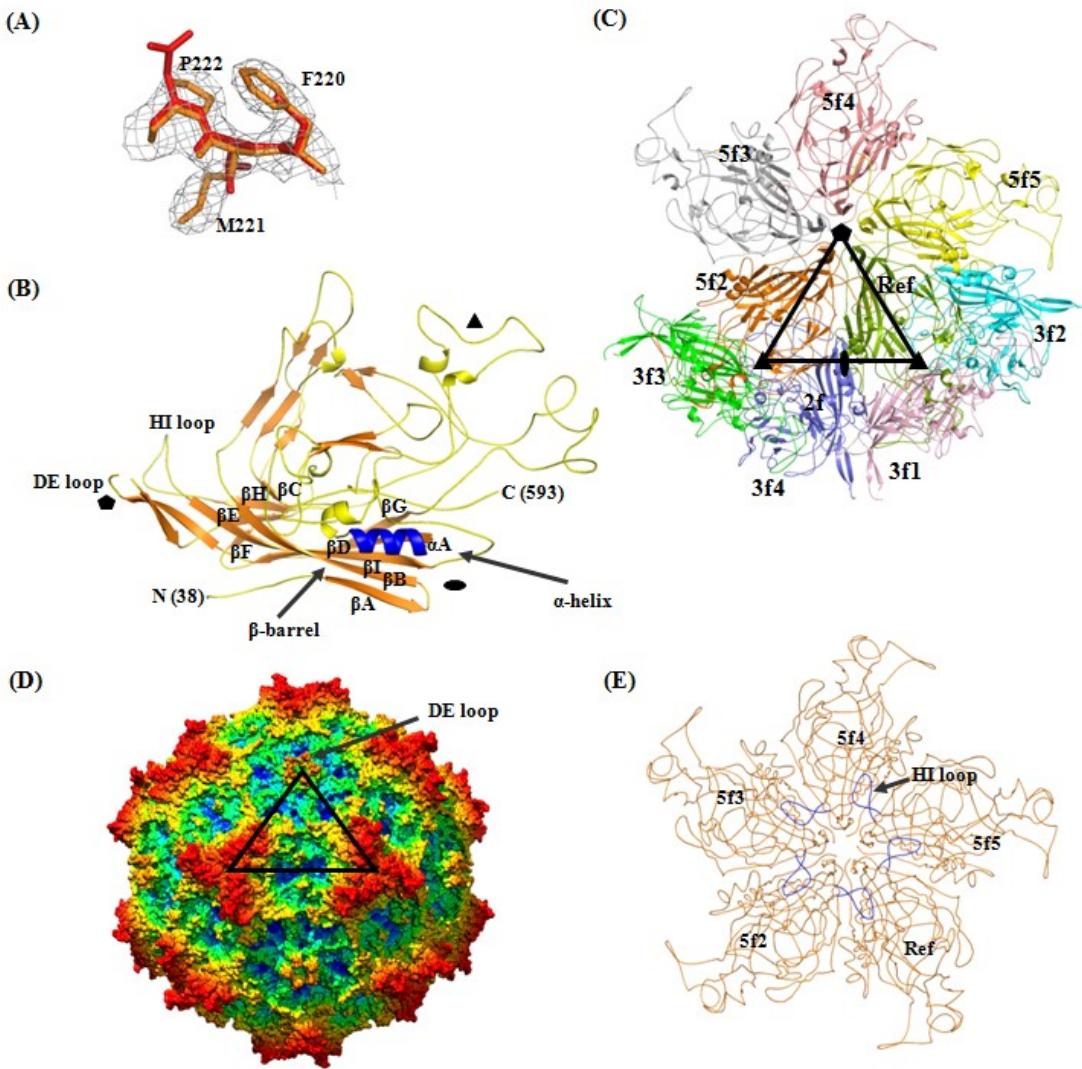


Figure 3-3. Structure of H-1PV. (A) Section of the  $2F_o-F_c$  electron density map contoured at  $1.5\sigma$  (gray mesh) of H-1PV virions (orange) for residues F220 to P222 that correspond to the MVMP (red) sequence C216 to D218. The H-1PV and MVMP coordinates are shown in stick form. (B) Ribbon diagram representation of H-1PV VP2 monomer highlighting the conserved  $\beta$ -barrel core motif ( $\beta$ BIDG- $\beta$ CHEF, orange) and the  $\alpha$ A helix (blue). Loop regions (yellow) containing small stretches of  $\beta$ -strand structure (orange), the first N-terminal residue observed (38), the C-terminal residue (593), DE and HI loops are labeled. Approximate positions of icosahedral two-, three-, and fivefold symmetry axes are depicted as filled ovals, triangles, and pentagons, respectively. (C) H-1PV VP2 related to the reference (Ref) monomer by fivefold (5f2-5f5), threefold (3f1-3f2) and twofold (2f) symmetry relationships. The black triangle depicts a viral asymmetric unit. (D) Depth-cued surface representation of the H-1PV capsid showing the topological features. (E) Ribbon diagram showing the interaction between fivefold related monomers. The HI loop is shown in blue.

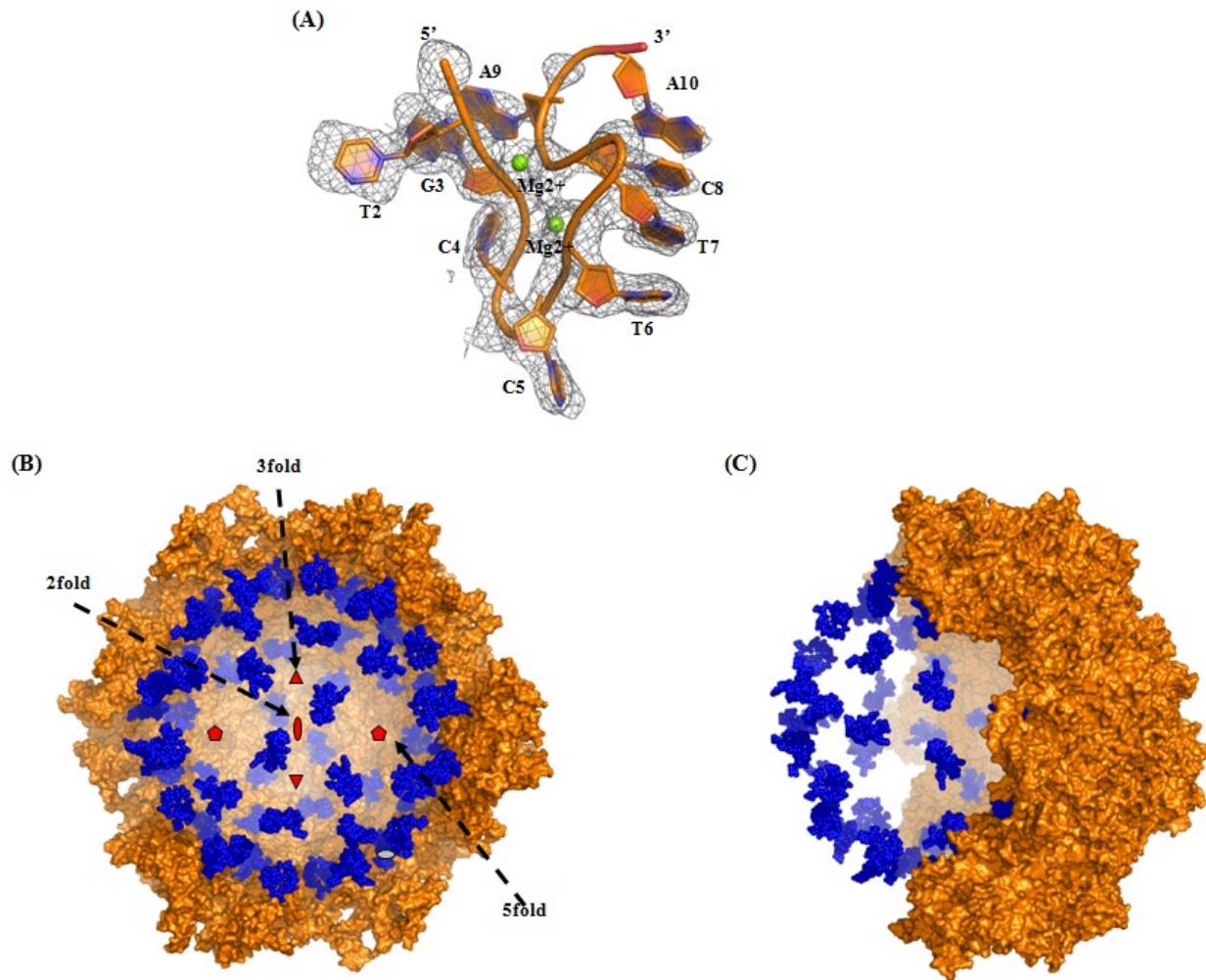


Figure 3-4. ssDNA observed in H-1PV virions. (A) The ssDNA sequence interpreted as a stretch of 9 nucleotides is shown inside the  $2F_o - F_c$  electron density map (gray mesh) and contoured at  $1.2 \sigma$ . The sugars, bases, and the phosphodiester backbone is colored in orange. The 5' and 3' ends and the nucleotides are labeled. The  $Mg^{2+}$  ions are shown as green spheres and labeled (B) A cross-section of the interior surface of full H-1PV capsid is shown in surface representation (orange) viewed down the twofold axis. The ssDNA cage that binds to the cavities at the twofold axis is shown in blue spheres (C) A  $180^\circ$  side view of the image in (B) showing the ssDNA cage. Approximate positions of icosahedral two-, three-, and fivefold symmetry axes are depicted as red filled ovals, triangles, and pentagons, respectively. These figures were generated with PyMol program (82).

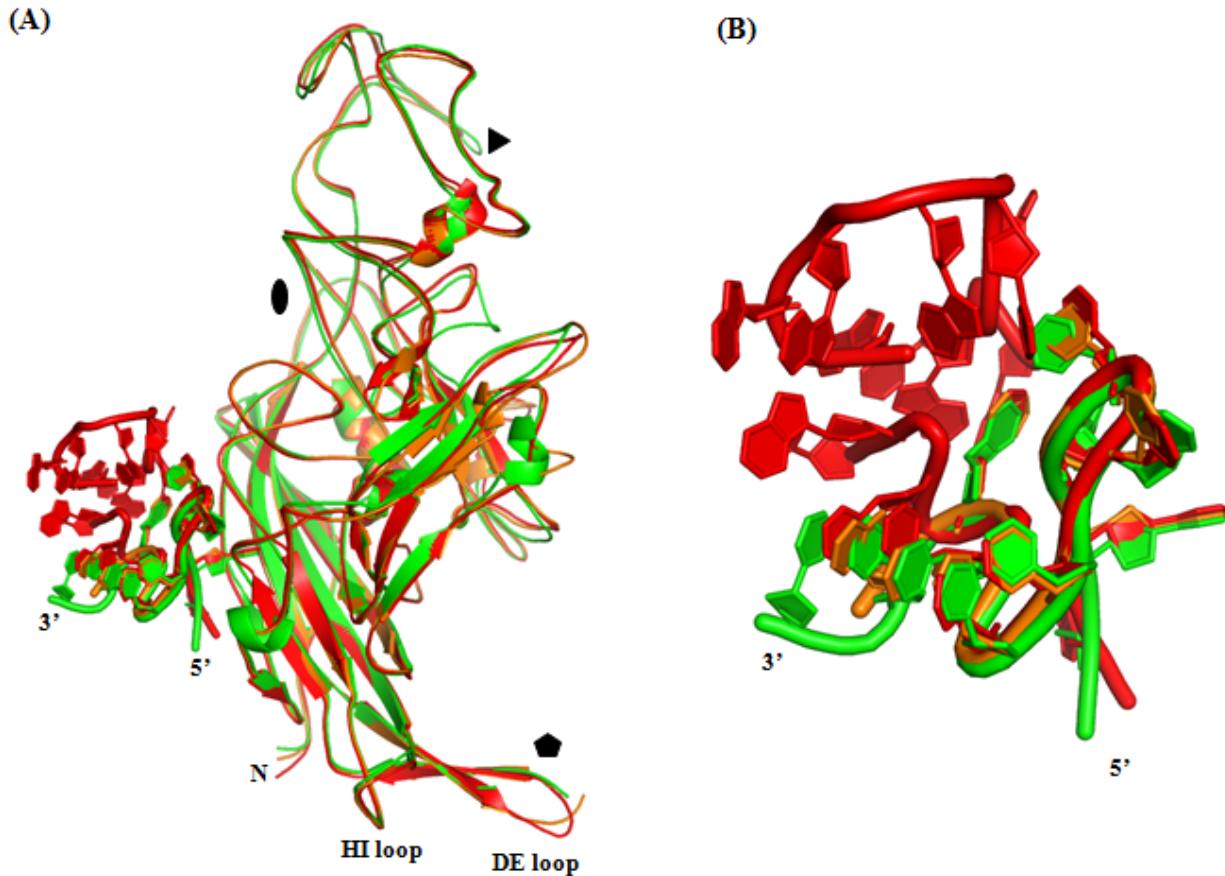


Figure 3-5. ssDNA topology observed in members of the parvovirus genus. (A and B) The conserved nucleotide binding pocket showing the ssDNA interacting with the VP2 of H-1PV (orange), MVMi (red) and CPV (green) at the capsid interior at the twofold axis. The 5' and 3' ends are labeled. (B) Close-up view of the superposition of the ssDNA observed in H-1PV, MVMi and CPV virions. Approximate positions of icosahedral two-, three-, and fivefold symmetry axes are depicted as black filled ovals, triangles, and pentagons, respectively. These figures were generated with PyMol program (82).

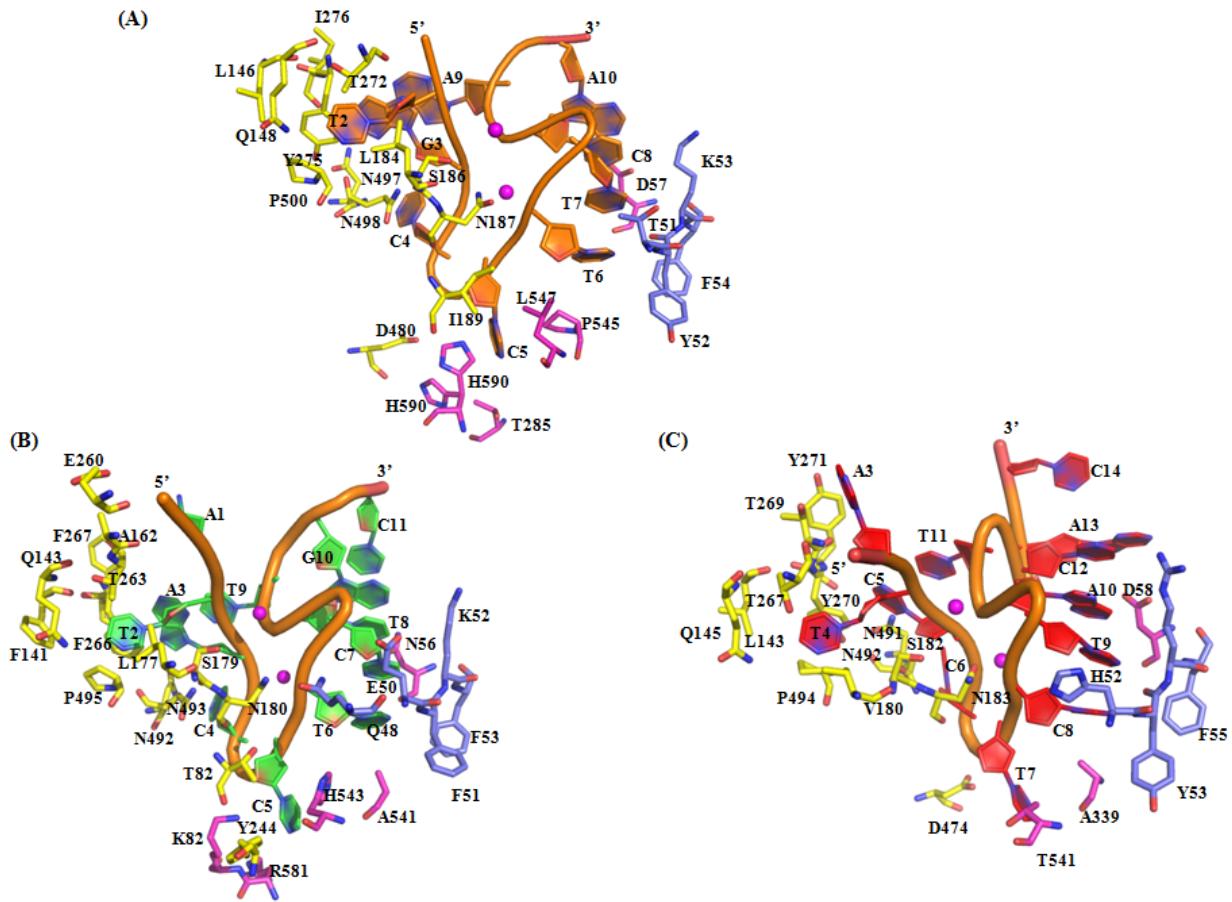


Figure 3-6. ssDNA-protein interactions in autonomous parvoviruses. ssDNA-protein interactions in (A) H-1PV, (B) CPV, and (D) MVMi are shown. Amino acids within 2.4 Å to 4.0 Å of the ordered DNA density are depicted in the stick model and labeled. The interacting residues are colored differently in yellow, pink and blue to show the contribution from the symmetry related monomers. The sugars and bases are colored according to virus type: H-1PV (orange), CPV (green), and MVMi (red), and the phosphodiester backbone colored orange. The 5' and 3' ends and the nucleotides are labeled. The Mg<sup>2+</sup> ions are depicted as magenta colored spheres.

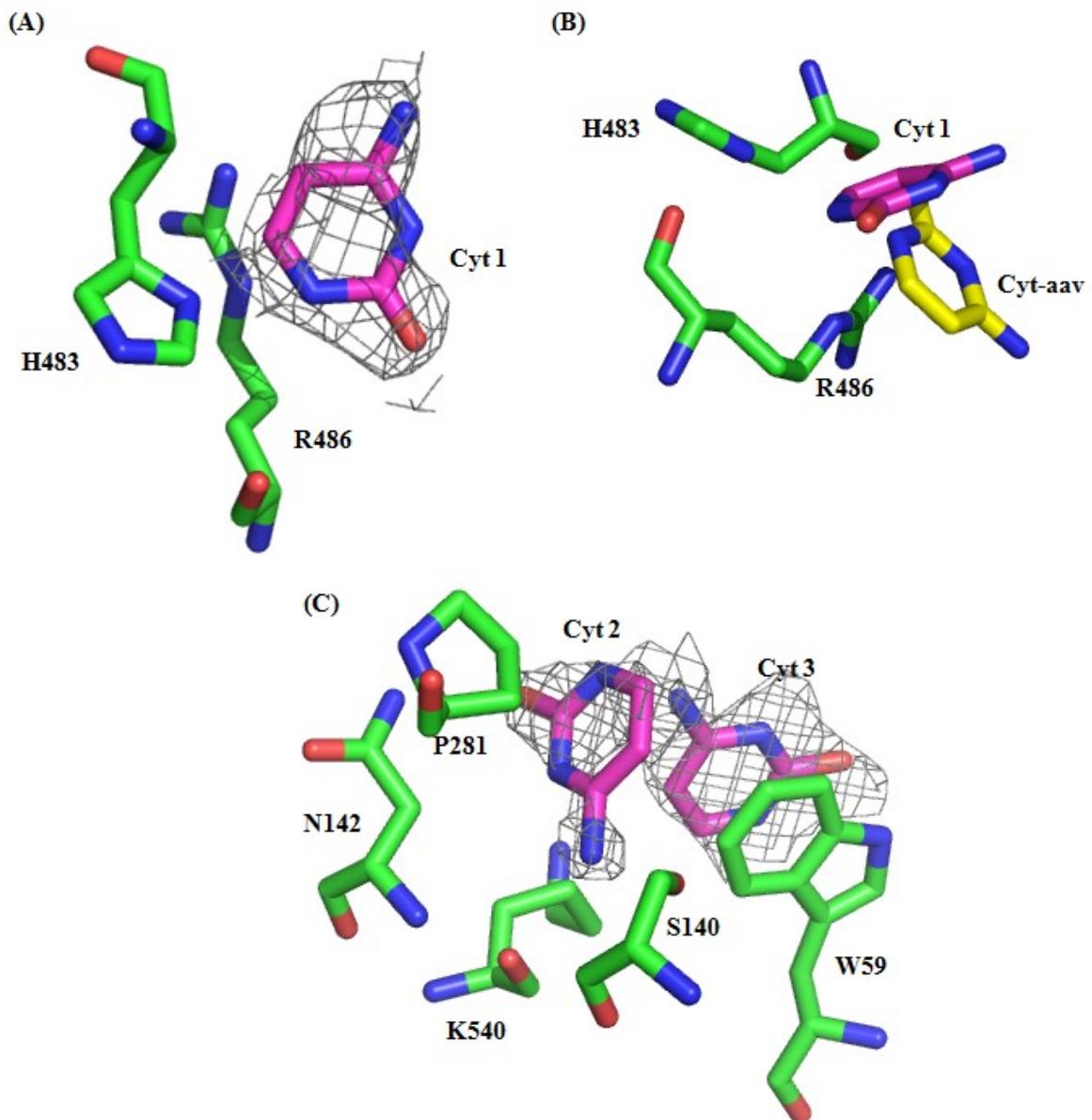


Figure 3-7. Cytosines observed in the full H-1PV structure. The amino acid residues interacting with the three cytosines are shown as green stick model, colored according to atom type. The cytosines are colored pink. In (B) the cytosine found in AAV6 structure is superposed on the full H-1PV structure and is colored yellow. The nucleotides are shown inside the  $2F_o - F_c$  difference density map (gray mesh) contoured at  $0.8\sigma$ . The amino acid residues and the nucleotides are labeled.

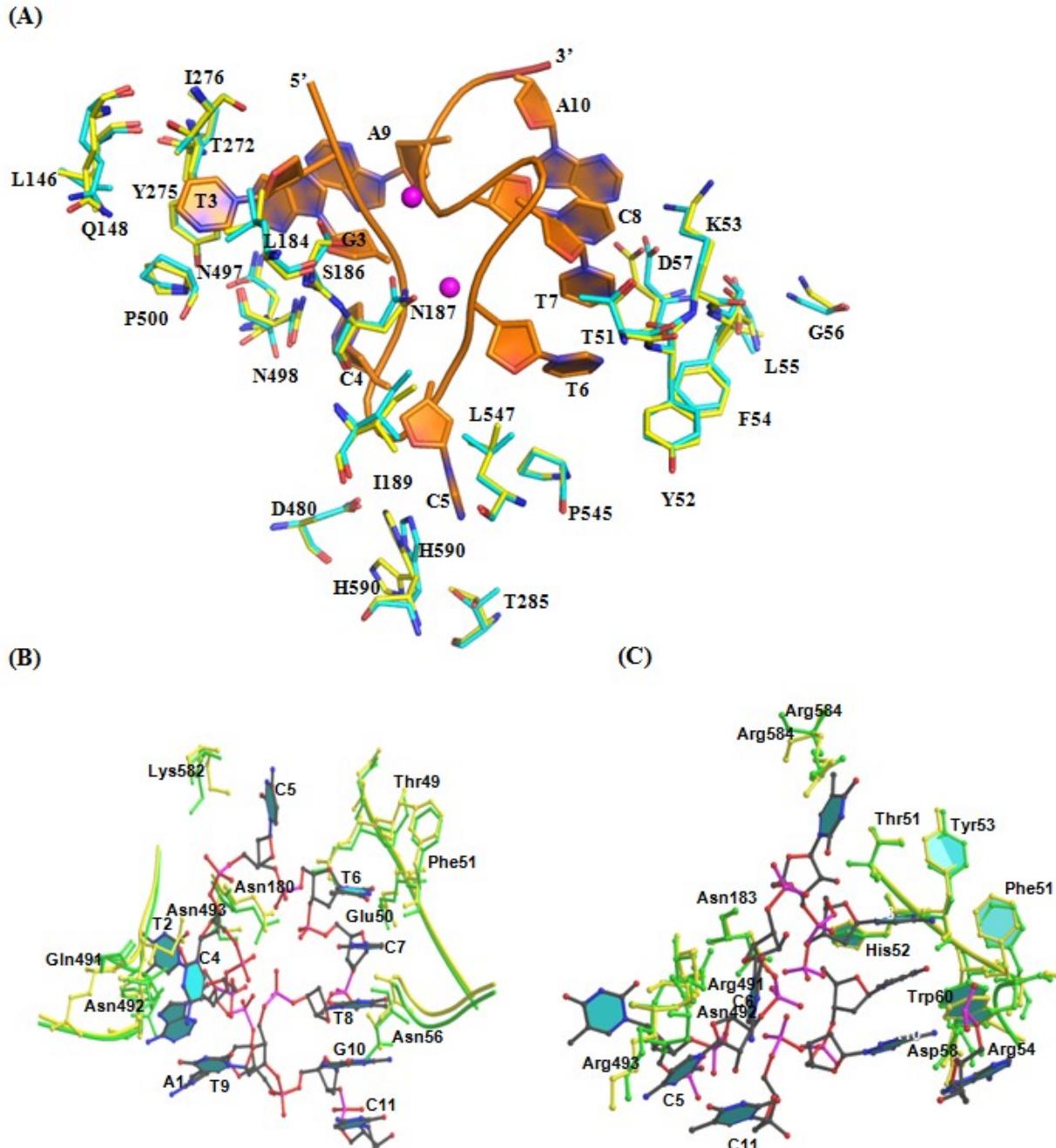


Figure 3-8. Conformational variation between empty and full capsids. (A) Conformational variation observed between VP2 structure of the full H-1PV (yellow) and empty H-1PV capsids (cyan) at the nucleotide binding pocket are shown. The sugars, bases and the phosphodiester backbone are colored orange. In (B) and (C), the full (green) and empty capsids (yellow) of CPV and MVM are compared, respectively. The sugars and bases are colored blue, and the phosphodiester backbone is colored pink. The amino acid residues and the nucleotides are labeled. The  $Mg^{2+}$  ions are depicted as magenta colored spheres.

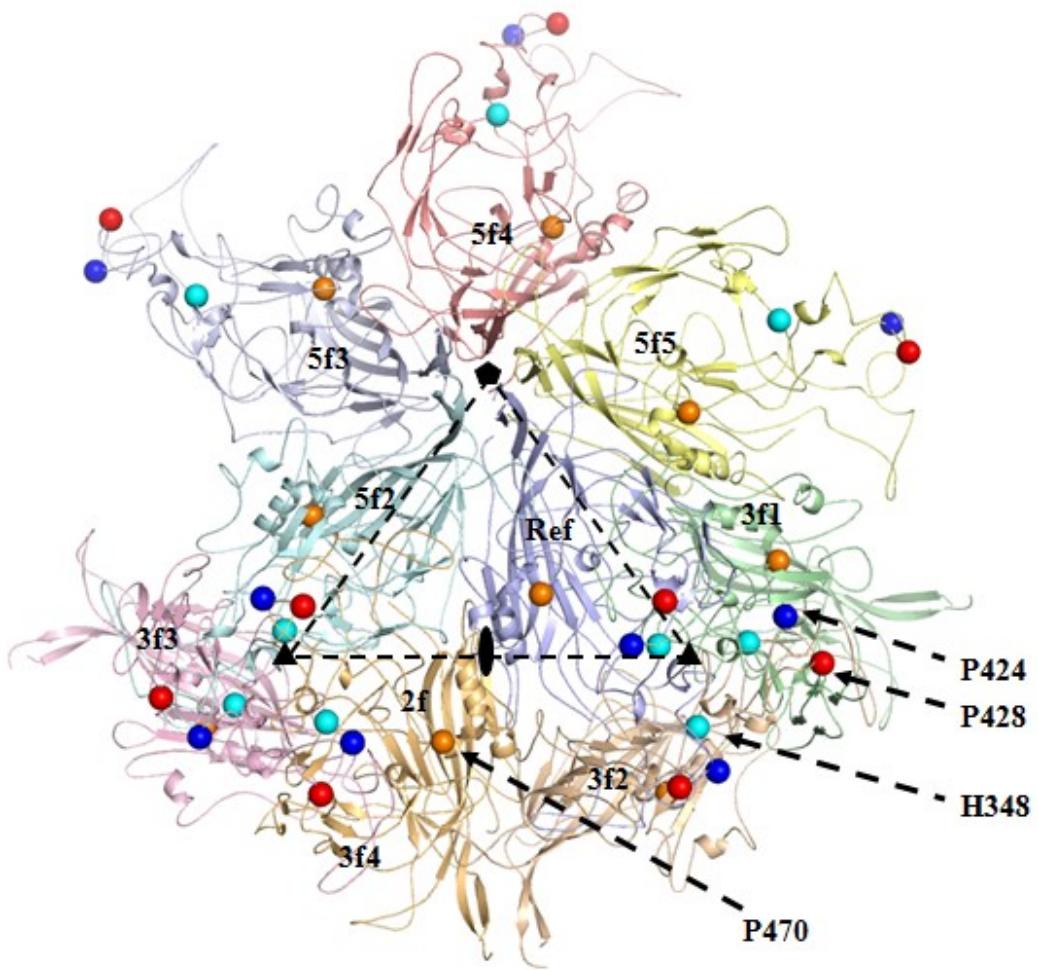


Figure 3-9. Location of the cis-peptide bonds observed in H-1PV structure. Ribbon diagram showing the secondary structure of H-1PV VP2 related to the reference (Ref) monomer by fivefold (5f<sub>2</sub>-5f<sub>5</sub>), threefold (3f<sub>1</sub>-3f<sub>2</sub>) and twofold (2f) symmetry relationships. The amino acid residues involved in cis-peptide bonds are depicted as spheres: P470 (orange), P424 (blue), P428 (red) and H348 (cyan). The black triangle depicts a viral asymmetric unit. Approximate positions of icosahedral two-, three-, and fivefold symmetry axes are depicted as filled ovals, triangles, and pentagons, respectively. This figure was generated with PyMol program (82).

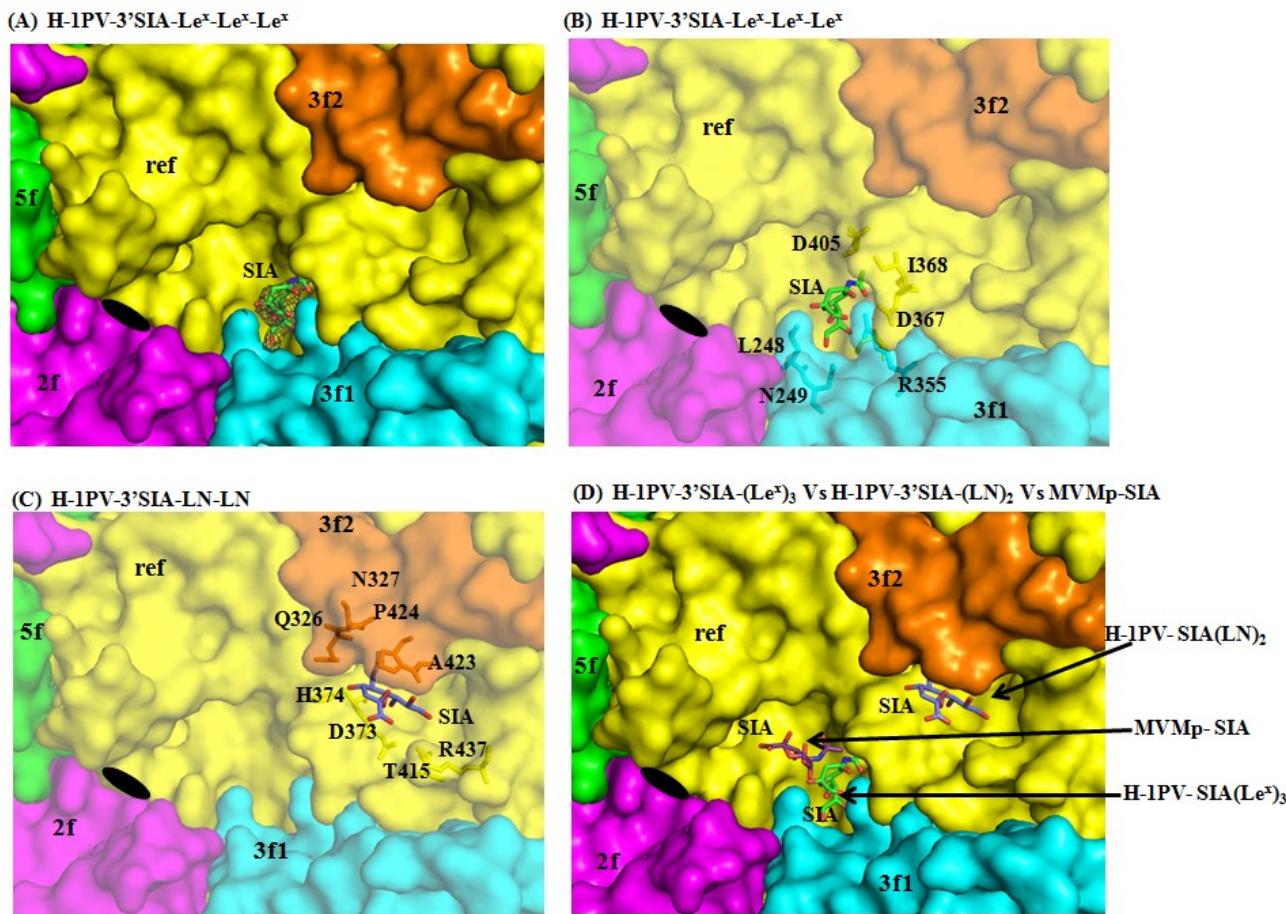


Figure 3-10. Sialic acid binding site on the H-1PV capsid. Surface representation of a close-up of the depression at the icosahedral twofold axes of the capsid showing the SIA binding site on H-1PV. In (A and B), the SIA binding site identified in the H-1PV-3'SIA( $\text{Le}^x$ )<sub>3</sub> complex is shown. The residues that interact with SIA are shown in stick model and are in the same color as the surface. The SIA is shown in stick model (carbon atoms in green) inside a 2F<sub>o</sub>-F<sub>c</sub> density map (dark grey) at contoured at 1 $\sigma$ . In (C) the SIA modeled (carbon atoms in blue) for the H-1PV-3'SIA(LN)<sub>2</sub> complex is shown. In (D), the SIA (carbon in purple) modeled on the MVMP capsid surface is superimposed on the H-1PV capsid surface. The approximate location of the icosahedral twofold axes is shown by the filled oval.

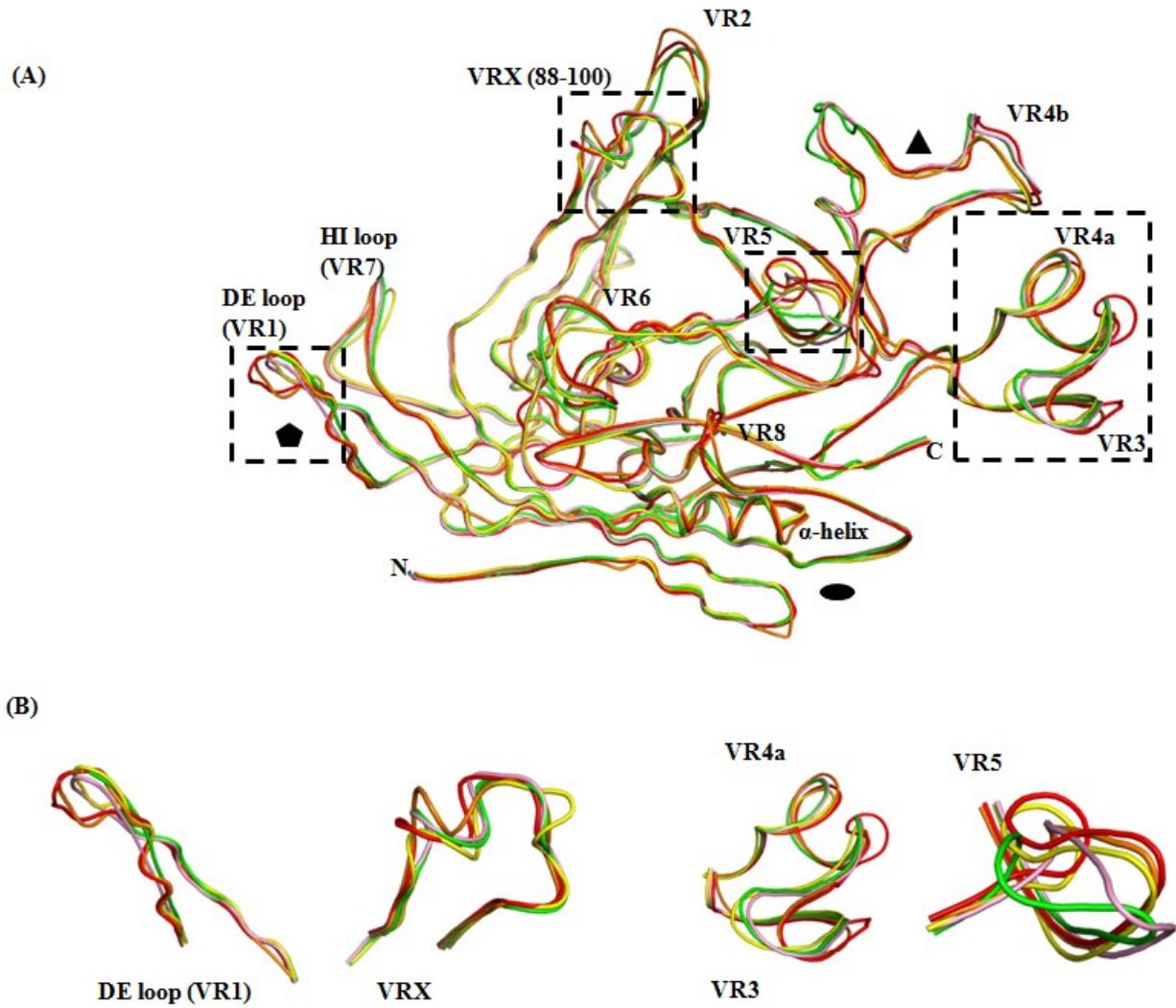


Figure 3-11. Comparison of VP2 structures of H-1PV, MVM, PPV, FPV and CPV. (A) Superimposition of the coil representation of VP2 monomer structures for members of parvovirus genus, H-1PV (orange), MVM (red), PPV (yellow), FPV (pink), and CPV (green). The common variable regions are labeled. The DE and HI loops are labeled. The approximate icosahedral twofold (filled oval), threefold (filled triangle), and fivefold (filled pentagon) axes are shown. (B) Close-up views of the regions inside dashed boxes. These figures were generated with PyMol program (82).

CHAPTER 4  
PARVOVIRUS CAPSID DYNAMICS ASSOCIATED WITH ENDOSOMAL  
TRAFFICKING

**Background**

The ssDNA parvoviruses enter host cells through receptor mediated endocytosis, and undergo endosomal pH mediated processing in the early to late endosomes as well as in lysosomes prior to nuclear entry for replication (infectious life cycle of parvoviruses described in detail in Chapter 1). It is known that the capsid undergoes structural transitions during the entry process so as to enable the (i) externalization of VP1u that harbors PLA<sub>2</sub> activity essential for during endosomal escape, and also nuclear localization signals (NLS) required for nuclear entry, (ii) exposure of the VP2 N-terminus that contains nuclear export signal (NES) required to traffick newly formed virions out of nucleus and the cleavage of VP2 N-terminus to form VP3 in the autonomous parvoviruses that undergo a maturation step, and (iii) extrusion of the genome. Mutagenesis studies on AAV2 and MVM have predicted the fivefold channel to be the portal for these exposures (32, 74, 100, 101, 121, 170, 240, 279, 323). These extrusion events can be mimicked using limited heat shock and/or coupled with pH treatment with the capsid remaining intact (65, 67, 100, 132, 170, 191, 252, 312). However, the mechanisms of capsid endosomal processing during trafficking to the nucleus for genome replication are poorly understood. The available parvovirus capsid structures show that the residues within these N-terminal regions cannot be accommodated through the fivefold channel without structural re-arrangement of the surrounding β-ribbons and immunological studies support the "breathing" of this pore (65). To gain insight into the capsid structural transitions during infection, the crystal structures of AAV8 were determined at pH 4.0, pH 5.5, pH 6.0, and pH 7.5 after incubation at pH 4.0,

to mimic the conditions encountered during trafficking to early endosomes, late endosomes, lysosomes and endosomal escape to the cytosol prior to nuclear entry (208). Significant amino acid side chain conformational transitions were observed inside the capsid at the threefold axis and on the capsid exterior at the icosahedral twofold axis. At low pH, conformational changes at the capsid interior disrupted the interaction of the VP with the ordered nucleic acid density observed in crystal structures and this was suggested to prime genomic uncoating. Structural changes in the capsid exterior at the icosahedral twofold axis resulted in a reduction in the number of inter monomer interactions at the twofold as pH was reduced. This “weakening” of the interface was proposed to be involved in capsid destabilization events that enable AAV VP1u externalization without capsid disassembly. Similar conformational changes were observed in an equivalent site in AAV1 (unpublished data). The structural transitions induced in AAV8 at low pH were observed to be reversible that would stabilize the capsid following VP1u exposure. This was the first study on the dependoviruses characterizing the capsid dynamics at endosomal pHs. It has been observed that the structurally variable capsid surface residues dictate the differences in tissue tropism and transduction efficiency observed among the AAV serotypes (unpublished data).

For the autonomous parvoviruses, capsid dynamics have been structurally characterized only for CPV and FPV. The capsid structure of CPV was determined at pH 7.5 and pH 5.5 to mimic the endosomal pH. Structures for CPV and FPV were also determined at pH 7.5 and pH 6.2 in the presence or absence of  $\text{Ca}^{2+}$  ions to understand the differences in the hemagglutination ability of CPV and FPV (274). While, CPV is able to hemagglutinate at pH 6.2 and pH 7.5, FPV can only do so at pH 6.2 (52). This

study showed that the largest structural difference occurred in a surface loop (also called the 'flexible' or hemagglutinating' loop in previous studies and consists of CPV VP2 residues 359-375) located at the wall of the twofold depression and colocalizes with the SIA binding site and host range determinants (274). This flexible loop was disordered in CPV at pH 5.5 and pH 6.2, and in FPV in the presence of EDTA. This loop binds to  $\text{Ca}^{2+}$  ion in FPV (only at pH 7.5) but not in CPV (at any of the studied pHs) and is adjacent to a double  $\text{Ca}^{2+}$  binding site which is conserved in both FPV and CPV. These results showed that the capsid dynamics in CPV and FPV are influenced by the  $\text{Ca}^{2+}$  ion concentration and are related to their hemagglutinating activity (274). Utilizing proteinases, minor structural variations in the surface loops of the CPV capsids were observed at pH 4.5 but not at pH 5.5 (213). However, these studies on CPV and FPV did not provide information on the capsid dynamics at pH 4.0 where the conformational changes observed in AAV8 were more pronounced. The capsid surface residues at the twofold depression differ between CPV, FPV and MVM, thus it is likely that the local surface charge variation among the autonomous parvoviruses might result in different structural transitions during endosomal trafficking. To understand the mechanism of structural transitions observed in MVM capsids during endosomal trafficking, the structures of MVM capsids were determined in this study at pH 4.0, pH 5.6 and pH 6.0 to mimic the pH encountered in the early endosomes, late endosomes and lysosomes, respectively.

## Experimental Methods

### Virus Production and Purification

The MVMp VLPs were produced and purified as described in the Experimental Methods section of chapter 2. For the pH experiments, the purified virus samples were

dialyzed into the appropriate buffers; pH 4.0 and pH 5.6 (150 mM Sodium Acetate buffer), pH 6.0 (150 mM NaCl, 10 mM bis-Tris-HCl). The purity and integrity of the capsids were verified by SDS-PAGE and negative stain EM, respectively (as in chapter 2).

### **Crystallization, Data Collection and Processing**

Crystals of MVMP VLPs at the various pHs were grown using the hanging-drop vapor-diffusion method (199) with VDX 24-well plates and siliconized cover slips (Hampton Research, Laguna Niguel, CA, USA) in screening solutions containing PEG 8000 (1-3% w/v) and 150 mM NaCl in 10 mM bis-Tris-HCl (pH 6.0) or 150 mM Sodium Acetate buffer (pH 4.0 and 5.6). The drops were prepared by mixing 2  $\mu$ l of virus solution (10 mg/ml) with 2  $\mu$ l of reservoir solution and then equilibrated against 1 ml reservoir solution at RT. For pH 6.0 and 5.6, crystals isomorphous to those for MVMP-VLPs at pH 7.5 were obtained, but for pH 4.0, ‘flat’ crystals were observed. Crystals obtained were soaked for 30 s in cryoprotectant solution containing the precipitant solution with 10% PEG 8000 and 30% glycerol and flash cooled in liquid nitrogen vapor prior to X-ray diffraction data collection.

Diffraction data for the MVMP-pH 6.0 crystals was collected at the F1 beamline at CHESS on an ADSC Quantum 270 CCD detector, for the MVMP-pH 5.6 crystals at the SERCAT 22ID beamline at APS, and for the MVMP-pH 4.0 crystals at the X29 beamline at BNL on an ADSC Quantum 315 CCD detector. For MVMP-pH 6.0, a total of 200 usable images were collected from one crystal with a crystal-to-detector distance of 300 mm, an oscillation angle of 0.3° per image, and an exposure time of 25 s per image at a wavelength of  $\lambda = 0.9180 \text{ \AA}$ . A total of 500 usable images were collected from three crystals of MVMP-pH 5.6 at a wavelength of  $\lambda = 1.000 \text{ \AA}$ , with a crystal-to-detector

distance of 350 mm, an oscillation angle of 0.3° per image, and an exposure time of 4 s per image. For MVMP-pH 4.0, a total of 230 usable images were collected from two crystals with a crystal-to-detector distance of 300 mm, an oscillation angle of 0.3° per image, and an exposure time of 8 s per image at a wavelength of  $\lambda = 1.0809 \text{ \AA}$ . The crystals diffracted X-rays to beyond 3.8 Å, 3.3 Å and 3.2 Å resolution for pH 6.0, 5.6 and 4.0, respectively. The measured diffraction intensities were indexed and integrated with the HKL2000 suite of programs (224), and scaled and merged with SCALEPACK (224). The space group for the MVM pH 6.0, pH 5.6 and pH 4.0 MVMP crystals was determined to be C2. The unit cell volume for the pH 4.0 crystals was smaller and half of that of pH 6.0 and pH 5.6 crystals. The MVMP pH 6.0, pH 5.6 and pH 4.0 data sets scaled with an  $R_{\text{sym}}$  of 16.0% (48.4% completeness), 13.3% (94.6% completeness), and 11.9% (72% completeness), respectively. The data processing statistics are summarized in Table 4-1.

### Structure Determination

The diffraction intensity data sets were converted to structure factor amplitudes using the TRUNCATE program from CCP4 (Collaborative Computational Project, Number 4) (62). For the MVMP pH 6.0 and pH 5.6 data, the C2 unit cell contains two half particles with different orientations in the crystallographic asymmetric unit, as described for the previously solved MVMi and MVMP-VLP structures (7, 168, 184). The MVMP VP2 VLP structure coordinates (PDB accession no. 1Z14) (168) was used as the phasing model to initiate molecular replacement and refinement was carried out using the CNS program (46) as detailed in the Experimental Methods section of chapter 2. The refinement statistics are given in Table 4-1. The figures were generated using program PYMOL (82).

For the MVMP pH 4.0 data, the Matthew's coefficient ( $V_M$ ) was calculated to be  $3.16 \text{ \AA}^3 \text{ Da}^{-1}$ , corresponding to a solvent contents of 61%, with two particles in the unit cell related by a crystallographic twofold axis and the icosahedral twofold axis coincident with the crystallographic twofold axis (i.e., one half capsid or 30mer occupying a crystallographic asymmetric unit) (195). The particle orientation in the crystal unit cell was determined with a self-rotation function using the General Lock Rotation Function (GLRF) program (299) computed with 10% of the observed data between 10.0 and 5.0 Å resolution as large terms at  $\kappa=72^\circ$ ,  $120^\circ$  and  $180^\circ$  to search for the twofold, threefold and fivefold icosahedral symmetry axes, respectively. The radius of integration was set to 120 Å. The rotation function for  $\kappa=180^\circ$  established the orientation of the icosahedral symmetry axes relative to the crystal axes and confirmed that the two particles in the unit cell are related by the crystallographic twofold axis (b axis). A polyalanine model of the MVMP VP2 crystal structure (PDB accession no. 1Z14) (168) was generated by the MOLEMAN program (167) and expanded to 30 subunits (half capsid or 30-mer) using the Oligomer Generator subroutine available at the VIPER database (50) for use as a molecular replacement phasing model (258). Molecular replacement was carried out using the AutoMR subroutine in PHENIX (2). The best solution gave a Log Likelihood Gain (LLG) score of 2208, z-score for rotation function of 42.3 and z-score for translation function of 50.2 and  $R_{\text{factor}}$  of 53.8%, with an orientation, in Eulerian angles, of  $\alpha=179.99^\circ$ ,  $\beta=65.64^\circ$ , and  $\gamma=270.04^\circ$  and fractional coordinate position of 0.5003, 0.0000, and -0.0000. The initial phases for the oriented and positioned MVMP polyalanine model were further improved by refinement in the CNS program using simulated annealing, energy minimization, conventional positional,

and individual temperature factor (B-factor) refinement (46), while applying strict 30-fold noncrystallographic symmetry (NCS). The coordinates of the MVMP-VLP structure at pH 7.5 (PDB accession no. 1Z14) (168) was superposed onto the refined MVMP structures at pH 6.0, pH 5.6 and pH 4.0 using the LSQ subroutine in COOT (97) to identify conformational changes.

## Results

Diffraction data sets were obtained for MVMP VLP at pH 6.0, pH 5.6 and pH 4.0. The MVMP crystals at pH 5.6 and pH 6.0 were isomorphous to those of MVMP-VLP and wt MVMi for which structures are available (7, 168). Residues 39 to 587 of the MVMP VP2 were built into the MVMP electron density maps. The averaged density maps were not interpretable beyond N-terminal residue 39 of VP2, as was previously reported for the MVMP and MVMi VLP structures (168). The resolution for the MVMP data at pH 6.0 is low (3.8 Å) and the refinement statistics ( $R_{\text{factor}}/R_{\text{free}}$  is 40.2/40.6) are also not comparable to previously solved structures for MVM (7, 168, 184). The refined structure at resolution of 3.8 Å did not show any detectable differences in main chain or side chain conformations in comparison to MVMP structure determined at pH 7.5. The MVMP pH 5.6 structure was determined from a very complete data set (94.6% complete) and the current refinement  $R_{\text{factor}}/R_{\text{free}}$  is 30.5/30.6. The overall topology of VP2 at pH 5.6 is very similar to that at pH 7.5, with no main chain differences but significant amino acid side chain conformational changes at four regions on the capsid surface. The first is on the exterior capsid surface at the interface between the fivefold monomers and involves amino acid residue E79. Reducing the pH to 5.6 disturbs the interaction of E79 with the neighboring residues, R500 and R518 from the same monomer and K204 from a fivefold related monomer (Figure 4-1). The second change is on the capsid exterior on

the wall of the fivefold channel. At pH 5.6, K166 in the β-ribbon in DE loop orients away from E157 of the same monomer and towards D507 of a fivefold related monomer (Figure 4-1). The third change involves residue Q158, which is involved in fivefold symmetry related interactions at the neck of the fivefold channel. At pH 5.6, Q158 moves away by 1 Å towards the wall of the channel, thus widening the pore diameter by 2 Å (Figure 4-1). The fourth change occurs inside the capsid at the icosahedral twofold axis near the nucleotide binding pocket identified for CPV and MVMi (7, 306), where the side chain of R54 orients towards E62 from the same monomer at pH 5.6 (Figure 4-2A). The structure refinement for MVMP pH 4.0 data is not complete yet. However, preliminary analysis of the averaged density map and comparison to MVMP structure at pH 7.5 shows main chain differences at the HI loop.

### Discussion

The structures of MVMP VLPs have been determined at pHs that mimic the environment encountered by MVMP during endosomal trafficking. The data for MVMP at pH 5.6 is the most complete. Comparison to MVMP structure at pH 7.5, identified amino acid side chain conformational changes mainly at subunit interfaces and in the capsid interior at the twofold axis near the nucleotide binding pocket. The conformational change of E79 reduces the intersubunit contacts between the VP monomers at the fivefold interface. The change in the orientation of K166 to interact with D507 might contribute to the DE loop flexibility at reduced pH. The Q158 of five fivefold symmetry related monomers act as sentries at the opening of the fivefold channel. The side chain amino acid changes observed for E79, K166 and Q158 at or near the fivefold axis of symmetry might be involved in capsid destabilization that primes the capsid for VP1u or genome extrusion through the fivefold channel. Conformational changes that 'weaken'

the intersubunit interfaces was also observed in the AAV8 structures solved at low pHs (208). Residues K166 and E79 are conserved amongst the members of parvovirus genus. Q158 is conserved for the rodent parvoviruses (S158 in CPV, FPV and PPV).

In full MVMi capsids (genome containing), R54 interacts with an adenine at pH 7.5 (Figure 4-2B). The R54 conformation observed in the MVMP VLP at pH 5.6 would prevent its interaction with the nucleotide. The weakening of the capsid-DNA interaction at pH 5.6 might prepare the genome for its release from the capsid. Similar observations were made for the AAV8 structure solved at pH 4.0, where the conformation change in H632 resulted in loss of interaction with the ordered nucleotide (208). R54 is replaced with K in CPV, FPV, H-1PV, Lull and with Q in PPV. While in MVMi, R54 interacts with adenine, in CPV and H-1PV the interacting nucleotide is a cytosine and thymine, respectively ((7, 306), and data from studies conducted in chapter 3). It is possible that due to the differences in capsid-DNA interactions the autonomous parvoviruses might respond differently to the endosomal pH and exhibit different genome uncoating kinetics. The conformational change of the HI loop observed at pH 4.0 might be related to the opening of the fivefold channel as was observed in AAV2 upon HS binding (178). The structural transitions observed for the MVMP structures at the different pHs were subtle compared to that observed in CPV and FPV, where a loop adopts a different conformation or becomes disordered depending on the pH (274).

The MVMP structures determined at pH values of 6.0, 5.6 and 4.0 to mimic the endosomal trafficking environment demonstrated that although the capsid VP topologies of all the structures were very similar, significant amino acid side chain

conformational changes were observed at reduced pH on (a) the interior surface of the capsid at the icosahedral twofold axis near the nucleotide binding pocket, and (b) the exterior capsid surface at or near the fivefold channel. These structural transitions that disrupt intersubunit contacts and capsid-DNA interactions and widen the fivefold channel are consistent with capsid destabilization events that likely facilitate VP1 $\alpha$  extrusion for PLA2 activity, VP2 N-terminus exposure for capsid maturation and capsid priming for genome release.

Table 4-1. Data processing and refinement statistics

Parameter	MVMp-pH 6.0	MVMp-pH 5.6	MVMp-pH 4.0
Wavelength ( $\lambda$ , Å)	0.918	1.000	1.0809
Space Group	C2	C2	C2
Unit cell parameters (Å, °)	a=437.9, b=408.7, c=299.6, $\beta=95.5$	a=442.0, b=411.7, c=301.8, $\beta=95.9$	a=250.7, b=433.9, c=250.7, $\beta=117.1$
Resolution (Å)	50-3.8 (3.9-3.8) <sup>a</sup>	40-3.3 (3.4-3.3) <sup>a</sup>	40-3.2 (3.3-3.2)
Completeness (%)	48.4	94.6	72.0
Redundancy	1.9	2.6	2.0
$R_{\text{sym}}^{\text{b}}$ (%)	16.0	13.3	13.1
$R_{\text{factor}}^{\text{c}}/R_{\text{free}}^{\text{d}}$ (%)	40.2/40.6	30.5/30.6	33.1/33.3

<sup>a</sup>Values in parenthesis are for highest resolution shell.

<sup>b</sup> $R_{\text{sym}} = (\sum |I - \langle I \rangle| / \sum \langle I \rangle) \times 100$ .

<sup>c</sup> $R_{\text{factor}} = (\sum |F_o - |F_c|| / \sum |F_o|) \times 100$ .

<sup>d</sup> $R_{\text{free}}$  is calculated the same as  $R_{\text{factor}}$ , except it uses 5% of reflection data omitted from refinement.

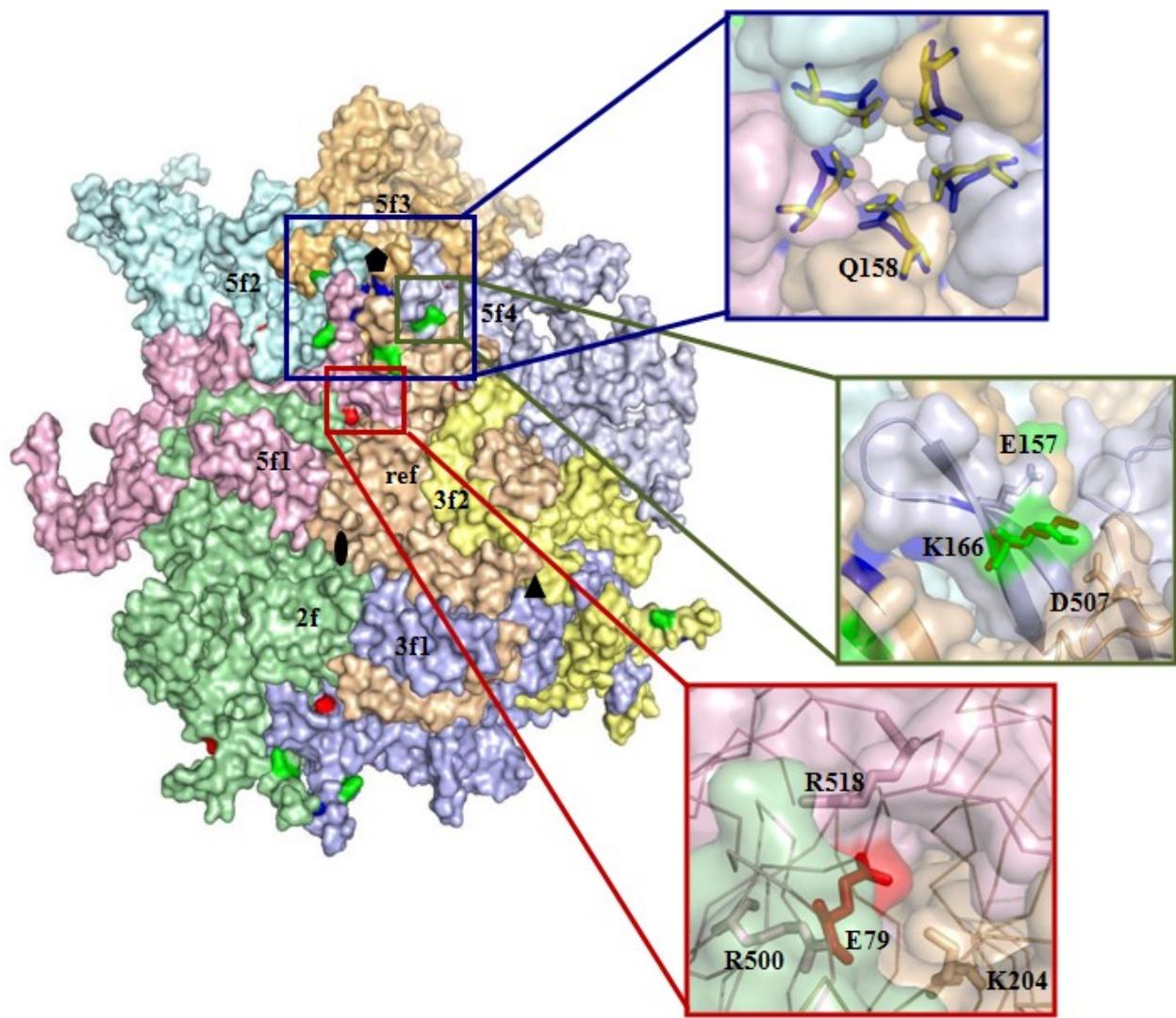


Figure 4-1. MVMp capsid dynamics on the exterior surface. The amino acid residues that show side-chain conformational changes in the MVMp structure at pH 5.6 are colored as red (E79), blue (Q158) and green (K166), respectively. The symmetry related VP2 monomers are colored differently and labeled. The inset provides a close-up view of the structural variations; At the shoulder of the canyon, K166 at pH 5.6 (green) moves closer to D507 (wheat) of a fivefold related monomer and moves away from E157 (light blue) of the same reference monomer, as compared to at pH 7.5 (K166 in brown); inside the fivefold pore of the MVMp capsid, Q158 from the fivefold related monomers at pH 5.6 (in blue) and pH 7.5 (in yellow) is shown. The side chains of the interacting residues are shown in stick model. The approximate icosahedral twofold (filled oval), threefold (filled triangle), and fivefold (filled pentagon) axes are shown. These figures were generated with PyMol program (82).

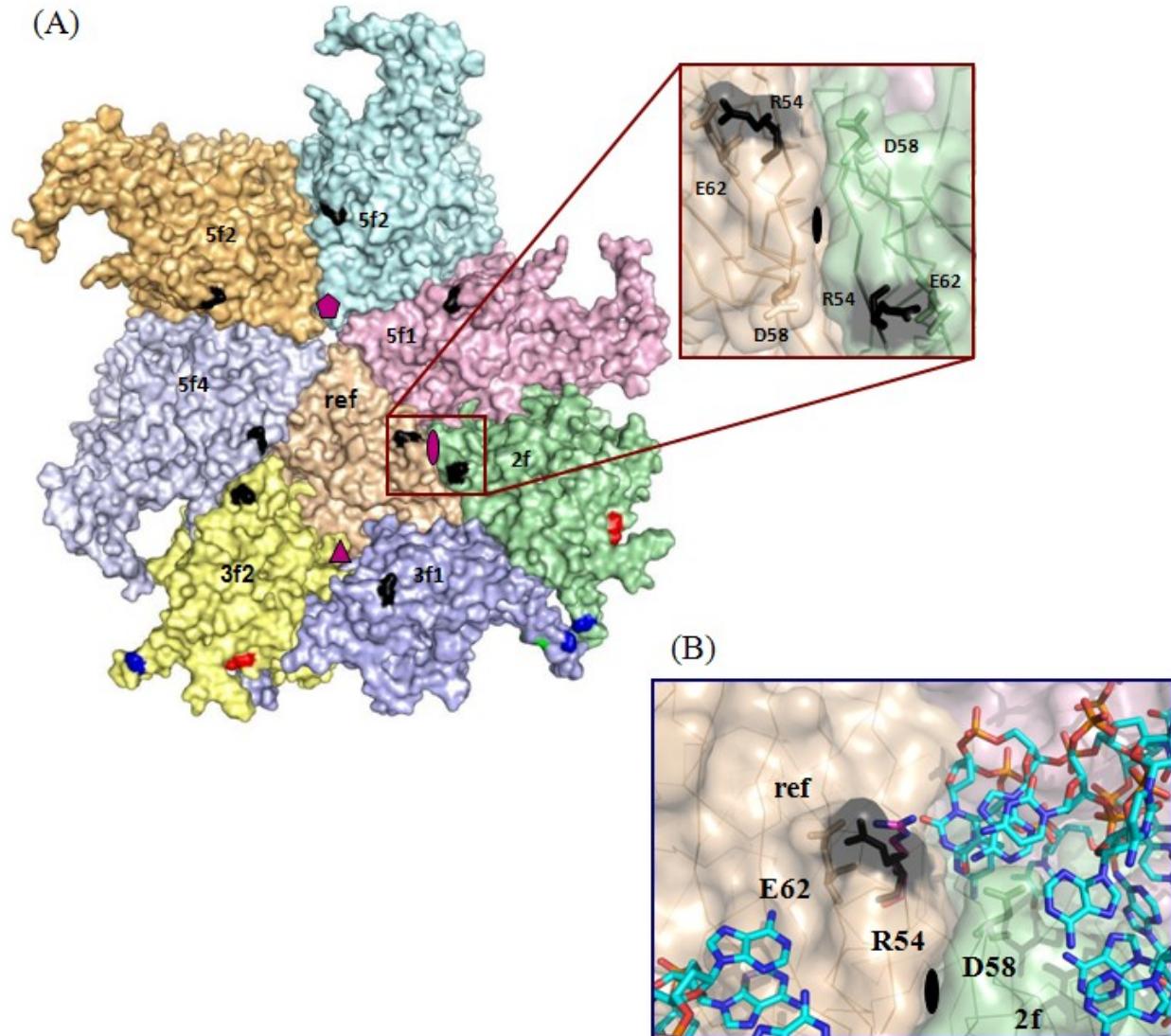


Figure 4-2. Capsid dynamics in the interior of the MVMp capsid. (A) The amino acid residues that show side chain differences at pH 5.6 are colored as red (E79), blue (Q158), green (K166), and black (R54), respectively. Inset shows a close-up view of the twofold axis. The reference (Ref) monomer, the twofold (2f) related monomer and the residue R54 are depicted in wheat, green and black, respectively. (B) The side chains of R54 and D58 interact with ordered nucleotides in infectious virions. Nucleotides and amino acid side chains are shown as stick representation and colored according to atom type. The side chain of R54 interacts with DNA at pH 7.5 (carbon atoms in magenta, nitrogen in blue and oxygen in red), but moves away at pH 5.6 (R54 colored in black).

## CHAPTER 5

### SUMMARY AND FUTURE DIRECTIONS

The objective of this study was to utilize MVM and H-1PV, autonomous members of the *Parvoviridae* family, as models to map the capsid regions that dictate surface glycan receptor attachment, cell tropism, host range and endosomal trafficking of these very similar viruses. The structure determination studies of MVM capsid-SIA glycan complexes, showed that all the glycans bound in the vicinity of the twofold depression on the MVM capsid that had been identified as the SIA binding site in previous structure studies (188). In this study we were able to visualize a longer oligosaccharide in the binding pocket as compared to the previous studies. The binding pocket seemed to accommodate the various glycans in different conformations and the capsid interactions with the terminal SIA were not conserved, except in the MVMP-3'SIA(Le<sup>X</sup>)<sub>3</sub> and MVMP-3'SIA(LN)<sub>3</sub> complexes. Although, the capsid-glycan interactions were not conserved, a few capsid surface residues made non-specific contacts with the glycans on all the MVM capsids. This study structurally verified the role of residues 362, 368, 399 and 558 that are also cell tropism and host pathogenicity determinants for MVM, in the SIA glycan binding. It also identified the possible role of residues 321 and 366 in the recognition of α2-8 linked SIA glycans by MVMI. However, the heterogeneity of the terminal SIA binding observed for these structures, which could be due to the promiscuous nature of the binding pocket or the inherent limitations of the crystallography technique, precluded the identification of capsid surface determinants involved in the specificity of the terminal SIA binding. To define the specificity of the MVM capsid interactions with the commonly recognized terminal SIA and to minimize the heterogeneity inherent in the crystallographic technique, multiple diffraction data

sets for native MVMP capsid and for MVMP complexed with SIA alone need to be separately solved (no merging of data from different crystals) and  $F_o$ - $F_e$  difference Fourier maps calculated. If this strategy shows SIA binding on the MVMP capsid in one particular conformation, then it could be compared to SIA binding on the MVMi capsid to analyze the differences in receptor binding. In addition to the structural studies, residues E321 and M366 in MVMi could be mutated to verify their suggested role in recognition of  $\alpha$ 2-8 linked SIA glycans. Also, the mutant viruses could then be screened on the glycan array to determine any changes in binding specificities. Cell binding competition assays using these recognized glycans to correlate the glycan screening, glycomic profiling and structural data have been initiated and will be completed. The utilization of a common SIA binding pocket by all MVM viruses that colocalizes with tropism and pathogenicity determinants suggests the role of this receptor binding pocket in cell recognition and the differences in SIA receptor interactions could explain the differences in pathogenicity but the mechanism for differences in cell tropism is still unclear.

Screening of MVM viruses on the derivatized SGM showed that both MVMP and MVMi specifically bound only to the  $\alpha$ 2-3 linked sialylated derivatives, which was consistent with the previous glycan array screening (209). It also identified 9-O-methylation (and 9-O-acetylation and 9-O-lactoylation for MVMP viruses) as an additional component of their SIA recognition. However, preliminary modeling of these SIA derivatives in the glycan binding pocket failed to provide an understanding of the mechanism of specific recognition of these derivatives. The 9-O acetylated and 9-O-lactoylated SIA derivatives are present in nature, but the 9-O-methylated derivative has not yet been isolated from natural sources, so the significance of this recognition is unknown. The binding profile

for VLPs, empty and full particles on the SGM was similar which validates the use of VLPs in lieu of infectious virions for structural and biochemical studies that examine receptor interactions. In addition, all the MVM viruses bound to a biantennary SIA glycan with 3'SIA-LN motif that was not present on the previously used glycan arrays. Glycomic profiling of the cell lines permissive for MVMP and/or MVMi validated the presence of the glycans that were recognized by the MVM viruses in the previous glycan array and the current SGM screening, such as glycans with SIA-LN motif and SIA-Le<sup>x</sup> motif, the α2,8 multisialylated glycans, and the biantennary glycans. The SIA-Le<sup>x</sup> cancer cell motif is only expressed on the NB324K cell surface which is a SV40 transformed kidney fibroblast cell line and this was consistent with MVM's oncotropism. The similar glycan profile of the A9 and EL4 T cells failed to explain the differences in cell tropism between these MVM viruses based on differential receptor recognition. Based on previous published data (15, 112, 196) and the results of this study, it is evident that the capsid twofold pocket plays a role in dictating differences in cell tropism post cellular entry. The post entry event could be interaction with a cellular host factor present only in permissive cell line or the correct cell signaling pathway initiated in the permissive cell line triggered by ambient utilization of SIA receptor. Covalent tagging of MVM capsids with peptides such as poly-Histidine, FLAG, or GST (Glutathione) followed by immunoprecipitation experiments using cell lysate could be conducted to identify the cellular host proteins that interact with the capsid post entry and during trafficking. The differences in protein expression or abundance between the permissive and restrictive cell lines following initiation of cell signaling pathways could be studied by Two-Dimensional Difference Gel Electrophoresis (2D-DIGE), wherein the proteins

expressed in the two cell lines would be labeled with different fluorescent dyes and then the samples mixed and separated based on isoelectric point and molecular weight in the same gel. This experiment could also be utilized to compare protein expression before and after virus infection. The role of cell signaling could be probed by conducting a genome wide microarray before and after virus infection in the different cell lines to identify differences in cellular transcription. Also, small molecule inhibitors of specific cell signaling pathways could be used to confirm the role of a particular pathway in MVM infection. Recently developed techniques, such as Fluorescent Cell Barcoding (where samples are labeled with different intensities of a single fluorophore), in combination with phospho-epitope-specific flow cytometry (flow cytometry based analysis of different phosphorylated proteins using specific antibodies) would allow for high throughput detection of differences in cell signaling networks.

The structure of H-1PV, determined to aid the mapping of SIA site(s), is similar to that of other parvoviruses, especially MVM, except for differences at previously defined variable loops. H-1PV capsids screened on a glycan microarray revealed specificity for α2,3 linked SIA glycans, including the SIA-Le<sup>X</sup> tumor cell marker motif, similar to previous observations for MVM. The common recognition of the 3'SIA-Le<sup>X</sup> tumor cell marker by MVM and H-1PV explains their oncotropism. Structural studies of the H-1PV complexed with SIA glycans, including one with a 3'SIA-Le<sup>X</sup> motif, also identified the same twofold depression as MVM, as its SIA binding site. This observation indicates that parvovirus capsids utilize common regions as determinants of cell tropism, pathogenicity, and host range. Out of the four host range switch mutations *in vitro* in MVMP that conferred the ability to grow in rat fibroblasts (similar to H-1PV's cell

tropism), only two of those (residues 560 and 584 in H-1PV) are structurally variable between MVMP and H-1PV and are located at the twofold depression and might play a role in the virus adaptation to a new host. There are other residues that differ between these viruses at the SIA binding site and adopt different side chain conformations and might be host range determinants. The H-1PV capsid-glycan complex structures need to be completely refined and only preliminary conclusions can be made. The glycans on the H-1PV capsid are involved in non-specific interactions with homologous residues on MVMP that are also involved in the cell tropism and pathogenicity determination, and SIA binding for MVM. Especially, D405 (D399 in MVMP) seems to be an α2,3 linked SIA binding determinant on both the viruses. In a recent study, it was shown that the H-1PV capsid mutants, H374R and I368S, demonstrated much reduced cell binding as compared to the homologous MVMP mutants (10, 188). The residue 374 (H374 in H-1PV and K368 in MVMP), although a conservative mutation between the two capsids could contribute to the differences in the neuraminidase sensitivity or SIA binding affinity, similar to observations with the MVMP virulent mutants. The structural differences observed between MVM and H-1PV at the SIA receptor binding site might dictate the differences in oncotropic properties observed for these viruses. The H374R and I368S H-1PV mutant capsids (constructs available) could be produced in large quantities to re-screen on the glycan array and also tested in cell binding competition assays with the 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan to verify the glycan binding specificity of the mutant capsids. Lulli, another oncotropic rodent parvovirus has been shown to bind to 3'SIA(Le<sup>X</sup>)<sub>3</sub> glycan which highlights the common theme of utilization of this motif by these oncotropic viruses to bind to cancer cells. Structure of Lulli complexed to

3'SIA(Le<sup>X</sup>)<sub>3</sub> would provide structural verification for this common property amongst the rodent parvoviruses and provide information that would aid in elucidating the mechanism of differential tumor tropism between these viruses.

Ordered ssDNA in a conserved binding site inside the H-1PV virion suggested a potential role in genome packaging and capsid stability. The differences in the genome sequence observed in the H-1PV, CPV and MVM viruses could be attributed to the slight differences in the amino acid composition of the nucleotide binding pocket. Mutagenesis of the amino acids that are involved in unique interactions with the H-1PV genome could be done to investigate the significance of these capsid-DNA interactions in genome-specific packaging. Also, amino acids involved in conserved capsid-DNA interactions could be mutated (apart from the ones made in a previous study (247)) to understand their role in DNA packaging and capsid stability. To study the mechanism of DNA ordering, the structure of Lulli, a member of the parvovirus genus that packages minus strand and plus strand into different particles, could be determined. The information from this study would tell us if the structural ordering is due to a specific genome sequence of the parvovirus genus or because of a conserved DNA binding pocket. Also, the four cis-peptide bonds identified in the H-1PV virion structure might play a role in capsid assembly and this could be verified by mutating the virus at those positions and monitoring intact capsid production. The information on capsid regions involved in cellular specificity and genome interactions would aid in the development of oncotropic parvovirus based gene therapy vectors with improved cancer cell targeting and therapeutic genome packaging efficiency.

Structural studies of the MVMP capsids at the pHs encountered during endosomal trafficking provided insights into the capsid dynamics. Conformational changes were observed, especially at the icosahedral twofold regions, which alter capsid-DNA interaction(s) inside the capsid, possibly in preparation for DNA uncoating, and at the fivefold region, consistent with the opening of the channel, possibly for VP1 externalization. The structure of the MVMP-VLPs at pH 4.0 needs to be completely refined. Diffraction data for DNA containing infectious virions at low pHs has been collected and will be processed to provide structural verification for the disruption of capsid-DNA interactions at the twofold during endosomal trafficking prior to genome uncoating. Since it is not known whether receptor remains bound or dissociates from the capsid following endocytosis, diffraction data for MVMi-VLP and H-1PV crystals grown in the presence of glycan receptor at the various pHs have been collected and will be solved to compare any differences in the capsid dynamics for the receptor bound and unbound state. Time-resolved limited proteolysis with/without heat shock and site specific labeling in combination with peptide mass spectrometry has also been initiated to provide information on the most dynamic and accessible regions of the capsid.

## LIST OF REFERENCES

1. **Abschuetz, A., T. Kehl, R. Geibig, B. Leuchs, J. Rommelaere, and A. Regnier-Vigouroux.** 2006. Oncolytic murine autonomous parvovirus, a candidate vector for glioma gene therapy, is innocuous to normal and immunocompetent mouse glial cells. *Cell Tissue Res* **325**:423-436.
2. **Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**:213-221.
3. **Agbandje, M., S. Kajigaya, R. McKenna, N. S. Young, and M. G. Rossmann.** 1994. The structure of human parvovirus B19 at 8 Å resolution. *Virology* **203**:106-115.
4. **Agbandje, M., R. McKenna, M. G. Rossmann, M. L. Strassheim, and C. R. Parrish.** 1993. Structure determination of feline panleukopenia virus empty particles. *Proteins* **16**:155-171.
5. **Agbandje-McKenna, M., and M. S. Chapman.** 2006. Correlating structure with function in viral capsid, p. 125-139. In J. R. Kerr, S. F. Cotmore, M. E. Bloom, R. M. Linden, and C. R. Parrish (ed.), *Parvoviruses*. Hodder Arnold, London.
6. **Agbandje-McKenna, M., and J. Kleinschmidt.** 2011. AAV capsid structure and cell interactions. *Methods Mol Biol* **807**:47-92.
7. **Agbandje-McKenna, M., A. L. Llamas-Saiz, F. Wang, P. Tattersall, and M. G. Rossmann.** 1998. Functional implications of the structure of the murine parvovirus, minute virus of mice. *Structure* **6**:1369-1381.
8. **Akache, B., D. Grimm, K. Pandey, S. R. Yant, H. Xu, and M. A. Kay.** 2006. The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. *J Virol* **80**:9831-9836.
9. **Alexandersen, S., M. E. Bloom, and S. Perryman.** 1988. Detailed transcription map of Aleutian mink disease parvovirus. *J Virol* **62**:3684-3694.
10. **Allaume, X., N. El-Andaloussi, B. Leuchs, S. Bonifati, A. Kulkarni, T. Marttila, J. K. Kaufmann, D. M. Nettelbeck, J. Kleinschmidt, J. Rommelaere, and A. Marchini.** 2012. Retargeting of Rat Parvovirus H-1PV to Cancer Cells through Genetic Engineering of the Viral Capsid. *J Virol* **86**:3452-3465.
11. **Angata, T., and A. Varki.** 2002. Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev* **102**:439-469.

12. **Anouja, F., R. Wattiez, S. Mousset, and P. CailletFauquet.** 1997. The cytotoxicity of the parvovirus minute virus of mice nonstructural protein NS1 is related to changes in the synthesis and phosphorylation of cell proteins. *Journal of Virology* **71**:4671-4678.
13. **Asokan, A., J. B. Hamra, L. Govindasamy, M. Agbandje-McKenna, and R. J. Samulski.** 2006. Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry. *J Virol* **80**:8961-8969.
14. **Astell, C. R., M. Thomson, M. Merchlinsky, and D. C. Ward.** 1983. The complete DNA sequence of minute virus of mice, an autonomous parvovirus. *Nucleic Acids Res* **11**:999-1018.
15. **Ball-Goodrich, L. J., and P. Tattersall.** 1992. Two amino acid substitutions within the capsid are coordinately required for acquisition of fibrotropism by the lymphotropic strain of minute virus of mice. *J Virol* **66**:3415-3423.
16. **Bansal, G. P., J. A. Hatfield, F. E. Dunn, A. A. Kramer, F. Brady, C. H. Riggin, M. S. Collett, K. Yoshimoto, S. Kajigaya, and N. S. Young.** 1993. Candidate recombinant vaccine for human B19 parvovirus. *J Infect Dis* **167**:1034-1044.
17. **Bantel-Schaal, U., I. Braspenning-Wesch, and J. Kartenbeck.** 2009. Adeno-associated virus type 5 exploits two different entry pathways in human embryo fibroblasts. *J Gen Virol* **90**:317-322.
18. **Bantel-Schaal, U., B. Hub, and J. Kartenbeck.** 2002. Endocytosis of adeno-associated virus type 5 leads to accumulation of virus particles in the Golgi compartment. *J Virol* **76**:2340-2349.
19. **Barbis, D. P., S. F. Chang, and C. R. Parrish.** 1992. Mutations adjacent to the dimple of the canine parvovirus capsid structure affect sialic acid binding. *Virology* **191**:301-308.
20. **Bartlett, J. S., R. Wilcher, and R. J. Samulski.** 2000. Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol* **74**:2777-2785.
21. **Basak, S., and H. Turner.** 1992. Infectious entry pathway for canine parvovirus. *Virology* **186**:368-376.
22. **Bauer, P. H., R. T. Bronson, S. C. Fung, R. Freund, T. Stehle, S. C. Harrison, and T. L. Benjamin.** 1995. Genetic and structural analysis of a virulence determinant in polyomavirus VP1. *J Virol* **69**:7925-7931.
23. **Bell, C. L., L. H. Vandenberghe, P. Bell, M. P. Limberis, G. P. Gao, K. Van Vliet, M. Agbandje-McKenna, and J. M. Wilson.** 2011. The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice. *J Clin Invest* **121**:2427-2435.

24. **Benson, S. D., J. K. Bamford, D. H. Bamford, and R. M. Burnett.** 2004. Does common architecture reveal a viral lineage spanning all three domains of life? *Mol Cell* **16**:673-685.
25. **Bergeron, J., B. Hebert, and P. Tijssen.** 1996. Genome organization of the Kresse strain of porcine parvovirus: identification of the allotropic determinant and comparison with those of NADL-2 and field isolates. *J Virol* **70**:2508-2515.
26. **Bergoin, M., J. Kleinschmidt, J. M. Almendral, K. Hedman, Y. Li, M. Agbandje-McKenna, P. Tattersall, P. Tijssen, D. J. Pintel, and T. W. Flegel.** 2011. Ninth Report of the International Committee on Taxonomy of Viruses. In A. M. Q. King, Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (ed.), Virus Taxonomy. Elsevier.
27. **Berns, K., and C. R. Parrish.** 2007. Parvoviridae, p. 2437-2478. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, Fifth ed, vol. Two. Lippincott Williams and Wilkins.
28. **Blackburn, S. D., S. E. Cline, J. P. Hemming, and F. B. Johnson.** 2005. Attachment of bovine parvovirus to O-linked alpha 2,3 neuraminic acid on glycophorin A. *Arch Virol* **150**:1477-1484.
29. **Blackburn, S. D., R. A. Steadman, and F. B. Johnson.** 2006. Attachment of adeno-associated virus type 3H to fibroblast growth factor receptor 1. *Arch Virol* **151**:617-623.
30. **Blechacz, B., and S. J. Russell.** 2004. Parvovirus vectors: use and optimisation in cancer gene therapy. *Expert Rev Mol Med* **6**:1-24.
31. **Bleker, S., M. Pawlita, and J. A. Kleinschmidt.** 2006. Impact of capsid conformation and Rep-capsid interactions on adeno-associated virus type 2 genome packaging. *J Virol* **80**:810-820.
32. **Bleker, S., F. Sonntag, and J. A. Kleinschmidt.** 2005. Mutational analysis of narrow pores at the fivefold symmetry axes of adeno-associated virus type 2 capsids reveals a dual role in genome packaging and activation of phospholipase A2 activity. *J Virol* **79**:2528-2540.
33. **Bodendorf, U., C. Cziepluch, J. C. Jauniaux, J. Rommelaere, and N. Salomé.** 1999. Nuclear export factor CRM1 interacts with nonstructural proteins NS2 from parvovirus minute virus of mice. *J Virol* **73**:7769-7779.
34. **Boisvert, M., S. Fernandes, and P. Tijssen.** 2010. Multiple pathways involved in porcine parvovirus cellular entry and trafficking toward the nucleus. *J Virol* **84**:7782-7792.

35. **Bonsch, C., C. Zuercher, P. Lieby, C. Kempf, and C. Ros.** 2010. The globoside receptor triggers structural changes in the B19 virus capsid that facilitate virus internalization. *J Virol* **84**:11737-11746.
36. **Bothner, B., X. F. Dong, L. Bibbs, J. E. Johnson, and G. Siuzdak.** 1998. Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. *J Biol Chem* **273**:673-676.
37. **Bowles, D. E., S. W. McPhee, C. Li, S. J. Gray, J. J. Samulski, A. S. Camp, J. Li, B. Wang, P. E. Monahan, J. E. Rabinowitz, J. C. Grieger, L. Govindasamy, M. Agbandje-McKenna, X. Xiao, and R. J. Samulski.** 2012. Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Mol Ther* **20**:443-455.
38. **Bowles, D. E., J. E. Rabinowitz, and S. R.J.** 2006. The Genus Dependovirus, p. 15-24. *In* J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), *Parvoviruses*. Hodder Arnold, London.
39. **Brandenburger, A., D. Legendre, B. Avalosse, and J. Rommelaere.** 1990. NS-1 and NS-2 proteins may act synergistically in the cytopathogenicity of parvovirus MVMP. *Virology* **174**:576-584.
40. **Brandts, J. F., H. R. Halvorson, and M. Brennan.** 1975. Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**:4953-4963.
41. **Brockhausen, I., H. Schachter, and P. Stanley.** 2009. O-GalNAc Glycans. *In* A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, and M. E. Etzler (ed.), *Essentials of Glycobiology*. The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor NY.
42. **Brown, C. S., S. Welling-Wester, M. Feijlbrief, J. W. Van Lent, and W. J. Spaan.** 1994. Chimeric parvovirus B19 capsids for the presentation of foreign epitopes. *Virology* **198**:477-488.
43. **Brown, K. E., S. M. Anderson, and N. S. Young.** 1993. Erythrocyte P antigen: cellular receptor for B19 parvovirus. *Science* **262**:114-117.
44. **Brownstein, D. G., A. L. Smith, E. A. Johnson, D. J. Pintel, L. K. Naeger, and P. Tattersall.** 1992. The pathogenesis of infection with minute virus of mice depends on expression of the small nonstructural protein NS2 and on the genotype of the allotropic determinants VP1 and VP2. *J Virol* **66**:3118-3124.
45. **Brunger, A. T.** 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**:472-475.

46. **Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grossé-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren.** 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**:905-921.
47. **Burmeister, W. P., D. Guilligay, S. Cusack, G. Wadell, and N. Arnberg.** 2004. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J Virol* **78**:7727-7736.
48. **Caillet-Fauquet, P., M. Perros, A. Brandenburger, P. Spegelaere, and J. Rommelaere.** 1990. Programmed killing of human cells by means of an inducible clone of parvoviral genes encoding non-structural proteins. *EMBO J* **9**:2989-2995.
49. **Carrasco, C., A. Carreira, I. A. Schaap, P. A. Serena, J. Gomez-Herrero, M. G. Mateu, and P. J. de Pablo.** 2006. DNA-mediated anisotropic mechanical reinforcement of a virus. *Proc Natl Acad Sci U S A* **103**:13706-13711.
50. **Carrillo-Tripp, M., C. M. Shepherd, I. A. Borelli, S. Venkataraman, G. Lander, P. Natarajan, J. E. Johnson, C. L. Brooks, 3rd, and V. S. Reddy.** 2009. VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res* **37**:D436-442.
51. **Cater, J. E., and D. J. Pintel.** 1992. The small non-structural protein NS2 of the autonomous parvovirus minute virus of mice is required for virus growth in murine cells. *J Gen Virol* **73 ( Pt 7)**:1839-1843.
52. **Chang, S. F., J. Y. Sgro, and C. R. Parrish.** 1992. Multiple amino acids in the capsid structure of canine parvovirus coordinately determine the canine host range and specific antigenic and hemagglutination properties. *J Virol* **66**:6858-6867.
53. **Chapman, M. S., and M. Agbandje-McKenna.** 2006. Atomic structure of viral particles, p. 107-123. In J. R. Kerr, S. F. Cotmore, M. E. Bloom, R. M. Linden, and C. R. Parrish (ed.), *Parvoviruses*. Hodder Arnold, London.
54. **Chejanovsky, N., and B. J. Carter.** 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. *Virology* **173**:120-128.
55. **Chen, K. C., B. C. Shull, M. Lederman, E. R. Stout, and R. C. Bates.** 1988. Analysis of the termini of the DNA of bovine parvovirus: demonstration of sequence inversion at the left terminus and its implication for the replication model. *J Virol* **62**:3807-3813.
56. **Chen, V. B., W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson.** 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**:12-21.

57. **Chipman, P. R., M. Agbandje-McKenna, S. Kajigaya, K. E. Brown, N. S. Young, T. S. Baker, and M. G. Rossmann.** 1996. Cryo-electron microscopy studies of empty capsids of human parvovirus B19 complexed with its cellular receptor. *Proc Natl Acad Sci U S A* **93**:7502-7506.
58. **Clemens, K. E., and D. J. Pintel.** 1988. The two transcription units of the autonomous parvovirus minute virus of mice are transcribed in a temporal order. *J Virol* **62**:1448-1451.
59. **Cohen, S., A. R. Behzad, J. B. Carroll, and N. Pante.** 2006. Parvoviral nuclear import: bypassing the host nuclear-transport machinery. *J Gen Virol* **87**:3209-3213.
60. **Cohen, S., A. K. Marr, P. Garcin, and N. Pante.** 2011. Nuclear envelope disruption involving host caspases plays a role in the parvovirus replication cycle. *J Virol* **85**:4863-4874.
61. **Cohen, S., and N. Pante.** 2005. Pushing the envelope: microinjection of Minute virus of mice into Xenopus oocytes causes damage to the nuclear envelope. *J Gen Virol* **86**:3243-3252.
62. **Collaborative Computational Project, N.** 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**:760-763.
63. **Colomar, M. C., B. Hirt, and P. Beard.** 1998. Two segments in the genome of the immunosuppressive minute virus of mice determine the host-cell specificity, control viral DNA replication and affect viral RNA metabolism. *J Gen Virol* **79** ( Pt 3):581-586.
64. **Cornelis, J. J., P. Becquart, N. Duponchel, N. Salome, B. L. Avalosse, M. Namba, and J. Rommelaere.** 1988. Transformation of human fibroblasts by ionizing radiation, a chemical carcinogen, or simian virus 40 correlates with an increase in susceptibility to the autonomous parvoviruses H-1 virus and minute virus of mice. *J Virol* **62**:1679-1686.
65. **Cotmore, S. F., M. D'Abramo A, Jr., C. M. Ticknor, and P. Tattersall.** 1999. Controlled conformational transitions in the MVM virion expose the VP1 N-terminus and viral genome without particle disassembly. *Virology* **254**:169-181.
66. **Cotmore, S. F., A. M. D'Abramo, Jr., L. F. Carbonell, J. Bratton, and P. Tattersall.** 1997. The NS2 polypeptide of parvovirus MVM is required for capsid assembly in murine cells. *Virology* **231**:267-280.
67. **Cotmore, S. F., S. Hafenstein, and P. Tattersall.** 2010. Depletion of virion-associated divalent cations induces parvovirus minute virus of mice to eject its genome in a 3'-to-5' direction from an otherwise intact viral particle. *J Virol* **84**:1945-1956.

68. **Cotmore, S. F., V. C. McKie, L. J. Anderson, C. R. Astell, and P. Tattersall.** 1986. Identification of the major structural and nonstructural proteins encoded by human parvovirus B19 and mapping of their genes by prokaryotic expression of isolated genomic fragments. *J Virol* **60**:548-557.
69. **Cotmore, S. F., and P. Tattersall.** 1989. A genome-linked copy of the NS-1 polypeptide is located on the outside of infectious parvovirus particles. *J Virol* **63**:3902-3911.
70. **Cotmore, S. F., and P. Tattersall.** 2006. A Rolling-hairpin Strategy: Basic Mechanism of DNA Replication in the Parvoviruses, p. 171-188. *In* J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), *Parvoviruses*. Hodder Arnold, London.
71. **Cotmore, S. F., and P. Tattersall.** 1990. Alternate splicing in a parvoviral nonstructural gene links a common amino-terminal sequence to downstream domains which confer radically different localization and turnover characteristics. *Virology* **177**:477-487.
72. **Cotmore, S. F., and P. Tattersall.** 2005. Encapsidation of minute virus of mice DNA: aspects of the translocation mechanism revealed by the structure of partially packaged genomes. *Virology* **336**:100-112.
73. **Cotmore, S. F., and P. Tattersall.** 2005. Genome packaging sense is controlled by the efficiency of the nick site in the right-end replication origin of parvoviruses minute virus of mice and Lulli. *J Virol* **79**:2287-2300.
74. **Cotmore, S. F., and P. Tattersall.** 2012. Mutations at the Base of the Icosahedral Five-Fold Cylinders of Minute Virus of Mice Induce 3'-to-5' Genome Uncoating and Critically Impair Entry Functions. *J Virol* **86**:69-80.
75. **Cotmore, S. F., and P. Tattersall.** 2007. Parvoviral host range and cell entry mechanisms. *Adv Virus Res* **70**:183-232.
76. **Cotmore, S. F., and P. Tattersall.** 2006. Structure and Organization of the Viral Genome, p. 73-94. *In* J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), *Parvoviruses*. Hodder Arnold, London.
77. **Cotmore, S. F., and P. Tattersall.** 1987. The autonomously replicating parvoviruses of vertebrates. *Advances in Virus Research* **33**:91-174.
78. **Cotmore, S. F., and P. Tattersall.** 1987. The autonomously replicating parvoviruses of vertebrates. *Adv Virus Res* **33**:91-174.
79. **D'Abramo, A. M., Jr., A. A. Ali, F. Wang, S. F. Cotmore, and P. Tattersall.** 2005. Host range mutants of Minute Virus of Mice with a single VP2 amino acid change require additional silent mutations that regulate NS2 accumulation. *Virology* **340**:143-154.

80. **Daeffler, L., R. Horlein, J. Rommelaere, and J. Nuesch.** 2003. Modulation of minute virus of mice cytotoxic activities through site-directed mutagenesis within the NS coding region. *Journal of Virology* **77**:12466-12478.
81. **de la Roche Saint Andre, C., F. Harper, and J. Feunteun.** 1990. Analysis of the hamster polyomavirus infection in vitro: host-restricted productive cycle. *Virology* **177**:532-540.
82. **DeLano, W. L.** 2002. The PyMOL Molecular Graphic System. DeLano Scientific, San Carlos, CA.
83. **Di Pasquale, G., and J. A. Chiorini.** 2006. AAV transcytosis through barrier epithelia and endothelium. *Mol Ther* **13**:506-516.
84. **Di Pasquale, G., B. L. Davidson, C. S. Stein, I. Martins, D. Scudiero, A. Monks, and J. A. Chiorini.** 2003. Identification of PDGFR as a receptor for AAV-5 transduction. *Nat Med* **9**:1306-1312.
85. **Di Pasquale, G., N. Kaludov, M. Agbandje-McKenna, and J. A. Chiorini.** 2010. BAAV transcytosis requires an interaction with beta-1-4 linked- glucosamine and gp96. *PLoS One* **5**:e9336.
86. **Ding, W., L. Zhang, Z. Yan, and J. F. Engelhardt.** 2005. Intracellular trafficking of adeno-associated viral vectors. *Gene Ther* **12**:873-880.
87. **Ding, W., L. N. Zhang, C. Yeaman, and J. F. Engelhardt.** 2006. rAAV2 traffics through both the late and the recycling endosomes in a dose-dependent fashion. *Mol Ther* **13**:671-682.
88. **DiPrimio, N., A. Asokan, L. Govindasamy, M. Agbandje-McKenna, and R. J. Samulski.** 2008. Surface loop dynamics in adeno-associated virus capsid assembly. *J Virol* **82**:5178-5189.
89. **Doerig, C., B. Hirt, P. Beard, and J. P. Antonietti.** 1988. Minute virus of mice non-structural protein NS-1 is necessary and sufficient for trans-activation of the viral P39 promoter. *J Gen Virol* **69** ( Pt 10):2563-2573.
90. **Dormitzer, P. R., Z. Y. Sun, G. Wagner, and S. C. Harrison.** 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J* **21**:885-897.
91. **Dorsch, S., G. Liebisch, B. Kaufmann, P. von Landenberg, J. H. Hoffmann, W. Drobnik, and S. Modrow.** 2002. The VP1 unique region of parvovirus B19 and its constituent phospholipase A2-like activity. *J Virol* **76**:2014-2018.
92. **Douar, A. M., K. Poulard, D. Stockholm, and O. Danos.** 2001. Intracellular trafficking of adeno-associated virus vectors: routing to the late endosomal compartment and proteasome degradation. *J Virol* **75**:1824-1833.

93. **Duan, D., Q. Li, A. W. Kao, Y. Yue, J. E. Pessin, and J. F. Engelhardt.** 1999. Dynamin is required for recombinant adeno-associated virus type 2 infection. *J Virol* **73**:10371-10376.
94. **Duan, D., Y. Yue, Z. Yan, J. Yang, and J. F. Engelhardt.** 2000. Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus. *J Clin Invest* **105**:1573-1587.
95. **Dubielzig, R., J. A. King, S. Weger, A. Kern, and J. A. Kleinschmidt.** 1999. Adeno-associated virus type 2 protein interactions: formation of pre-encapsidation complexes. *J Virol* **73**:8989-8998.
96. **Eichwald, V., L. Daeffler, M. Klein, J. Rommelaere, and N. Salome.** 2002. The NS2 proteins of parvovirus minute virus of mice are required for efficient nuclear egress of progeny virions in mouse cells. *J Virol* **76**:10307-10319.
97. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**:2126-2132.
98. **Enderlin, M., E. V. Kleinmann, S. Struyf, C. Buracchi, A. Vecchi, R. Kinscherf, F. Kiessling, S. Paschek, S. Sozzani, J. Rommelaere, J. J. Cornelis, J. Van Damme, and C. Dinsart.** 2009. TNF-alpha and the IFN-gamma-inducible protein 10 (IP-10/CXCL-10) delivered by parvoviral vectors act in synergy to induce antitumor effects in mouse glioblastoma. *Cancer Gene Ther* **16**:149-160.
99. **Etingov, I., R. Itah, M. Minberg, A. Keren-Naus, H. J. Nam, M. Agbandje-McKenna, and C. Davis.** 2008. An extension of the Minute Virus of Mice tissue tropism. *Virology* **379**:245-255.
100. **Farr, G. A., S. F. Cotmore, and P. Tattersall.** 2006. VP2 cleavage and the leucine ring at the base of the fivefold cylinder control pH-dependent externalization of both the VP1 N terminus and the genome of minute virus of mice. *J Virol* **80**:161-171.
101. **Farr, G. A., and P. Tattersall.** 2004. A conserved leucine that constricts the pore through the capsid fivefold cylinder plays a central role in parvoviral infection. *Virology* **323**:243-256.
102. **Farr, G. A., L. G. Zhang, and P. Tattersall.** 2005. Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry. *Proc Natl Acad Sci U S A* **102**:17148-17153.
103. **Feschotte, C., and C. Gilbert.** 2012. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* **13**:283-296.
104. **Fisher, A. J., and J. E. Johnson.** 1993. Ordered duplex RNA controls capsid architecture in an icosahedral animal virus. *Nature* **361**:176-179.

105. **Fox, D., X. He, A. Abe, T. Hollander, L. Li, L. Kan, A. Friedman, Y. Shimizu, J. Shayman, and K. Kozarsky.** 2001. The T lymphocyte structure CD60 contains a sialylated carbohydrate epitope that is expressed on both gangliosides and glycoproteins. *Immunological Investigations* **30**:67-85.
106. **Fox, J. M., and M. E. Bloom.** 1999. Identification of a cell surface protein from Crandell feline kidney cells that specifically binds Aleutian mink disease parvovirus. *J Virol* **73**:3835-3842.
107. **Fox, J. P.** 1984. Modes of action of poliovirus vaccines and relation to resulting immunity. *Rev Infect Dis* **6 Suppl 2**:S352-355.
108. **Fricks, C. E., and J. M. Hogle.** 1990. Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. *J Virol* **64**:1934-1945.
109. **Fry, E. E., S. M. Lea, T. Jackson, J. W. Newman, F. M. Ellard, W. E. Blakemore, R. Abu-Ghazaleh, A. Samuel, A. M. King, and D. I. Stuart.** 1999. The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex. *EMBO J* **18**:543-554.
110. **Fuks, F., L. Deleu, C. Dinsart, J. Rommelaere, and S. Faisst.** 1996. ras oncogene-dependent activation of the P4 promoter of minute virus of mice through a proximal P4 element interacting with the Ets family of transcription factors. *J Virol* **70**:1331-1339.
111. **Gamblin, S. J., L. F. Haire, R. J. Russell, D. J. Stevens, B. Xiao, Y. Ha, N. Vasisht, D. A. Steinhauer, R. S. Daniels, A. Elliot, D. C. Wiley, and J. J. Skehel.** 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* **303**:1838-1842.
112. **Gardiner, E. M., and P. Tattersall.** 1988. Mapping of the fibrotropic and lymphotropic host range determinants of the parvovirus minute virus of mice. *J Virol* **62**:2605-2613.
113. **Geletneky, K., J. Huesing, J. Rommelaere, J. R. Schlehofer, M. Dahm, O. Krebs, M. v. Knebel Doeberitz, B. Huber, and J. Hajda.** 2012. Phase I/Ila study of intratumoral/intracerebral or intravenous/intracerebral administration of Parvovirus H-1 (ParvOryx) in patients with progressive primary or recurrent glioblastoma multiforme: ParvOryx01 protocol. *BMC Cancer* **12**:99.
114. **Geletneky, K., I. Kiprianova, A. Ayache, R. Koch, Y. C. M. Herrero, L. Deleu, C. Sommer, N. Thomas, J. Rommelaere, and J. R. Schlehofer.** 2010. Regression of advanced rat and human gliomas by local or systemic treatment with oncolytic parvovirus H-1 in rat models. *Neuro Oncol* **12**:804-814.

115. **Girod, A., C. E. Wobus, Z. Zadori, M. Ried, K. Leike, P. Tijssen, J. A. Kleinschmidt, and M. Hallek.** 2002. The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J Gen Virol* **83**:973-978.
116. **Goodman, L. B., S. M. Lyi, N. C. Johnson, J. O. Cifuentes, S. L. Hafenstein, and C. R. Parrish.** 2010. Binding site on the transferrin receptor for the parvovirus capsid and effects of altered affinity on cell uptake and infection. *J Virol* **84**:4969-4978.
117. **Goodman, L. B., S. M. Lyi, N. C. Johnson, J. O. Cifuentes, S. L. Hafenstein, and C. R. Parrish.** Binding site on the transferrin receptor for the parvovirus capsid and effects of altered affinity on cell uptake and infection. *J Virol* **84**:4969-4978.
118. **Govindasamy, L., K. Hueffer, C. R. Parrish, and M. Agbandje-McKenna.** 2003. Structures of host range-controlling regions of the capsids of canine and feline parvoviruses and mutants. *J Virol* **77**:12211-12221.
119. **Govindasamy, L., E. Padron, R. McKenna, N. Muzychka, N. Kaludov, J. A. Chiorini, and M. Agbandje-McKenna.** 2006. Structurally mapping the diverse phenotype of adeno-associated virus serotype 4. *J Virol* **80**:11556-11570.
120. **Grekova, S., R. Zawatzky, R. Horlein, C. Cziepluch, M. Mincberg, C. Davis, J. Rommelaere, and L. Daeffler.** 2010. Activation of an antiviral response in normal but not transformed mouse cells: a new determinant of minute virus of mice oncotropicism. *J Virol* **84**:516-531.
121. **Grieger, J. C., J. S. Johnson, B. Gurda-Whitaker, M. Agbandje-McKenna, and R. J. Samulski.** 2007. Surface-exposed adeno-associated virus Vp1-NLS capsid fusion protein rescues infectivity of noninfectious wild-type Vp2/Vp3 and Vp3-only capsids but not that of fivefold pore mutant virions. *J Virol* **81**:7833-7843.
122. **Grieger, J. C., S. Snowdy, and R. J. Samulski.** 2006. Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. *J Virol* **80**:5199-5210.
123. **Gurda, B. L., K. N. Parent, H. Bladek, R. S. Sinkovits, M. A. DiMatta, C. Rence, A. Castro, R. McKenna, N. Olson, K. Brown, T. S. Baker, and M. Agbandje-McKenna.** 2010. Human bocavirus capsid structure: insights into the structural repertoire of the parvoviridae. *J Virol* **84**:5880-5889.
124. **Haag, A., P. Menten, J. Van Damme, C. Dinsart, J. Rommelaere, and J. J. Cornelis.** 2000. Highly efficient transduction and expression of cytokine genes in human tumor cells by means of autonomous parvovirus vectors; generation of antitumor responses in recipient mice. *Hum Gene Ther* **11**:597-609.

125. **Hafenstein, S., V. D. Bowman, T. Sun, C. D. Nelson, L. M. Palermo, P. R. Chipman, A. J. Battisti, C. R. Parrish, and M. G. Rossmann.** 2009. Structural comparison of different antibodies interacting with parvovirus capsids. *J Virol* **83**:5556-5566.
126. **Halder, S., R. Ng, and M. Agbandje-McKenna.** 2012. Parvoviruses: structure and infection. *Future Virology* **7**:253-278.
127. **Hansen, J., K. Qing, H. J. Kwon, C. Mah, and A. Srivastava.** 2000. Impaired intracellular trafficking of adeno-associated virus type 2 vectors limits efficient transduction of murine fibroblasts. *J Virol* **74**:992-996.
128. **Hansen, J., K. Qing, and A. Srivastava.** 2001. Adeno-associated virus type 2-mediated gene transfer: altered endocytic processing enhances transduction efficiency in murine fibroblasts. *J Virol* **75**:4080-4090.
129. **Harbison, C. E., J. A. Chiorini, and C. R. Parrish.** 2008. The parvovirus capsid odyssey: from the cell surface to the nucleus. *Trends Microbiol* **16**:208-214.
130. **Harbison, C. E., S. M. Lyi, W. S. Weichert, and C. R. Parrish.** 2009. Early steps in cell infection by parvoviruses: host-specific differences in cell receptor binding but similar endosomal trafficking. *J Virol* **83**:10504-10514.
131. **Hauck, B., W. Zhao, K. High, and W. Xiao.** 2004. Intracellular viral processing, not single-stranded DNA accumulation, is crucial for recombinant adeno-associated virus transduction. *J Virol* **78**:13678-13686.
132. **Hernando, E., A. L. Llamas-Saiz, C. Foces-Foces, R. McKenna, I. Portman, M. Agbandje-McKenna, and J. M. Almendral.** 2000. Biochemical and physical characterization of parvovirus minute virus of mice virus-like particles. *Virology* **267**:299-309.
133. **Heuser, J.** 1989. Changes in lysosome shape and distribution correlated with changes in cytoplasmic pH. *J Cell Biol* **108**:855-864.
134. **Horiuchi, M., N. Ishiguro, H. Goto, and M. Shinagawa.** 1992. Characterization of the stage(s) in the virus replication cycle at which the host-cell specificity of the feline parvovirus subgroup is regulated in canine cells. *Virology* **189**:600-608.
135. **Hueffer, K., L. Govindasamy, M. Agbandje-McKenna, and C. R. Parrish.** 2003. Combinations of two capsid regions controlling canine host range determine canine transferrin receptor binding by canine and feline parvoviruses. *J Virol* **77**:10099-10105.
136. **Hueffer, K., J. S. Parker, W. S. Weichert, R. E. Geisel, J. Y. Sgro, and C. R. Parrish.** 2003. The natural host range shift and subsequent evolution of canine parvovirus resulted from virus-specific binding to the canine transferrin receptor. *J Virol* **77**:1718-1726.

137. **Hueffer, K., and C. R. Parrish.** 2003. Parvovirus host range, cell tropism and evolution. *Curr Opin Microbiol* **6**:392-398.
138. **Inoko, E., Y. Nishiura, H. Tanaka, T. Takahashi, K. Furukawa, K. Kitajima, and C. Sato.** 2010. Developmental stage-dependent expression of an alpha2,8-trisialic acid unit on glycoproteins in mouse brain. *Glycobiology* **20**:916-928.
139. **Irimura, T., S. Nakamori, Y. Matsushita, Y. Taniuchi, N. Todoroki, T. Tsuji, Y. Izumi, Y. Kawamura, S. D. Hoff, K. R. Cleary, and et al.** 1993. Colorectal cancer metastasis determined by carbohydrate-mediated cell adhesion: role of sialyl-LeX antigens. *Semin Cancer Biol* **4**:319-324.
140. **Jabs, A., M. S. Weiss, and R. Hilgenfeld.** 1999. Non-proline cis peptide bonds in proteins. *J Mol Biol* **286**:291-304.
141. **Jnaoui, K., M. Minet, and T. Michiels.** 2002. Mutations that affect the tropism of DA and GDVII strains of Theiler's virus in vitro influence sialic acid binding and pathogenicity. *J Virol* **76**:8138-8147.
142. **Johnson, F. B., L. B. Fenn, T. J. Owens, L. J. Faucheuex, and S. D. Blackburn.** 2004. Attachment of bovine parvovirus to sialic acids on bovine cell membranes. *J Gen Virol* **85**:2199-2207.
143. **Johnson, F. B., and M. D. Hoggan.** 1973. Structural proteins of HADEN virus. *Virology* **51**:129-137.
144. **Johnson, J. S., M. Gentzsch, L. Zhang, C. M. Ribeiro, B. Kantor, T. Kafri, R. J. Pickles, and R. J. Samulski.** 2011. AAV exploits subcellular stress associated with inflammation, endoplasmic reticulum expansion, and misfolded proteins in models of cystic fibrosis. *PLoS Pathog* **7**:e1002053.
145. **Johnson, J. S., C. Li, N. DiPrimio, M. S. Weinberg, T. J. McCown, and R. J. Samulski.** 2010. Mutagenesis of adeno-associated virus type 2 capsid protein VP1 uncovers new roles for basic amino acids in trafficking and cell-specific transduction. *J Virol* **84**:8888-8902.
146. **Johnson, J. S., and R. J. Samulski.** 2009. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J Virol* **83**:2632-2644.
147. **Johnston, J. B., S. H. Nazarian, R. Natale, and G. McFadden.** 2005. Myxoma virus infection of primary human fibroblasts varies with cellular age and is regulated by host interferon responses, p. 235-248, *Virology*, vol. 332, United States.

148. **Kaludov, N., K. E. Brown, R. W. Walters, J. Zabner, and J. A. Chiorini.** 2001. Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol* **75**:6884-6893.
149. **Kamerling, J. P., and G. J. Gerwig.** 2006. Structural analysis of naturally occurring sialic acids. *Methods Mol Biol* **347**:69-91.
150. **Kaminsky, P. M., N. W. Keiser, Z. Yan, D. C. Lei-Butters, and J. F. Engelhardt.** 2012. Directing Integrin-linked Endocytosis of Recombinant AAV Enhances Productive FAK-dependent Transduction. *Mol Ther*.
151. **Kannagi, R.** 2004. Molecular mechanism for cancer-associated induction of sialyl Lewis X and sialyl Lewis A expression-The Warburg effect revisited. *Glycoconj J* **20**:353-364.
152. **Kannagi, R., M. Izawa, T. Koike, K. Miyazaki, and N. Kimura.** 2004. Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci* **95**:377-384.
153. **Kashiwakura, Y., K. Tamayose, K. Iwabuchi, Y. Hirai, T. Shimada, K. Matsumoto, T. Nakamura, M. Watanabe, K. Oshimi, and H. Daida.** 2005. Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J Virol* **79**:609-614.
154. **Kaufmann, B., U. Baxa, P. R. Chipman, M. G. Rossmann, S. Modrow, and R. Seckler.** 2005. Parvovirus B19 does not bind to membrane-associated globoside in vitro. *Virology* **332**:189-198.
155. **Kaufmann, B., P. R. Chipman, V. A. Kostyuchenko, S. Modrow, and M. G. Rossmann.** 2008. Visualization of the externalized VP2 N termini of infectious human parvovirus B19. *J Virol* **82**:7306-7312.
156. **Kaufmann, B., M. El-Far, P. Plevka, V. D. Bowman, Y. Li, P. Tijssen, and M. G. Rossmann.** 2011. Structure of *Bombyx mori* densovirus 1, a silkworm pathogen. *J Virol* **85**:4691-4697.
157. **Kaufmann, B., A. Lopez-Bueno, M. G. Mateu, P. R. Chipman, C. D. Nelson, C. R. Parrish, J. M. Almendral, and M. G. Rossmann.** 2007. Minute virus of mice, a parvovirus, in complex with the Fab fragment of a neutralizing monoclonal antibody. *J Virol* **81**:9851-9858.
158. **Kaufmann, B., A. A. Simpson, and M. G. Rossmann.** 2004. The structure of human parvovirus B19. *Proc Natl Acad Sci U S A* **101**:11628-11633.

159. **Kelly, B. N., S. Kyere, I. Kinde, C. Tang, B. R. Howard, H. Robinson, W. I. Sundquist, M. F. Summers, and C. P. Hill.** 2007. Structure of the antiviral assembly inhibitor CAP-1 complex with the HIV-1 CA protein. *J Mol Biol* **373**:355-366.
160. **Kern, A., K. Schmidt, C. Leder, O. J. Muller, C. E. Wobus, K. Bettinger, C. W. Von der Lieth, J. A. King, and J. A. Kleinschmidt.** 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J Virol* **77**:11072-11081.
161. **Kestler, J., B. Neeb, S. Struyf, J. Van Damme, S. F. Cotmore, A. D'Abramo, P. Tattersall, J. Rommelaere, C. Dinsart, and J. J. Cornelis.** 1999. cis requirements for the efficient production of recombinant DNA vectors based on autonomous parvoviruses. *Hum Gene Ther* **10**:1619-1632.
162. **Khayat, R., and J. E. Johnson.** 2011. Pass the jelly rolls. *Structure* **19**:904-906.
163. **Kimsey, P. B., H. D. Engers, B. Hirt, and C. V. Jongeneel.** 1986. Pathogenicity of fibroblast- and lymphocyte-specific variants of minute virus of mice. *J Virol* **59**:8-13.
164. **King, J. A., R. Dubielzig, D. Grimm, and J. A. Kleinschmidt.** 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20**:3282-3291.
165. **Klein, A., S. Diaz, I. Ferreira, G. Lamblin, P. Roussel, and A. E. Manzi.** 1997. New sialic acids from biological sources identified by a comprehensive and sensitive approach: liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) of SIA quinoxalinones. *Glycobiology* **7**:421-432.
166. **Kleywegt, G. J.** 2007. Crystallographic refinement of ligand complexes. *Acta Crystallogr D Biol Crystallogr* **63**:94-100.
167. **Kleywegt, G. J., and T. A. Jones.** 1997. Model building and refinement practice. *Methods Enzymol* **277**:208-230.
168. **Kontou, M., L. Govindasamy, H. J. Nam, N. Bryant, A. L. Llamas-Saiz, C. Foces-Foces, E. Hernando, M. P. Rubio, R. McKenna, J. M. Almendral, and M. Agbandje-McKenna.** 2005. Structural determinants of tissue tropism and in vivo pathogenicity for the parvovirus minute virus of mice. *J Virol* **79**:10931-10943.
169. **Krissinel, E., and K. Henrick.** 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**:2256-2268.
170. **Kronenberg, S., B. Bottcher, C. W. von der Lieth, S. Bleker, and J. A. Kleinschmidt.** 2005. A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J Virol* **79**:5296-5303.

171. **Kronenberg, S., J. A. Kleinschmidt, and B. Bottcher.** 2001. Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids. *EMBO Rep* **2**:997-1002.
172. **Kruger, L., H. Eskerski, C. Dinsart, J. Cornelis, J. Rommelaere, U. Haberkorn, and J. A. Kleinschmidt.** 2008. Augmented transgene expression in transformed cells using a parvoviral hybrid vector. *Cancer Gene Ther* **15**:252-267.
173. **Lacroix, J., B. Leuchs, J. Li, G. Hristov, H. E. Deubzer, A. E. Kulozik, J. Rommelaere, J. R. Schlehofer, and O. Witt.** 2010. Parvovirus H1 selectively induces cytotoxic effects on human neuroblastoma cells. *Int J Cancer* **127**:1230-1239.
174. **Lang, S. I., S. Boelz, A. Y. Stroh-Dege, J. Rommelaere, C. Dinsart, and J. J. Cornelis.** 2005. The infectivity and lytic activity of minute virus of mice wild-type and derived vector particles are strikingly different. *J Virol* **79**:289-298.
175. **Lederman, M., R. C. Bates, and E. R. Stout.** 1983. In vitro and in vivo studies of bovine parvovirus proteins. *J Virol* **48**:10-17.
176. **Legrand, C., J. Rommelaere, and P. Caillet-Fauquet.** 1993. MVM(p) NS-2 protein expression is required with NS-1 for maximal cytotoxicity in human transformed cells. *Virology* **195**:149-155.
177. **Lerch, T. F., Q. Xie, and M. S. Chapman.** 2010. The structure of adeno-associated virus serotype 3B (AAV-3B): insights into receptor binding and immune evasion. *Virology* **403**:26-36.
178. **Levy, H. C., V. D. Bowman, L. Govindasamy, R. McKenna, K. Nash, K. Warrington, W. Chen, N. Muzyczka, X. Yan, T. S. Baker, and M. Agbandje-McKenna.** 2009. Heparin binding induces conformational changes in Adeno-associated virus serotype 2. *J Struct Biol* **165**:146-156.
179. **Li, Q., A. G. Yafal, Y. M. Lee, J. Hogle, and M. Chow.** 1994. Poliovirus neutralization by antibodies to internal epitopes of VP4 and VP1 results from reversible exposure of these sequences at physiological temperature. *J Virol* **68**:3965-3970.
180. **Ling, C., Y. Lu, J. K. Kalsi, G. R. Jayandharan, B. Li, W. Ma, B. Cheng, S. W. Gee, K. E. McGoogan, L. Govindasamy, L. Zhong, M. Agbandje-McKenna, and A. Srivastava.** 2010. Human hepatocyte growth factor receptor is a cellular coreceptor for adeno-associated virus serotype 3. *Hum Gene Ther* **21**:1741-1747.
181. **Linser, P., H. Bruning, and R. W. Armentrout.** 1977. Specific binding sites for a parvovirus, minute virus of mice, on cultured mouse cells. *J Virol* **24**:211-221.

182. **Liu, F., H. L. Qi, and H. L. Chen.** 2001. Regulation of differentiation- and proliferation-inducers on Lewis antigens, alpha-fucosyltransferase and metastatic potential in hepatocarcinoma cells. *Br J Cancer* **84**:1556-1563.
183. **Liu, J., and S. C. Thorp.** 2002. Cell surface heparan sulfate and its roles in assisting viral infections. *Med Res Rev* **22**:1-25.
184. **Llamas-Saiz, A. L., M. Agbandje-McKenna, W. R. Wikoff, J. Bratton, P. Tattersall, and M. G. Rossmann.** 1997. Structure determination of minute virus of mice. *Acta Crystallogr D Biol Crystallogr* **53**:93-102.
185. **Lombardo, E., J. C. Ramirez, M. Agbandje-McKenna, and J. M. Almendral.** 2000. A beta-stranded motif drives capsid protein oligomers of the parvovirus minute virus of mice into the nucleus for viral assembly. *J Virol* **74**:3804-3814.
186. **Lombardo, E., J. C. Ramirez, J. Garcia, and J. M. Almendral.** 2002. Complementary roles of multiple nuclear targeting signals in the capsid proteins of the parvovirus minute virus of mice during assembly and onset of infection. *J Virol* **76**:7049-7059.
187. **Lopez de Turiso, J. A., E. Cortes, C. Martinez, R. Ruiz de Ybanez, I. Simarro, C. Vela, and I. Casal.** 1992. Recombinant vaccine for canine parvovirus in dogs. *J Virol* **66**:2748-2753.
188. **Lopez-Bueno, A., M. P. Rubio, N. Bryant, R. McKenna, M. Agbandje-McKenna, and J. M. Almendral.** 2006. Host-selected amino acid changes at the sialic acid binding pocket of the parvovirus capsid modulate cell binding affinity and determine virulence. *J Virol* **80**:1563-1573.
189. **Lopez-Bueno, A., J. C. Segovia, J. A. Bueren, M. G. O'Sullivan, F. Wang, P. Tattersall, and J. M. Almendral.** 2008. Evolution to pathogenicity of the parvovirus minute virus of mice in immunodeficient mice involves genetic heterogeneity at the capsid domain that determines tropism. *J Virol* **82**:1195-1203.
190. **Lux, K., N. Goerlitz, S. Schlemminger, L. Perabo, D. Goldnau, J. Endell, K. Leike, D. M. Kofler, S. Finke, M. Hallek, and H. Buning.** 2005. Green fluorescent protein-tagged adeno-associated virus particles allow the study of cytosolic and nuclear trafficking. *J Virol* **79**:11776-11787.
191. **Mani, B., C. Baltzer, N. Valle, J. M. Almendral, C. Kempf, and C. Ros.** 2006. Low pH-dependent endosomal processing of the incoming parvovirus minute virus of mice virion leads to externalization of the VP1 N-terminal sequence (N-VP1), N-VP2 cleavage, and uncoating of the full-length genome. *J Virol* **80**:1015-1024.
192. **Mani, B., M. Gerber, P. Lieby, N. Boschetti, C. Kempf, and C. Ros.** 2007. Molecular mechanism underlying B19 virus inactivation and comparison to other parvoviruses. *Transfusion* **47**:1765-1774.

193. **Maroto, B., J. C. Ramirez, and J. M. Almendral.** 2000. Phosphorylation status of the parvovirus minute virus of mice particle: mapping and biological relevance of the major phosphorylation sites. *J Virol* **74**:10892-10902.
194. **Maroto, B., N. Valle, R. Saffrich, and J. M. Almendral.** 2004. Nuclear export of the nonenveloped parvovirus virion is directed by an unordered protein signal exposed on the capsid surface. *J Virol* **78**:10685-10694.
195. **Matthews, B. W.** 1968. Solvent content of protein crystals. *J Mol Biol* **33**:491-497.
196. **Maxwell, I. H., A. L. Spitzer, F. Maxwell, and D. J. Pintel.** 1995. The capsid determinant of fibrotropism for the MVMP strain of minute virus of mice functions via VP2 and not VP1. *J Virol* **69**:5829-5832.
197. **McKenna, R., L. L. Ilag, and M. G. Rossmann.** 1994. Analysis of the Single-stranded DNA Bacteriophage [phi] X174, Refined at a Resolution of 3· 0 Å. *Journal of molecular biology* **237**:517-543.
198. **McKenna, R., N. H. Olson, P. R. Chipman, T. S. Baker, T. F. Booth, J. Christensen, B. Aasted, J. M. Fox, M. E. Bloom, J. B. Wolfinbarger, and M. Agbandje-McKenna.** 1999. Three-dimensional structure of Aleutian mink disease parvovirus: implications for disease pathogenicity. *J Virol* **73**:6882-6891.
199. **McPherson, A.** 1982. *Preparation and Analysis of Protein Crystals* 1ed. Wiley, New York.
200. **Miller, C. L., and D. J. Pintel.** 2002. Interaction between parvovirus NS2 protein and nuclear export factor Crm1 is important for viral egress from the nucleus of murine cells. *J Virol* **76**:3257-3266.
201. **Moehler, M., B. Blechacz, N. Weiskopf, M. Zeidler, W. Stremmel, J. Rommelaere, P. R. Galle, and J. J. Cornelis.** 2001. Effective infection, apoptotic cell killing and gene transfer of human hepatoma cells but not primary hepatocytes by parvovirus H1 and derived vectors. *Cancer Gene Ther* **8**:158-167.
202. **Moehler, M., M. Zeidler, J. Schede, J. Rommelaere, P. R. Galle, J. J. Cornelis, and M. Heike.** 2003. Oncolytic parvovirus H1 induces release of heat-shock protein HSP72 in susceptible human tumor cells but may not affect primary immune cells. *Cancer Gene Ther* **10**:477-480.
203. **Monahan, P. E., C. D. Lothrop, J. Sun, M. L. Hirsch, T. Kafri, B. Kantor, R. Sarkar, D. M. Tillson, J. R. Elia, and R. J. Samulski.** 2010. Proteasome inhibitors enhance gene delivery by AAV virus vectors expressing large genomes in hemophilia mouse and dog models: a strategy for broad clinical application. *Mol Ther* **18**:1907-1916.
204. **Mueller, C., and T. R. Flotte.** 2008. Clinical gene therapy using recombinant adeno-associated virus vectors. *Gene Ther* **15**:858-863.

205. **Munakata, Y., T. Saito-Ito, K. Kumura-Ishii, J. Huang, T. Kodera, T. Ishii, Y. Hirabayashi, Y. Koyanagi, and T. Sasaki.** 2005. Ku80 autoantigen as a cellular coreceptor for human parvovirus B19 infection. *Blood* **106**:3449-3456.
206. **Naeger, L. K., J. Cater, and D. J. Pintel.** 1990. The small nonstructural protein (NS2) of the parvovirus minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner. *J Virol* **64**:6166-6175.
207. **Naeger, L. K., N. Salomé, and D. J. Pintel.** 1993. NS2 is required for efficient translation of viral mRNA in minute virus of mice-infected murine cells. *J Virol* **67**:1034-1043.
208. **Nam, H. J., B. L. Gurda, R. McKenna, M. Potter, B. Byrne, M. Salganik, N. Muzychka, and M. Agbandje-McKenna.** 2011. Structural studies of adeno-associated virus serotype 8 capsid transitions associated with endosomal trafficking. *J Virol* **85**:11791-11799.
209. **Nam, H. J., B. Gurda-Whitaker, W. Y. Gan, S. Ilaria, R. McKenna, P. Mehta, R. A. Alvarez, and M. Agbandje-McKenna.** 2006. Identification of the sialic acid structures recognized by minute virus of mice and the role of binding affinity in virulence adaptation. *J Biol Chem* **281**:25670-25677.
210. **Nam, H. J., M. D. Lane, E. Padron, B. Gurda, R. McKenna, E. Kohlbrenner, G. Aslanidi, B. Byrne, N. Muzychka, S. Zolotukhin, and M. Agbandje-McKenna.** 2007. Structure of adeno-associated virus serotype 8, a gene therapy vector. *J Virol* **81**:12260-12271.
211. **Nathwani, A. C., Tuddenham, E.G.D, Rangarajan, S., Rosales, C., McIntosh, J., Linch, D.C., Chowdary, P., Riddell, A., Pie, J., Harrington, C., O'Beirne, J., Smith, K., Pasi, J., Glader, B., Rustagi, P., Ng, C.Y.C., Kay, M.A., Zhou, J., Spence, Y., Morton, C.L., Allay, J., Coleman, J., Sleep, S., Cunningham, J.M., Srivastava, D., Basner-Tschakarjan, E., Mingozi, F., High, K.A., Gray, J.T., Reiss, U.M., Nienhuis, A.W., Davidoff, A.M.** 2011. Adenovirus-Associated Virus Vector–Mediated Gene Transfer in Hemophilia B. *the New England Journal of Medicine*.
212. **Navaza, J.** 2001. Implementation of molecular replacement in AMoRe. *Acta Crystallogr D Biol Crystallogr* **57**:1367-1372.
213. **Nelson, C. D., E. Minkkinen, M. Bergkvist, K. Hoelzer, M. Fisher, B. Bothner, and C. R. Parrish.** 2008. Detecting small changes and additional peptides in the canine parvovirus capsid structure. *J Virol* **82**:10397-10407.
214. **Ng, R., L. Govindasamy, B. L. Gurda, R. McKenna, O. G. Kozyreva, R. J. Samulski, K. N. Parent, T. S. Baker, and M. Agbandje-McKenna.** 2010. Structural characterization of the dual glycan binding adeno-associated virus serotype 6. *J Virol* **84**:12945-12957.

215. **Nonnenmacher, M., and T. Weber.** 2011. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe* **10**:563-576.
216. **North, S., H. Huang, S. Sundaram, J. Jang-Lee, A. Etienne, A. Trollope, S. Chalabi, A. Dell, P. Stanley, and S. Haslam.** 2010. Glycomics Profiling of Chinese Hamster Ovary Cell Glycosylation Mutants Reveals N-Glycans of a Novel Size and Complexity. *Journal of Biological Chemistry* **285**:5759-5775.
217. **Nuesch, J., and J. Rommelaere.** 2006. NS1 interaction with CKII alpha: Novel protein complex mediating parvovirus-induced cytotoxicity. *Journal of Virology* **80**:4729-4739.
218. **Nüesch, J. P. F.** 2006. Regulation of non-structural protein functions by differential synthesis, modification and trafficking, p. 275-289. *In* J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), *Parvoviruses*. Hodder Arnold, London.
219. **O'Donnell, J., K. A. Taylor, and M. S. Chapman.** 2009. Adeno-associated virus-2 and its primary cellular receptor--Cryo-EM structure of a heparin complex. *Virology* **385**:434-443.
220. **Ohshima, T., T. Nakajima, T. Oishi, N. Imamoto, Y. Yoneda, A. Fukamizu, and K. Yagami.** 1999. CRM1 mediates nuclear export of nonstructural protein 2 from parvovirus minute virus of mice. *Biochem Biophys Res Commun* **264**:144-150.
221. **Olijslagers, S., A. Y. Dege, C. Dinsart, M. Voorhoeve, J. Rommelaere, M. H. Noteborn, and J. J. Cornelis.** 2001. Potentiation of a recombinant oncolytic parvovirus by expression of Apoptin. *Cancer Gene Ther* **8**:958-965.
222. **Olofsson, S., and T. Bergstrom.** 2005. Glycoconjugate glycans as viral receptors. *Ann Med* **37**:154-172.
223. **Opie, S. R., K. H. Warrington, Jr., M. Agbandje-McKenna, S. Zolotukhin, and N. Muzyczka.** 2003. Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J Virol* **77**:6995-7006.
224. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**:307-326.
225. **Padron, E., V. Bowman, N. Kaludov, L. Govindasamy, H. Levy, P. Nick, R. McKenna, N. Muzyczka, J. A. Chiorini, T. S. Baker, and M. Agbandje-McKenna.** 2005. Structure of adeno-associated virus type 4. *J Virol* **79**:5047-5058.

226. **Paglino, J., and P. Tattersall.** 2011. The parvoviral capsid controls an intracellular phase of infection essential for efficient killing of stepwise-transformed human fibroblasts. *Virology* **416**:32-41.
227. **Pajusola, K., M. Gruchala, H. Joch, T. F. Luscher, S. Yla-Herttuala, and H. Bueler.** 2002. Cell-type-specific characteristics modulate the transduction efficiency of adeno-associated virus type 2 and restrain infection of endothelial cells. *J Virol* **76**:11530-11540.
228. **Palermo, L. M., S. L. Hafenstein, and C. R. Parrish.** 2006. Purified feline and canine transferrin receptors reveal complex interactions with the capsids of canine and feline parvoviruses that correspond to their host ranges. *J Virol* **80**:8482-8492.
229. **Palermo, L. M., K. Hueffer, and C. R. Parrish.** 2003. Residues in the apical domain of the feline and canine transferrin receptors control host-specific binding and cell infection of canine and feline parvoviruses. *J Virol* **77**:8915-8923.
230. **Parker, J. S., W. J. Murphy, D. Wang, S. J. O'Brien, and C. R. Parrish.** 2001. Canine and feline parvoviruses can use human or feline transferrin receptors to bind, enter, and infect cells. *J Virol* **75**:3896-3902.
231. **Parker, J. S., and C. R. Parrish.** 2000. Cellular uptake and infection by canine parvovirus involves rapid dynamin-regulated clathrin-mediated endocytosis, followed by slower intracellular trafficking. *J Virol* **74**:1919-1930.
232. **Parrish, C. R.** 2006. Autonomous Parvovirus Variation and Evolution, p. 47-54. *In* J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), *Parvoviruses*. Hodder Arnold, London.
233. **Parrish, C. R., C. F. Aquadro, M. L. Strassheim, J. F. Evermann, J. Y. Sgro, and H. O. Mohammed.** 1991. Rapid antigenic-type replacement and DNA sequence evolution of canine parvovirus. *J Virol* **65**:6544-6552.
234. **Parrish, C. R., P. Have, W. J. Foreyt, J. F. Evermann, M. Senda, and L. E. Carmichael.** 1988. The global spread and replacement of canine parvovirus strains. *J Gen Virol* **69 ( Pt 5)**:1111-1116.
235. **Parrish, C. R., and Y. Kawaoka.** 2005. The origins of new pandemic viruses: the acquisition of new host ranges by canine parvovirus and influenza A viruses. *Annu Rev Microbiol* **59**:553-586.
236. **Perez, R., M. Castellanos, A. Rodriguez-Huete, and M. G. Mateu.** 2011. Molecular determinants of self-association and rearrangement of a trimeric intermediate during the assembly of a parvovirus capsid. *J Mol Biol* **413**:32-40.

237. **Perros, M., L. Deleu, J. M. Vanacker, Z. Kherrouche, N. Spruyt, S. Faisst, and J. Rommelaere.** 1995. Upstream CREs participate in the basal activity of minute virus of mice promoter P4 and in its stimulation in ras-transformed cells. *J Virol* **69**:5506-5515.
238. **Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin.** 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**:1605-1612.
239. **Pintel, D., D. Dadachanji, C. R. Astell, and D. C. Ward.** 1983. The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. *Nucleic Acids Res* **11**:1019-1038.
240. **Plevka, P., S. Hafenstein, L. Li, A. D'Abregamo, Jr., S. F. Cotmore, M. G. Rossmann, and P. Tattersall.** 2011. Structure of a packaging-defective mutant of minute virus of mice indicates that the genome is packaged via a pore at a 5-fold axis. *J Virol* **85**:4822-4827.
241. **Prasad, K. M., and J. P. Trempe.** 1995. The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology* **214**:360-370.
242. **Previsani, N., S. Fontana, B. Hirt, and P. Beard.** 1997. Growth of the parvovirus minute virus of mice MVMp3 in EL4 lymphocytes is restricted after cell entry and before viral DNA amplification: cell-specific differences in virus uncoating in vitro. *J Virol* **71**:7769-7780.
243. **Qing, K., C. Mah, J. Hansen, S. Zhou, V. Dwarki, and A. Srivastava.** 1999. Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat Med* **5**:71-77.
244. **Ramirez, J. C., A. Fairen, and J. M. Almendral.** 1996. Parvovirus minute virus of mice strain i multiplication and pathogenesis in the newborn mouse brain are restricted to proliferative areas and to migratory cerebellar young neurons. *J Virol* **70**:8109-8116.
245. **Reardon, J. E., and T. Spector.** 1989. Herpes simplex virus type 1 DNA polymerase. Mechanism of inhibition by acyclovir triphosphate. *J Biol Chem* **264**:7405-7411.
246. **Reguera, J., A. Carreira, L. Riolobos, J. M. Almendral, and M. G. Mateu.** 2004. Role of interfacial amino acid residues in assembly, stability, and conformation of a spherical virus capsid. *Proc Natl Acad Sci U S A* **101**:2724-2729.
247. **Reguera, J., E. Grueso, A. Carreira, C. Sanchez-Martinez, J. M. Almendral, and M. G. Mateu.** 2005. Functional relevance of amino acid residues involved in interactions with ordered nucleic acid in a spherical virus. *J Biol Chem* **280**:17969-17977.

248. **Riolobos, L., J. Reguera, M. G. Mateu, and J. M. Almendral.** 2006. Nuclear transport of trimeric assembly intermediates exerts a morphogenetic control on the icosahedral parvovirus capsid. *J Mol Biol* **357**:1026-1038.
249. **Ritter, G., E. Boosfeld, E. Markstein, R. K. Yu, S. L. Ren, W. B. Stallcup, H. F. Oettgen, L. J. Old, and P. O. Livingston.** 1990. Biochemical and serological characteristics of natural 9-O-acetyl GD3 from human melanoma and bovine buttermilk and chemically O-acetylated GD3. *Cancer Res* **50**:1403-1410.
250. **Rogers, G. N., J. C. Paulson, R. S. Daniels, J. J. Skehel, I. A. Wilson, and D. C. Wiley.** 1983. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature* **304**:76-78.
251. **Rommelaere, J., and J. J. Cornelis.** 1991. Antineoplastic activity of parvoviruses. *J Virol Methods* **33**:233-251.
252. **Ros, C., C. Baltzer, B. Mani, and C. Kempf.** 2006. Parvovirus uncoating in vitro reveals a mechanism of DNA release without capsid disassembly and striking differences in encapsidated DNA stability. *Virology* **345**:137-147.
253. **Ros, C., C. J. Burckhardt, and C. Kempf.** 2002. Cytoplasmic trafficking of minute virus of mice: low-pH requirement, routing to late endosomes, and proteasome interaction. *J Virol* **76**:12634-12645.
254. **Ros, C., M. Gerber, and C. Kempf.** 2006. Conformational changes in the VP1-unique region of native human parvovirus B19 lead to exposure of internal sequences that play a role in virus neutralization and infectivity. *J Virol* **80**:12017-12024.
255. **Ros, C., and C. Kempf.** 2004. The ubiquitin-proteasome machinery is essential for nuclear translocation of incoming minute virus of mice. *Virology* **324**:350-360.
256. **Rose, J. A., J. V. Maizel, Jr., J. K. Inman, and A. J. Shatkin.** 1971. Structural proteins of adenovirus-associated viruses. *J Virol* **8**:766-770.
257. **Rosenfeld, S. J., K. Yoshimoto, S. Kajigaya, S. Anderson, N. S. Young, A. Field, P. Warrener, G. Bansal, and M. S. Collett.** 1992. Unique region of the minor capsid protein of human parvovirus B19 is exposed on the virion surface. *J Clin Invest* **89**:2023-2029.
258. **Rossmann, M. G., R. McKenna, L. Tong, D. Xia, J. Dai, H. Wu, H. K. Choi, and R. E. Lynch.** 1992. Molecular replacement real-space averaging. *Journal of applied crystallography* **25**:166-180.
259. **Rubio, M. P., A. Lopez-Bueno, and J. M. Almendral.** 2005. Virulent variants emerging in mice infected with the apathogenic prototype strain of the parvovirus minute virus of mice exhibit a capsid with low avidity for a primary receptor. *J Virol* **79**:11280-11290.

260. **Saliki, J. T., B. Mizak, H. P. Flore, R. R. Gettig, J. P. Burand, L. E. Carmichael, H. A. Wood, and C. R. Parrish.** 1992. Canine parvovirus empty capsids produced by expression in a baculovirus vector: use in analysis of viral properties and immunization of dogs. *J Gen Virol* **73** ( Pt 2):369-374.
261. **Sanlioglu, S., P. K. Benson, J. Yang, E. M. Atkinson, T. Reynolds, and J. F. Engelhardt.** 2000. Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by rac1 and phosphatidylinositol-3 kinase activation. *J Virol* **74**:9184-9196.
262. **Sato, C., H. Fukuoka, K. Ohta, T. Matsuda, R. Koshino, K. Kobayashi, F. A. Troy, 2nd, and K. Kitajima.** 2000. Frequent occurrence of pre-existing alpha 2->8-linked disialic and oligosialic acids with chain lengths up to 7 Sia residues in mammalian brain glycoproteins. Prevalence revealed by highly sensitive chemical methods and anti-di-, oligo-, and poly-Sia antibodies specific for defined chain lengths. *J Biol Chem* **275**:15422-15431.
263. **Sato, C., T. Matsuda, and K. Kitajima.** 2002. Neuronal differentiation-dependent expression of the disialic acid epitope on CD166 and its involvement in neurite formation in Neuro2A cells, p. 45299-45305, *J Biol Chem*, vol. 277, United States.
264. **Schmidt, M., and J. A. Chiorini.** 2006. Gangliosides are essential for bovine adeno-associated virus entry. *J Virol* **80**:5516-5522.
265. **Schnaar, R. L., A. Suzuki, and P. Stanley.** 2009. Glycosphingolipids. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, and M. E. Etzler (ed.), *Essentials of Glycobiology*. The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor NY.
266. **Schwegmann-Wessels, C., and G. Herrler.** 2006. Sialic acids as receptor determinants for coronaviruses. *Glycoconjugate Journal* **23**:51-58.
267. **Sedlik, C., A. Dridi, E. Deriaud, M. F. Saron, P. Rueda, J. Sarraseca, J. I. Casal, and C. Leclerc.** 1999. Intranasal delivery of recombinant parvovirus-like particles elicits cytotoxic T-cell and neutralizing antibody responses. *J Virol* **73**:2739-2744.
268. **Sedlik, C., J. Sarraseca, P. Rueda, C. Leclerc, and I. Casal.** 1995. Immunogenicity of poliovirus B and T cell epitopes presented by hybrid porcine parvovirus particles. *J Gen Virol* **76** ( Pt 9):2361-2368.
269. **Segovia, J. C., J. M. Gallego, J. A. Bueren, and J. M. Almendral.** 1999. Severe leukopenia and dysregulated erythropoiesis in SCID mice persistently infected with the parvovirus minute virus of mice. *J Virol* **73**:1774-1784.
270. **Seiler, M. P., A. D. Miller, J. Zabner, and C. L. Halbert.** 2006. Adeno-associated virus types 5 and 6 use distinct receptors for cell entry. *Hum Gene Ther* **17**:10-19.

271. **Seki, T., and Y. Arai.** 1993. Highly polysialylated neural cell adhesion molecule (NCAM-H) is expressed by newly generated granule cells in the dentate gyrus of the adult rat. *J Neurosci* **13**:2351-2358.
272. **Shen, S., K. D. Bryant, S. M. Brown, S. H. Randell, and A. Asokan.** 2011. Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. *J Biol Chem* **286**:13532-13540.
273. **Shieh, M. T., and P. G. Spear.** 1994. Herpesvirus-induced cell fusion that is dependent on cell surface heparan sulfate or soluble heparin. *J Virol* **68**:1224-1228.
274. **Simpson, A. A., V. Chandrasekar, B. Hebert, G. M. Sullivan, M. G. Rossmann, and C. R. Parrish.** 2000. Host range and variability of calcium binding by surface loops in the capsids of canine and feline parvoviruses. *J Mol Biol* **300**:597-610.
275. **Simpson, A. A., P. R. Chipman, T. S. Baker, P. Tijssen, and M. G. Rossmann.** 1998. The structure of an insect parvovirus (*Galleria mellonella* densovirus) at 3.7 Å resolution. *Structure* **6**:1355-1367.
276. **Simpson, A. A., B. Hebert, G. M. Sullivan, C. R. Parrish, Z. Zadot, P. Tijssen, and M. G. Rossmann.** 2002. The structure of porcine parvovirus: comparison with related viruses. *J Mol Biol* **315**:1189-1198.
277. **Song, X., B. Xia, S. Stowell, Y. Lasanajak, D. Smith, and R. Cummings.** 2009. Novel Fluorescent Glycan Microarray Strategy Reveals Ligands for Galectins. *Chemistry & Biology* **16**:36-47.
278. **Song, X., H. Yu, X. Chen, Y. Lasanajak, M. Tappert, G. Air, V. Tiwari, H. Cao, H. Chokhawala, H. Zheng, R. Cummings, and D. Smith.** 2011. A Sialylated Glycan Microarray Reveals Novel Interactions of Modified Sialic Acids with Proteins and Viruses. *Journal of Biological Chemistry* **286**:31610-31622.
279. **Sonntag, F., S. Bleker, B. Leuchs, R. Fischer, and J. A. Kleinschmidt.** 2006. Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *J Virol* **80**:11040-11054.
280. **Sonntag, F., K. Kother, K. Schmidt, M. Weghofer, C. Raupp, K. Nieto, A. Kuck, B. Gerlach, B. Bottcher, O. J. Muller, K. Lux, M. Horer, and J. A. Kleinschmidt.** 2011. The assembly-activating protein promotes capsid assembly of different adeno-associated virus serotypes. *J Virol* **85**:12686-12697.
281. **Sonntag, F., K. Schmidt, and J. A. Kleinschmidt.** 2010. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci U S A* **107**:10220-10225.

282. **Spalholz, B. A., and P. Tattersall.** 1983. Interaction of minute virus of mice with differentiated cells: strain-dependent target cell specificity is mediated by intracellular factors. *J Virol* **46**:937-943.
283. **Stahnke, S., K. Lux, S. Uhrig, F. Kreppel, M. Hosei, O. Coutelle, M. Ogris, M. Hallek, and H. Buning.** 2011. Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. *Virology* **409**:77-83.
284. **Stanford, M. M., M. Shaban, J. W. Barrett, S. J. Werden, P. A. Gilbert, J. Bondy-Denomy, L. Mackenzie, K. C. Graham, A. F. Chambers, and G. McFadden.** 2008. Myxoma virus oncolysis of primary and metastatic B16F10 mouse tumors *in vivo*. *Mol Ther* **16**:52-59.
285. **Stieger, K., T. Cronin, J. Bennett, and F. Rolling.** 2011. Adeno-associated virus mediated gene therapy for retinal degenerative diseases. *Methods Mol Biol* **807**:179-218.
286. **Suikkanen, S., T. Aaltonen, M. Nevalainen, O. Valilehto, L. Lindholm, M. Vuento, and M. Vihinen-Ranta.** 2003. Exploitation of microtubule cytoskeleton and dynein during parvoviral traffic toward the nucleus. *J Virol* **77**:10270-10279.
287. **Suikkanen, S., M. Antila, A. Jaatinen, M. Vihinen-Ranta, and M. Vuento.** 2003. Release of canine parvovirus from endocytic vesicles. *Virology* **316**:267-280.
288. **Suikkanen, S., K. Saajarvi, J. Hirsimaki, O. Valilehto, H. Reunanan, M. Vihinen-Ranta, and M. Vuento.** 2002. Role of recycling endosomes and lysosomes in dynein-dependent entry of canine parvovirus. *J Virol* **76**:4401-4411.
289. **Summerford, C., J. S. Bartlett, and R. J. Samulski.** 1999. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat Med* **5**:78-82.
290. **Summerford, C., and R. J. Samulski.** 1998. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol* **72**:1438-1445.
291. **Takeuchi, Y., M. Akutsu, K. Murayama, N. Shimizu, and H. Hoshino.** 1991. Host range mutant of human immunodeficiency virus type 1: modification of cell tropism by a single point mutation at the neutralization epitope in the env gene. *J Virol* **65**:1710-1718.
292. **Talora, C., D. C. Sgroi, C. P. Crum, and G. P. Dotto.** 2002. Specific down-modulation of Notch1 signaling in cervical cancer cells is required for sustained HPV-E6/E7 expression and late steps of malignant transformation. *Genes Dev* **16**:2252-2263.

293. **Tattersall, P.** 2006. The Evolution of Parvovirus Taxonomy, p. 5-14. In J. R. Kerr, Cotmore, S.F., Bloom M.E., Linden, R.M., Parrish, C.R. (ed.), Parvoviruses. Hodder Arnold, London.
294. **Tattersall, P., and J. Bratton.** 1983. Reciprocal productive and restrictive virus-cell interactions of immunosuppressive and prototype strains of minute virus of mice. *J Virol* **46**:944-955.
295. **Tattersall, P., P. J. Cawte, A. J. Shatkin, and D. C. Ward.** 1976. Three structural polypeptides coded for by minite virus of mice, a parvovirus. *J Virol* **20**:273-289.
296. **Tattersall, P., A. J. Shatkin, and D. C. Ward.** 1977. Sequence homology between the structural polypeptides of minute virus of mice. *J Mol Biol* **111**:375-394.
297. **Thacker, T. C., and F. B. Johnson.** 1998. Binding of bovine parvovirus to erythrocyte membrane sialylglycoproteins. *J Gen Virol* **79 ( Pt 9)**:2163-2169.
298. **Theobald, D. L., R. M. Mitton-Fry, and D. S. Wuttke.** 2003. Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**:115-133.
299. **Tong, L., and M. G. Rossmann.** 1997. Rotation function calculations with GLRF program. *Methods Enzymol* **276**:594-611.
300. **Toolan, H. W.** 1990. The rodent parvoviruses, p. 159-176. In P. Tijssen (ed.), *Handbook of parvoviruses*. CRC Press, Boca Raton, Fla.
301. **Toolan, H. W., and N. Ledinko.** 1968. Inhibition by H-1 virus of the incidence of tumors produced by adenovirus 12 in hamsters. *Virology* **35**:475-478.
302. **Toolan, H. W., E. L. Saunders, C. M. Southam, A. E. Moore, and A. G. Levin.** 1965. H-1 VIRUS VIREMIA IN THE HUMAN. *Proc Soc Exp Biol Med* **119**:711-715.
303. **Tresnan, D. B., L. Southard, W. Weichert, J. Y. Sgro, and C. R. Parrish.** 1995. Analysis of the cell and erythrocyte binding activities of the dimple and canyon regions of the canine parvovirus capsid. *Virology* **211**:123-132.
304. **Truyen, U., A. Gruenberg, S. F. Chang, B. Obermaier, P. Veijalainen, and C. R. Parrish.** 1995. Evolution of the feline-subgroup parvoviruses and the control of canine host range in vivo. *J Virol* **69**:4702-4710.
305. **Truyen, U., and C. R. Parrish.** 1992. Canine and feline host ranges of canine parvovirus and feline panleukopenia virus: distinct host cell tropisms of each virus in vitro and in vivo. *J Virol* **66**:5399-5408.

306. **Tsao, J., M. S. Chapman, M. Agbandje, W. Keller, K. Smith, H. Wu, M. Luo, T. J. Smith, M. G. Rossmann, R. W. Compans, and et al.** 1991. The three-dimensional structure of canine parvovirus and its functional implications. *Science* **251**:1456-1464.
307. **Urzhumtseva, L., P. Afonine, P. Adams, and A. Urzhumtsev.** 2009. Crystallographic model quality at a glance. *Acta Crystallographica Section D-Biological Crystallography* **65**:297-300.
308. **Valegard, K., J. B. Murray, P. G. Stockley, N. J. Stonehouse, and L. Liljas.** 1994. Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* **371**:623-626.
309. **Varki, N. M., and A. Varki.** 2007. Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab Invest* **87**:851-857.
310. **Vihinen-Ranta, M., L. Kakkola, A. Kalela, P. Vilja, and M. Vuento.** 1997. Characterization of a nuclear localization signal of canine parvovirus capsid proteins. *Eur J Biochem* **250**:389-394.
311. **Vihinen-Ranta, M., A. Kalela, P. Makinen, L. Kakkola, V. Marjomaki, and M. Vuento.** 1998. Intracellular route of canine parvovirus entry. *J Virol* **72**:802-806.
312. **Vihinen-Ranta, M., D. Wang, W. S. Weichert, and C. R. Parrish.** 2002. The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection. *J Virol* **76**:1884-1891.
313. **Vihinen-Ranta, M., W. Yuan, and C. R. Parrish.** 2000. Cytoplasmic trafficking of the canine parvovirus capsid and its role in infection and nuclear transport. *J Virol* **74**:4853-4859.
314. **Walters, R. W., M. Agbandje-McKenna, V. D. Bowman, T. O. Moninger, N. H. Olson, M. Seiler, J. A. Chiorini, T. S. Baker, and J. Zabner.** 2004. Structure of adeno-associated virus serotype 5. *J Virol* **78**:3361-3371.
315. **Walters, R. W., S. M. Yi, S. Keshavjee, K. E. Brown, M. J. Welsh, J. A. Chiorini, and J. Zabner.** 2001. Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J Biol Chem* **276**:20610-20616.
316. **Weigel-Kelley, K. A., M. C. Yoder, and A. Srivastava.** 2003. Alpha5beta1 integrin as a cellular coreceptor for human parvovirus B19: requirement of functional activation of beta1 integrin for viral entry. *Blood* **102**:3927-3933.
317. **Weigel-Kelley, K. A., M. C. Yoder, and A. Srivastava.** 2001. Recombinant human parvovirus B19 vectors: erythrocyte P antigen is necessary but not sufficient for successful transduction of human hematopoietic cells. *J Virol* **75**:4110-4116.

318. **Weiss, M. S., A. Jabs, and R. Hilgenfeld.** 1998. Peptide bonds revisited. *Nat Struct Biol* **5**:676.
319. **Weller, M. L., P. Amornphimoltham, M. Schmidt, P. A. Wilson, J. S. Gutkind, and J. A. Chiorini.** 2010. Epidermal growth factor receptor is a co-receptor for adeno-associated virus serotype 6. *Nat Med* **16**:662-664.
320. **Wetzel, K., P. Menten, G. Opdenakker, J. Van Damme, H. J. Grone, N. Giese, A. Vecchi, S. Sozzani, J. J. Cornelis, J. Rommelaere, and C. Dinsart.** 2001. Transduction of human MCP-3 by a parvoviral vector induces leukocyte infiltration and reduces growth of human cervical carcinoma cell xenografts. *J Gene Med* **3**:326-337.
321. **Wikoff, W. R., G. Wang, C. R. Parrish, R. H. Cheng, M. L. Strassheim, T. S. Baker, and M. G. Rossmann.** 1994. The structure of a neutralized virus: canine parvovirus complexed with neutralizing antibody fragment. *Structure* **2**:595-607.
322. **Wu, H., and M. G. Rossmann.** 1993. The canine parvovirus empty capsid structure. *J Mol Biol* **233**:231-244.
323. **Wu, P., W. Xiao, T. Conlon, J. Hughes, M. Agbandje-McKenna, T. Ferkol, T. Flotte, and N. Muzyczka.** 2000. Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism. *J Virol* **74**:8635-8647.
324. **Wu, Z., A. Asokan, J. C. Grieger, L. Govindasamy, M. Agbandje-McKenna, and R. J. Samulski.** 2006. Single amino acid changes can influence titer, heparin binding, and tissue tropism in different adeno-associated virus serotypes. *J Virol* **80**:11393-11397.
325. **Wu, Z., E. Miller, M. Agbandje-McKenna, and R. J. Samulski.** 2006. Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol* **80**:9093-9103.
326. **Xiao, W., K. H. Warrington, Jr., P. Hearing, J. Hughes, and N. Muzyczka.** 2002. Adenovirus-facilitated nuclear translocation of adeno-associated virus type 2. *J Virol* **76**:11505-11517.
327. **Xiao, X., W. Xiao, J. Li, and R. J. Samulski.** 1997. A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. *J Virol* **71**:941-948.
328. **Xie, Q., W. Bu, S. Bhatia, J. Hare, T. Somasundaram, A. Azzi, and M. S. Chapman.** 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci U S A* **99**:10405-10410.
329. **Xie, Q., and M. S. Chapman.** 1996. Canine parvovirus capsid structure, analyzed at 2.9 Å resolution. *J Mol Biol* **264**:497-520.

330. **Xie, Q., T. F. Lerch, N. L. Meyer, and M. S. Chapman.** 2011. Structure-function analysis of receptor-binding in adeno-associated virus serotype 6 (AAV-6). *Virology* **420**:10-19.
331. **Yan, Z., D. C. Lei-Butters, X. Liu, Y. Zhang, L. Zhang, M. Luo, R. Zak, and J. F. Engelhardt.** 2006. Unique biologic properties of recombinant AAV1 transduction in polarized human airway epithelia. *J Biol Chem* **281**:29684-29692.
332. **Yan, Z., R. Zak, G. W. Luxton, T. C. Ritchie, U. Bantel-Schaal, and J. F. Engelhardt.** 2002. Ubiquitination of both adeno-associated virus type 2 and 5 capsid proteins affects the transduction efficiency of recombinant vectors. *J Virol* **76**:2043-2053.
333. **Yan, Z., R. Zak, Y. Zhang, W. Ding, S. Godwin, K. Munson, R. Peluso, and J. F. Engelhardt.** 2004. Distinct classes of proteasome-modulating agents cooperatively augment recombinant adeno-associated virus type 2 and type 5-mediated transduction from the apical surfaces of human airway epithelia. *J Virol* **78**:2863-2874.
334. **Yunoki, M., M. Tsujikawa, T. Urayama, Y. Sasaki, M. Morita, H. Tanaka, S. Hattori, K. Takechi, and K. Ikuta.** 2003. Heat sensitivity of human parvovirus B19. *Vox Sang* **84**:164-169.
335. **Zadori, Z., J. Szelei, M. C. Lacoste, Y. Li, S. Gariepy, P. Raymond, M. Allaire, I. R. Nabi, and P. Tijssen.** 2001. A viral phospholipase A2 is required for parvovirus infectivity. *Dev Cell* **1**:291-302.
336. **Zhong, L., B. Li, G. Jayandharan, C. S. Mah, L. Govindasamy, M. Agbandje-McKenna, R. W. Herzog, K. A. Weigel-Van Aken, J. A. Hobbs, S. Zolotukhin, N. Muzyczka, and A. Srivastava.** 2008. Tyrosine-phosphorylation of AAV2 vectors and its consequences on viral intracellular trafficking and transgene expression. *Virology* **381**:194-202.
337. **Zhong, L., B. Li, C. S. Mah, L. Govindasamy, M. Agbandje-McKenna, M. Cooper, R. W. Herzog, I. Zolotukhin, K. H. Warrington, Jr., K. A. Weigel-Van Aken, J. A. Hobbs, S. Zolotukhin, N. Muzyczka, and A. Srivastava.** 2008. Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci U S A* **105**:7827-7832.
338. **Zhou, L., Y. Luo, Y. Wu, J. Tsao, and M. Luo.** 2000. Sialylation of the host receptor may modulate entry of demyelinating persistent Theiler's virus. *J Virol* **74**:1477-1485.

## BIOGRAPHICAL SKETCH

Sujata Halder was born in 1983 in the capital city of New Delhi, India. She spent most of her childhood and schooling years in New Delhi. She completed her high school education at Birla Vidya Niketan, India in 2001. Towards the later years of her high school education, she became very interested in Biology, and wanted to become a doctor just like her father. However, after graduating from high school she enrolled in Bachelor of Science (B.Sc.) at Sri Venkateswara College, University of Delhi, India and graduated in 2004 with an honors in Biochemistry. During this time, she conducted summer research at School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India gaining experience in molecular biology techniques under the guidance of Dr. Sudha Bhattacharya. This experience kindled within her a keen interest for further research. She then enrolled in Masters of Science (M.Sc.) in Biotechnology from the esteemed Indian Institute of Technology (Roorkee), India and graduated in 2006. During her master's program, Sujata developed deep passion for serious scientific research, where under the supervision of Dr. Partha Roy she worked on her thesis in developing functional assays for endocrine disruptors present in the environment. After completion of her master's program, Sujata decided to explore new and better opportunities and decided to apply to the graduate schools in the United States of America for a PhD degree. However, she was undecided about any particular field of study and wanted to learn new techniques before committing to join a laboratory, and hence, only applied to schools with interdisciplinary programs that offered lab rotation experience. In the fall of 2006, she joined the Interdisciplinary Program (IDP) at University of Florida, Gainesville (UFL). Following the first year of core courses and lab rotations, she finally began research as a graduate assistant under the supervision of

Dr. Mavis Agbandje-McKenna (Professor, Department of Biochemistry and Molecular Biology, UFL) in the field of virus crystallography. Her areas of interest include virology, macromolecular crystallography, and cancer biology. Her favorite hobby is amateur photography.