

Structural Studies of Polyomavirus and Herpesvirus Attachment Processes

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von

Dipl. Biochem. Michael Henning Christian Buch
aus Herborn

Tübingen

2016

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 16. Februar 2016

Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Thilo Stehle
2. Berichterstatter:	Prof. John E. Johnson, PhD

TABLE OF CONTENTS

TABLE OF CONTENTS	I
ABSTRACT	III
ZUSAMMENFASSUNG	V
ABBREVIATIONS	VII
1. INTRODUCTION	1
1.1 The Discovery of Viruses and Birth of Virology	1
1.2 Structural Features of Viruses	2
1.3 Viral Cell Attachment and Entry Strategies	4
1.4 Polyomaviruses: General Information, Cell Attachment, and Entry	5
1.5 Herpes Simplex Viruses and How They Enter a Cell	10
1.5.1 <i>Structural and Biological Properties of Glycoprotein C</i>	14
1.6 The Roles of Glycans in Viral Infection	15
1.6.1 <i>Gangliosides</i>	16
1.6.1.1 Sialic Acids as Viral Receptors	17
1.6.2 <i>Glycosaminoglycans</i>	18
2. AIMS OF THIS DISSERTATION	20
3. SMALL CHANGES, BIG DIFFERENCES – HOW MINOR ALTERATIONS IN THE VP1 BINDING POCKET CAN AFFECT THE ATTACHMENT AND INFECTION PROPERTIES OF POLYOMAVIRUSES	21
3.1 Results	21
3.1.1 <i>Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity</i>	21
3.1.2 <i>Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site</i>	23
3.1.3 <i>Mutations in the GM1 Binding Site of Simian Virus 40 Alter Receptor Usage and Cell Tropism</i>	25
3.2 Discussion	27
4. DEVELOPMENT OF A PURIFICATION STRATEGY FOR HSV-1 GLYCOPROTEIN C	30
4.1 Material and Methods	30
4.1.1 <i>Materials</i>	30
4.1.1.1 Chemicals	30
4.1.1.2 Reagents, Enzymes, and Kits	30
4.1.1.3 Solutions and Buffers	31
4.1.1.4 Bacterial Strains and Vectors	35
4.1.1.5 Consumables	35
4.1.1.6 Instruments	36
4.1.1.7 Software and Online Tools	37
4.1.2 <i>Methods</i>	37
4.1.2.1 Microbiological Methods	37

4.1.2.2	Methods in Molecular Biology	38
4.1.2.3	Methods in Cell Culture	42
4.1.2.4	Methods in Protein Analysis	44
4.1.2.5	Protein Purification	46
4.2	Results	49
4.2.1	<i>Expression Constructs of gC II-5</i>	49
4.2.1.1	Cloning of gC II-5-His into pcDNA 3.1(-)	51
4.2.1.2	Cloning of gC II-5-Fc into pcDNA 3.1(-)	51
4.2.2	<i>Expression of gC II-5 in CHO Lec 3.2.8.1.</i>	52
4.2.2.1	Expression of gC II-5-His	52
4.2.2.2	Expression of gC II-5-Fc	53
4.2.3	<i>Purification of the B1C1B4 Monoclonal α-gC Antibody</i>	54
4.2.3.1	Protein G Affinity Chromatography	54
4.2.3.2	Analytical Size Exclusion Chromatography	54
4.2.4	<i>Purification of gC II-5-Fc</i>	55
4.2.4.1	Protein A Affinity Chromatography	56
4.2.4.2	TEV Protease Cleavage	57
4.2.4.3	Size Exclusion Chromatography	58
4.3	Discussion and Outlook	60
5.	APPENDIX	62
5.1	Acknowledgements	62
5.2	Licenses for Figures from other Publications	65
5.2.1	<i>Fig 3A: Yan et al., 1996; Structure</i>	65
5.2.2	<i>Fig 3B: Lee et al., 2008; Structure</i>	65
5.2.3	<i>Fig 6: Grünewald et al., 2003; Science</i>	65
5.2.4	<i>Fig 7: Bloom et al., 2010; Biochimica et Biophysica Acta – Gene Regulatory Mechanisms</i>	66
5.2.5	<i>Fig 11: Imberty et al., 2007; Carbohydrate Research</i>	66
5.3	Sequences	67
5.3.1	<i>gC II-5-His</i>	67
5.3.2	<i>gC II-5-Fc</i>	68
5.4	Provided gC II-5 Construct by Eurofins MWG Operon LLC	70
5.5	MALDI-MS Analysis of gC II-5-Fc	70
6.	REFERENCES	71
7.	PUBLICATIONS AND CONTRIBUTIONS	88

ABSTRACT

This dissertation addresses cell attachment processes of viruses and focuses on polyomaviruses and human herpes virus 1, commonly known as herpes simplex virus type 1.

Polyomaviruses (PyVs) belong to a continuously expanding family of double-stranded DNA viruses that infect a wide spectrum of fish, mammals, and birds. While they typically remain asymptomatic in healthy hosts, they can cause serious disease in immunocompromised individuals and have been linked to respiratory ailments, neurodegenerative symptoms, nephropathy, and cancer.

Most PyVs enter their host cells by binding to sialylated carbohydrate receptors that are attached to lipids, proteins, or both. This attachment process is mediated by the major capsid protein VP1, 360 copies of which form the icosahedral capsid of the virus. VP1 possesses a conserved β -barrel core structure, but the loops connecting the strands are highly divergent and define receptor specificity and virus pathogenicity; even small changes can have drastic effects on the viral life cycle. One example for this is the murine polyomavirus (MuPyV). It has long been known that single amino acid exchanges in the loops forming the receptor binding pocket of MuPyV VP1 drastically alter tumorigenicity and influence spread, but the molecular determinants are unknown. Three MuPyV strains with different pathogenicity, RA, PTA, and LID, have been investigated in this work. In addition to the already established ganglioside receptors GD1a and GT1b, a third receptor, GT1a, was identified, and its structure bound to MuPyV VP1 was solved by X-ray crystallography. Moreover, structures of GD1a as well as another glycan, DSLNT, in complex with the different MuPyV strains were elucidated in order to uncover possible strain-dependent binding modes. Interestingly, none of the introduced amino acid exchanges influence the binding mode for any of the investigated glycans. However, by employing crystallographic affinity measurements, it could be shown that the strains display altered affinities for their interaction partners, which suggests that the regulation of pathogenicity in RA, PTA, and LID is far more intricate than previously assumed.

In another project, the effects of amino acid exchanges in PyV VP1 were investigated in simian virus 40 (SV40). Here, three mutations were isolated that show abolished binding to the classical SV40 receptor GM1. Additionally, one mutant also shows altered tropism. It is unknown which ganglioside these mutants utilize, but modeling of the mutated VP1 residues showed that steric clashes result in abolished GM1 binding.

Finally, the VP1 structures of the newly discovered human polyomaviruses 6 and 7 (HPyV6 and HPyV7, respectively) were investigated. While they display an architecture similar to other PyV VP1 structures, their receptor binding pockets are altered so drastically that sialic acid can no longer be bound; this finding was confirmed by cell binding experiments and nuclear magnetic resonance spectroscopy. It is yet unknown which receptors are engaged by HPyV6 and HPyV7, and these two viruses have not yet been linked to disease. Due to their architecture, they have been grouped together with KI polyomavirus and WU polyomavirus in the recently established genus wukipolyomavirus.

Herpes simplex virus type 1 (HSV-1) is one of the most common human pathogens. Once transmitted, it establishes a lifelong, persistent infection and frequently erupts in orolabial cold sores. In some cases, HSV-1 can migrate to the brain and cause life-threatening encephalitis.

To gain entrance to a host cell, glycoprotein C (gC) of HSV-1 attaches to heparan sulfate (HS) on the cell surface, followed by membrane fusion, which is in turn mediated by glycoproteins B, D, H, and L. The gC-HS interaction has been shown to be non-critical for HSV-1 infection, but its abolishment drastically reduces infectivity. Furthermore, gC has been found to mediate release of viral progeny from a parent cell, and it also plays a defensive role by inhibiting the C3b complement system. Structures for the other HSV-1 glycoproteins are available, but no such information exists for gC. This work introduces an expression construct that can be purified from mammalian cell culture for structure determination by X-ray crystallography. Although no crystals are yet available, the purification strategy lays an important foundation for structural studies of this multifunctional protein.

ZUSAMMENFASSUNG

Diese Dissertation befasst sich mit den Zellanheftungsprozessen von Viren und konzentriert sich dabei auf Polyomaviren und das humane Herpesvirus 1, allgemein bekannt als Herpes simplex Virus Typ 1.

Polyomaviren (PyVs) gehören zu einer ständig wachsenden Familie von Doppelstrang-DNA Viren, die ein breites Spektrum von Fischen, Säugetieren und Vögeln infizieren. Während sie in gesunden Wirten in der Regel asymptomatisch bleiben, können sie ernste Erkrankungen bei immungeschwächten Personen verursachen und werden dabei mit Atemwegserkrankungen, neurodegenerativen Symptomen, Nephropathie und Krebs in Verbindung gebracht.

Die meisten PyVs binden an ihre Wirtszellen über sialylierte Kohlenhydratrezeptoren, die Teil von Lipiden, Proteinen oder beiden sein können. Diese Bindung wird durch das Haupt-Kapsidprotein VP1 vermittelt, von dem 360 Kopien die ikosahedrale Hülle des Virus bilden. VP1 besitzt eine konservierte β -barrel Kernstruktur, aber die loops, die die β -Stränge verbinden, sind sehr unterschiedlich und definieren Rezeptorspezifität und virale Pathogenität; selbst kleine Veränderungen können drastische Auswirkungen auf den viralen Lebenszyklus haben. Ein Beispiel hierfür ist das murine Polyomavirus (MuPyV). Es ist seit langem bekannt, dass der Austausch einzelner Aminosäuren in den loops, die die Rezeptor-Bindungstasche von MuPyV VP1 bilden, Tumorbildung und die Verbreitung des Virus drastisch verändern; die molekularen Grundsätze hierfür sind nicht bekannt. Drei MuPyV Stämme mit unterschiedlich starker Pathogenizität, RA, PTA und LID, wurden in dieser Arbeit untersucht. Zusätzlich zu den bereits etablierten Rezeptoren, den Gangliosiden GD1a und GT1b, wurde ein dritter Rezeptor, GT1a, identifiziert; seine Struktur im Komplex mit MuPyV VP1 wurde durch Röntgenkristallographie gelöst. Darüber hinaus wurden Strukturen von GD1a sowie eines weiteren Glykans, DSLNT, im Komplex mit den verschiedenen MuPyV Stämmen untersucht, um potentiell je nach Stamm veränderte Bindungsarten aufzuklären. Interessanterweise beeinflusst keine der eingeführten Aminosäuremutationen die Bindungsmodi für die untersuchten Glykane. Durch kristallographische Affinitätsmessungen konnte allerdings gezeigt werden, dass die Stämme eine veränderte Affinität für ihre Interaktionspartner aufweisen, was nahelegt, dass die Regulierung der Pathogenizität bei RA, PTA und LID weit komplizierter ist, als bisher angenommen.

In einem anderen Projekt wurden die Auswirkungen von Aminosäuremutationen in PyV VP1 im simian virus 40 (SV40) untersucht. Hier wurden drei Mutanten isoliert, die keine Bindung mehr an den klassischen SV40-Rezeptor GM1 zeigen. Darüber hinaus zeigt eine Mutante veränderten Tropismus. Es ist nicht bekannt, welche Ganglioside diese Mutanten verwenden, aber die Modellierung der mutierten VP1 Aminosäuren zeigt, dass die Bindung an GM1 aus sterischen Gründen nicht mehr möglich ist.

Schließlich wurden die VP1 Strukturen der neu entdeckten menschlichen Polyomaviren 6 und 7 (HPyV6 und HPyV7) untersucht. Während die grundlegende Form ihrer VP1 Pentamere ähnlich ist wie in anderen PyV VP1 Strukturen, sind ihre Rezeptorbindungstaschen so drastisch verändert, dass Sialinsäure nicht mehr gebunden werden kann; dieser Befund wurde durch Zellbindungsexperimente und

magnetische Kernresonanzspektroskopie bestätigt. Es ist noch nicht bekannt, welche Rezeptoren durch HPyV6 und HPyV7 gebunden werden; ferner wurden diese beiden Viren bisher noch nicht mit Krankheiten in Verbindung gebracht. Aufgrund ihrer stark veränderten VP1 Architektur wurden sie zusammen mit dem KI Polyomavirus und dem WU Polyomavirus in der kürzlich etablierten Gattung Wukipolyomavirus zusammengefasst.

Herpes simplex Virus Typ 1 (HSV-1) ist eines der häufigsten menschlichen Pathogene. Einmal übertragen, etabliert es eine lebenslang persistierende Infektion und bricht regelmäßig aus, begleitet von orolabialen Fieberbläschen. In einigen Fällen kann HSV-1 ins Gehirn wandern und dort eine lebensbedrohliche Enzephalitis auslösen.

Um Zugang zu einer Wirtszelle zu erhalten, bindet das Glycoprotein C (gC) von HSV-1 an Heparansulfat (HS) auf der Zelloberfläche, gefolgt von Membranfusion, die ihrerseits durch die Glycoproteine B, D, H und L vermittelt wird. Es wurde gezeigt, dass die Bindung von gC an HS nicht zwingend erforderlich für die HSV-1-Infektion ist, aber die Infektiosität stark abnimmt, wenn diese Wechselwirkung nicht stattfindet. Weiterhin wurde herausgefunden, dass gC eine Rolle bei der Freisetzung von neuen Viruspartikeln aus der Wirtszelle spielt; zuletzt ist es wichtig für den Schutz des Virus, indem es das C3b Komplementsystem inhibiert. Strukturen der anderen HSV-1 Glykoproteine sind bekannt, aber solche Informationen sind nicht für gC verfügbar. In dieser Arbeit wurde ein Expressionskonstrukt generiert, das Proteinreinigung aus der Säugetierzellkultur zur Strukturbestimmung durch Röntgenkristallographie erlaubt. Obwohl noch keine Kristalle verfügbar sind, stellt die hier gezeigte Reinigungsstrategie eine wichtige Grundlage für strukturelle Studien dieses multifunktionalen Proteins dar.

ABBREVIATIONS

A ₂₆₀	Absorption at 260 nm
A ₂₈₀	Absorption at 280 nm
aa	Amino acids
APS	Ammonium persulfate
Amp	Ampicillin
<i>Bam</i> HI	<i>Bacillus amyloliquefaciens</i> H
BKPyV	BK polyomavirus
bp	Base pairs
BSA	Bovine serum albumin
BSc	Bachelor of Science
Ccw	Counter-clockwise
CHO	Chinese hamster ovary
CNS	Central nervous system
CS	Chondroitin sulfate
Cw	Clockwise
Da	Dalton
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
Fc	Fragment, crystallizable
Fig.	Figure
FMDV	Foot-and-mouth-disease virus
G418	Geneticin®
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	<i>N</i> -acetyl galactosamine
GBP	Glycan-binding protein
Glc	glucose
GlcNAc	<i>N</i> -acetyl glucosamine
gB	Glycoprotein B
gC	Glycoprotein C
gD	Glycoprotein D
gH	Glycoprotein H
gK	Glycoprotein K
gL	Glycoprotein L
GPI	Glycophosphatidylinositol
HIV	Human immunodeficiency virus
HPyV6	Human polyomavirus 6
HPyV7	Human polyomavirus 7
HPyV9	Human polyomavirus 9
HPyV10	Human polyomavirus 10
HPyV12	Human polyomavirus 12

HS	Heparan sulfate
HSE	Herpes simplex encephalitis
HSV-1	Herpes simplex virus type 1
HSV-2	Herpes simplex virus type 2
HVEM	Herpes virus entry mediator
ICTV	International Committee on Taxonomy of Viruses
IgG	Immunoglobulin G
JCPyV	JC polyomavirus
KIPyV	Karolinska Institute polyomavirus
LacCer	Lactosylceramide
LAT	Latency-associated transcript
LPyV	B-lymphotropic polyomavirus
LSTc	Lactoseries tetrasaccharide c
mAb	Monoclonal antibody
MAG	myelin-associated glycoprotein
MALDI-MS	Matrix-assisted LASER desorption/ionization mass spectrometry
MCPyV	Merkel cell polyomavirus
MEF	Mouse embryonic fibroblast
miRNA	Micro RNA
MWPyV	Malawi polyomavirus
MuPyV	Murine polyomavirus
Neu5Ac	<i>N</i> -acetyl neuraminic acid, sialic acid
Neu5Gc	<i>N</i> -glycolyl neuraminic acid
NJPyV	New Jersey polyomavirus
NM-IIA	Non-muscle myosin IIA
<i>NotI</i>	<i>Nocardia otitidis-caviarum</i>
P.i.	Post infection
PCR	Polymerase chain reaction
PFU	Plaque-forming unit
PILR α	Paired immunoglobulin-like type 2 receptor- α
PML	Progressive multifocal leukoencephalopathy
PSA	Polysialic acid
<i>Pvul</i>	<i>Proteus vulgaris</i>
PyV	Polyomavirus
RMSD	Root mean square deviation
RNA	Ribonucleic acid
Rpm	Repetitions per minute
RT	Room temperature
SAP	Shrimp alkaline phosphatase
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis.
SEC	Size exclusion chromatography
<i>Smal</i>	<i>Serratia marcescens</i>
SPR	Surface plasmon resonance
STD-NMR	Saturation transfer difference nuclear magnetic resonance spectroscopy

STLPyV	St. Louis polyomavirus
SV40	Simian virus 40
Tab.	Table
<i>Taq</i>	<i>Thermus aquaticus</i>
TBS	Tris-buffered saline
TEV	Tobacco etch virus
TG	Trigeminal ganglion
TMV	Tobacco mosaic virus
TNF	Tumor necrosis factor
TSPyV	<i>Trichodysplasia spinulosa</i> polyomavirus
VLP	Virus-like particle
VP1	Polyomavirus major capsid protein 1
VP2, VP3	Polyomavirus minor capsid proteins 2 and 3
v/v	Volume per volume
WT	Wild type
WUPyV	Washington University polyomavirus
w/v	Weight per volume
<i>XhoI</i>	<i>Xanthomonas holcicola</i>

The common abbreviations are used for chemical and physical units, the essential amino acids, and bases occurring in DNA and RNA. Gangliosides are abbreviated in the style of L. Svennerholm [1].

1. INTRODUCTION

1.1 THE DISCOVERY OF VIRUSES AND BIRTH OF VIROLOGY

Throughout history, humans have been faced with diseases and the need to overcome them. Countermeasures against various sicknesses were developed early on, such as variolation against smallpox, a practice that originated in China and India in the early 11th century and was introduced to the western world in the 18th century [2]. This practice led to groundbreaking discoveries by Edward Jenner in the 1790s, who found that treatment of patients with cowpox material induced only very mild symptoms but protected these patients against the much more dangerous smallpox disease ([3], reviewed in [4]). Still, despite the somewhat successful attempts to control disease, the occurrence of illnesses such as cholera or the Bubonic plague was attributed to 'miasma' (ancient Greek for pollution), and the causative agents and the workings behind them remained elusive. With the discovery of the light microscope, Antonie van Leeuwenhoek (1623-1732) introduced an entire spectrum of new organisms, too small to be seen by the naked eye, to the world. The existence of these microorganisms was widely accepted by the early 19th century, and Louis Pasteur (1822-1895) not only demonstrated that sterilized liquids can remain free of any microorganism, provided that airborne entry is prevented, he also associated various groups of microorganisms with specific processes. The German physician Robert Koch (1843-1910) was the first to link disease with microorganisms by identifying pathogens such as *Bacillus anthracis* [5] and *Mycobacterium tuberculosis*, which awarded him the Nobel Prize in Physiology and Medicine in 1905 (reviewed in [6]). The pathological understanding that resulted from Koch's discoveries was an important step in the struggle to regulate disease, but by the end of the 19th century another pathological agent that defied the rules established by Koch was identified.

In 1892 the Russian scientist Dimitrii Ivanovsky (1864-1920) discovered that the pathogen that causes tobacco mosaic disease could not be retained by filters that were commonly used to remove bacteria. These findings could be reproduced in 1898 by Dutch researcher Martinus Beijerinck (1851-1931), and he concluded that the causative agent of the disease was small enough to pass through bacterial filters. Friedrich Löffler and Paul Frosch, both students of Robert Koch, obtained the same results for the pathogen of foot-and-mouth-disease. The fact that these new pathogens could only propagate in live host organisms and not in liquids that support bacterial growth further set them apart from other microorganisms and prompted Beijerinck to name the agent responsible *contagium vivum fluidum*. The name was later modified to virus (Latin for slime or noxious fluid), and the discipline of virology was born. At the time, it defined viruses as obligatory internal parasites, which replicate inside a cell and are characterized by their relatively small size and simplicity.

Shortly after the discovery of the tobacco mosaic virus (TMV) and the foot-and-mouth-disease virus (FMDV), the first human virus, which causes yellow fever, was identified. With the advancement of scientific instruments our knowledge of viruses expanded. The 2014 release of the International Committee on Taxonomy of Viruses (ICTV) lists 104 virus families (www.ictvonline.org/virus/Taxonomy.asp) containing 3,186 virus and viroid species (www.ictvonline.org/virusTaxInfo.asp).

1.2 STRUCTURAL FEATURES OF VIRUSES

With the increasing number of discovered viruses, the interest in their structures grew stronger as well. This curiosity could be satisfied with technological advancements such as the invention of the electron microscope in the 1930s and the emergence of X-ray crystallography. For example, in 1935 crystals of TMV were produced and shown to be infectious [7], and in 1939 Kausche *et al.* obtained the first electron micrographs of the virus [8]. The accumulating data on virus structures again demonstrated their simplicity and elegance, and that much of their life cycle and pathogenicity is dictated by their shape and morphology. In accordance with the Baltimore classification system [9], animal viruses are generally grouped into DNA and RNA viruses. DNA viruses include enveloped complex representatives such as poxviruses, while some of the most notable enveloped RNA viruses of helical shape are orthomyxoviruses (influenza), paramyxoviruses (measles), and filoviruses (Ebola). It is interesting to note that the majority of both DNA and RNA viruses feature protein capsids with icosahedral symmetry, some of which are naked (such as polyoma- and picornaviruses) and others are, again, enveloped by a membrane (for example, herpesviruses and retroviruses), with proteins embedded into the envelope.

Icosahedrons, the principal building blocks of icosahedral capsids, are polyhedrons defined by 20 equilaterally triangular faces, 30 edges, and 12 vertices. They are formed by applying three symmetry operators onto the structure: (1) a threefold axis on the triangular face, (2) a twofold axis on the edge that links two triangles with one another, and (3) a fivefold axis on the vertex (Fig. 1).

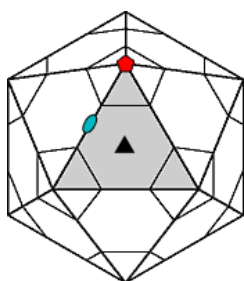


Fig. 1: General symmetry of an icosahedron. The structure is defined by 20 equilateral triangles (one shaded grey) that are arranged on the surface. These faces possess threefold symmetry (black triangles), and they are related to each other by twofold (blue ovals) symmetry axes between neighboring triangles and fivefold (red pentagons) symmetry axes on the vertices of the icosahedron. Figure modified from the VIPERdb server (<http://viperdb.scripps.edu>) [10, 11].

Since the icosahedron is composed of 20 triangles, the maximum number of elements used to build a capsid would be 60 proteins, arranged in trimers. However, viruses with larger genomes require larger capsids for packaging, which results in capsids that consist of more than 60 subunits and usually more than one kind of structural protein. In these capsids, the subunits form contacts that are still similar but no longer strictly equivalent. This concept is known as quasi-equivalency [10, 12], and it leads to capsids that bear pentameric structures at their vertices and a variable number of hexamers in between. The exact number of hexamers on each triangular face of the icosahedron as well as their orientation is dictated by the triangulation number, T . It is defined by the equation

$$T = h^2 + hk + k^2 \quad (1)$$

h and k ($h > k$) describe the size of the capsid in terms of how many sixfold symmetry entities (i.e. hexamers) one has to cross until the next pentamer is reached. As such, an icosahedral capsid always consists of 12 pentamers and $(T-1) \times 10$ hexamers. Fig. 2 displays several capsids with various triangulation numbers. Note that starting at $T=7$ there are sometimes two possibilities how the pentamers can be reached, which is why some (but not all) T numbers are amended by the qualifiers *dextro* (right) or *laevo* (left), e.g. T7l.

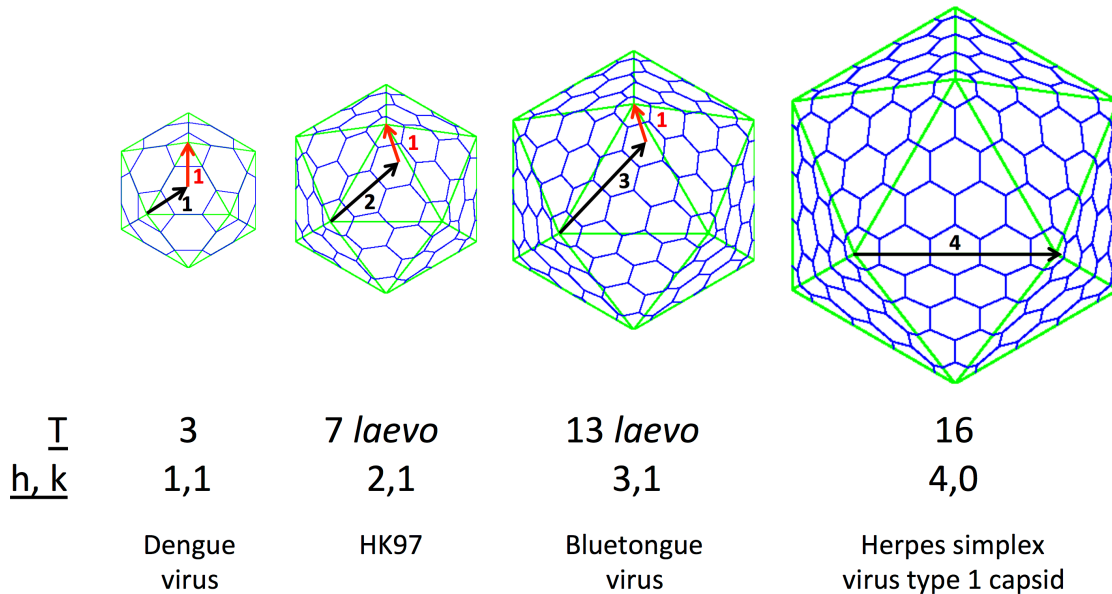


Fig. 2: Capsids with various T numbers and examples. The capsids are composed of 12 pentamers and a variable number of hexamers, depending on the triangulation number. The values h (black arrows) and k (red arrows) describe the number of hexamers that need to be crossed in order to get from one pentamer to the next. In larger capsids, two orientations are possible, termed *dextro* and *laevo*. Note that the $T=7$ and $T=13$ capsids are in the *laevo* conformation. Such a distinction is not possible for the other two capsids depicted here. Icosahedrons were taken from The Icosahedral Server (www.viperdb.scripps.edu). Figure adapted from [13].

Even though the quasi-equivalence model suggests pentamers and hexamers, the structural subunits on the hexameric sites can have other stoichiometries. One example for this is the family *Polyomaviridae*, whose members possess capsids that consist entirely of pentameric capsomeres [14]. The pentamers located at the 12 vertices of the capsid adhere to fivefold symmetry of the capsid and are termed strict, while the remaining pentamers only display fivefold symmetry with respect to each capsomer and are termed local pentamers (Fig. 3A) [15, 16].

The capsid can feature modifications, such the formation of isopeptide bonds, which greatly strengthen the durability of the capsid. Such isopeptide bonds have been observed in the bacteriophage HK97, and the capsid displays the properties of chain mail [17] as well as the ability to withstand an inner pressure of up to 50 atmospheres [18] (Fig. 3B). This type of topology is also found in herpesviruses [19].

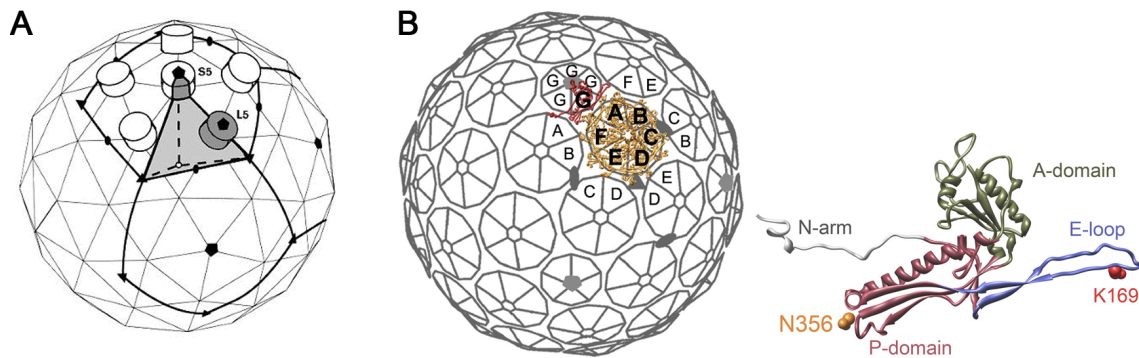


Fig. 3: Modifications in icosahedral capsids. Ovals, triangles, and pentagons depict twofold, threefold, and fivefold symmetry axes, respectively. **A** Pentameric capsomers of polyomavirus capsids. Pentamers situated on the vertices of the capsid adhere to global fivefold symmetry and are termed strict pentamers (S5), while the remaining pentamers still follow the icosahedral symmetry, but display a local fivefold symmetry and are therefore termed local pentamers (L5). Panel taken from [16], © Current Biology Ltd ISSN 0969-2126. **B** Inter-capsomer crosslinking in the Head-II structure of bacteriophage HK97. The capsid of the phage consists of 360 copies of the structural protein gp5*. In its final maturation stage, an isopeptide bond between K169 of one protein molecule and N356 of another fortifies the capsid's chain mail-like structure, which greatly enhances its durability. Panel modified from [20], © 2008, Elsevier Ltd.

1.3 VIRAL CELL ATTACHMENT AND ENTRY STRATEGIES

The principal goal of any virus is to attach to a host cell so that it may enter and hijack the cellular machinery to some degree in order to replicate its viral genome and produce progeny virus. When a virus and a potential host cell come into close contact with one another, the virus first has to latch onto the cell before it can enter.

Viral attachment to cells is generally mediated by viral receptors and, in some cases, attachment factors. Practically any molecule located on the cell surface can be exploited to serve as a receptor, and their function usually has nothing to do with the pathogen. Receptors are usually engaged specifically, which leads to subsequent steps that result in virus internalization by various endocytotic pathways [21]. Even though the interactions with receptors are usually very specific, the affinities can vary. One example are polyomaviruses, which have millimolar affinities for their glycan receptors [22, 23], a fact that is counteracted by the employment of multiple binding sites, so that avidity outweighs weak affinity. While viral receptors constitute a vast spectrum of molecules, it is interesting to note that proteins from the immunoglobulin superfamily [24] and integrins [25] are very widely used groups.

In contrast to receptors, attachment factors usually do not directly promote virus internalization, and their binding is mostly non-specific; they accumulate and concentrate viral particles on the cell membrane and thereby facilitate and promote entry and infection. Many viruses are known to bind more than one receptor and/or attachment factor.

Once they have bound an infectious receptor, the majority of animal viruses exploit the endocytotic pathways of the host cell to establish infection. This is the case for non-enveloped as well as enveloped viruses. While the latter may just fuse with the cell membrane, mediated by binding of attachment and/or fusion proteins, they still require a transportation system to circumnavigate further obstacles within

the cell [26, 27]. A number of endocytotic pathways, namely clathrin-dependent entry, macropinocytosis, and lipid rafts, will be discussed briefly.

Clathrin-mediated endocytosis occurs through vesicles that are covered with a cage formed by multiple copies of the triskelion-shaped protein clathrin [28]. This form of endocytosis is fast and has a high capacity [29], and it is usually employed to internalize compounds for recycling or degradation. Viruses exploit this mechanism by binding to receptors with appropriate internalization signals in clathrin-coated pits on the cell membrane. The vesicles are about 120 nm in size, but larger virus particles can also be accommodated [30].

Macropinocytosis is an internalization process that is usually hijacked by larger viruses, such as vaccinia or Ebola viruses [21, 31]. It is controlled by a complex signaling cascade, initiated by the activation of tyrosine kinases that results in a change in actin dynamics and the formation of tongue-like protrusions that envelop the cargo on the cell surface and internalize it in large vacuoles. As the method of choice for removing apoptotic debris in tissues, macropinocytosis, unlike phagocytosis, does not invoke inflammatory or innate immune responses.

Viral entry by caveolae and lipid rafts has first been observed for polyomaviruses. Caveolae are a special type of lipid rafts of about 50 – 100 nm and are found in many different cell types. Caveolin and cavin proteins [32] associate with lipid rafts and form oligomers that create invaginations in the cell membrane. The activation of tyrosine kinases, increased levels of cholesterol, and the insertion of the caveolin scaffolding domain into the cell membrane trigger vesicle formation, followed by fission, which is mediated by the GTPase, dynamin II (reviewed in [33, 34]). Interestingly, there are also variations of this pathway that are independent of caveolin and/or of dynamin [35]. It was believed that endocytosed particles were routed to an organelle called the caveosome; it has been shown, however, that this is actually a late endosome modified by overexpressed caveolin-1 awaiting degradation [36].

In the following chapters, the attachment and cell entry strategies of two virus types will be further discussed: (1) polyomaviruses and (2) herpesviruses.

1.4 POLYOMAVIRUSES: GENERAL INFORMATION, CELL ATTACHMENT, AND ENTRY

Polyomaviruses (PyV) were first identified in 1953, when Ludwik Gross discovered by accident that inoculation of mice with leukemic extracts also introduced bilateral neck tumors [37]. Five years later, Sarah Stewart and Bernice Eddy isolated the causative agent and termed it polyoma (Greek, many tumors) [38, 39]. The *Polyomaviridae* form an ever-expanding family, whose members infect a large range of birds, fish, and mammals, including humans. Apart from the first discovered PyV, murine polyomavirus (MuPyV, see above), simian virus 40 (SV40) [40] gained fame when a batch of polio vaccine was contaminated with the virus in the 1960s [40, 41]. In the 1970s, the human polyomaviruses BK (BKPyV) [42] and JC (JCPyV) [43] were discovered and identified as the causative agents for BK-associated nephropathy [44] and progressive multifocal leukoencephalopathy (PML) [45], respectively. More recently, two human polyomaviruses were discovered at the Karolinska Institute (Stockholm, Sweden, [46]) and the Washington University School of Medicine (St.

Louis, USA, [47]), followed by the identification of Merkel cell polyomavirus (MCPyV) [48], which is associated with a rare but aggressive form of skin cancer (reviewed in [49]). In 2010, two new human polyomaviruses termed HPyV6 and HPyV7 were found to be shed from skin, together with MCPyV [50]. In the same year, another human polyomavirus, which is directly associated with the skin disease *Trichodysplasia spinulosa*, was identified (TSPyV) [51], followed in subsequent years by human polyomavirus 9 (HPyV9) in 2011 [52], human polyomavirus 10 (HPyV10) [53-55], which is also known as Malawi polyomavirus (MWPyV [54]), in 2012, St. Louis polyomavirus (STLPyV) and human polyomavirus 12 (HPyV12) in 2013 [56, 57], and finally New Jersey polyomavirus (NJPyV) in 2014 [58]. Today, the family *Polyomaviridae* contains roughly 100 species. Based on host range, genetic makeup, and overall genome sequence, they are distributed in the three genera avipolyomaviruses, orthopolyomaviruses, and wukipolyomaviruses [59].

PyVs are small viruses, whose T=7d icosahedral capsids are roughly 45-50 nm in size and encapsidate a dsDNA genome of about 5,000 bp. In its early region, the genome encodes several non-structural proteins, such as small t-antigen and large T-antigen, which are responsible for DNA replication and cell transformation. Rodent polyomaviruses, such as MuPyV and hamster polyomavirus, also express middle T-antigen [60], and agnoprotein is produced by BKPyV, JCPyV, and SV40 ([61], reviewed in [62, 63]). In the late region, the PyV genome also encodes three structural proteins, the major capsid protein VP1, and the minor capsid proteins VP2 and VP3, although the latter is not present in all PyVs. The capsid is formed by 72 pentameric capsomeres of VP1 [15, 64], with one copy of VP2 or VP3 bound in the central pore of each pentamer [65, 66]. The minor capsid proteins share the same N-terminus, and VP2 is often myristylated [67]. SV40 additionally encodes VP4, a viroporin that facilitates viral release [68, 69].

The structure of VP1 is conserved among polyomaviruses, and it features a prominent β -barrel containing a jelly-roll motif comprised of two antiparallel β -sheets, termed BIDG and CHEF. While the β -strands make up the majority of the VP1 core structure, the loops connecting these strands are highly divergent among the different species and define the receptor specificity, the entry pathway, and hence, the pathogenicity. The capsids are formed by virtue of the C-terminal arms of VP1. Since pentameric structures have to accommodate penta- and hexavalent positions on the capsid, the associations between them are quite divergent, depending on their position. The C-terminus begins with a connector, followed by an α -helix that forms no contacts with the pentamer it originates from and is very exposed. Either two or three of these helices associate with each other and tie together adjacent pentamers. No such helices have been observed in monomers connected by a quasi-twofold symmetry axis, but their C-termini form non helical structures [70]. The final part of the C-terminal arm invades a neighboring monomer, and its penultimate segment forms a β -strand that aligns itself with the BIDG sheet [70]. The extreme C-terminal region reaches into the clockwise neighboring monomer, where it is held in place by that monomer's N-terminal arm, strand A, and helix 3A [70]. The capsid is further stabilized by disulfide bridges that involve cysteine residues located in the CD-loop [22].

Most PyVs initiate infection of a host cell by attachment of VP1 to glycans on the cell surface that are linked to lipids (for example, in the context of gangliosides,

see section 1.6.1) or to proteins. In most reported cases, these glycans feature sialic acid (Fig. 4A, see also section 1.6.1.1); in fact, even before it was established that MuPyV and SV40 recognize gangliosides [71], it was known that MuPyV requires sialic acid that is [α -2,3]-linked to galactose in order to establish infection [72, 73]. Depending on the virus species, sialic acid can be recognized in an [α -2,3], [α -2,6], or [α -2,8] linkage, and a number of PyVs have been shown to tolerate more than one interaction partner, bearing more than one linkage. Prominent examples are BKPyV, whose VP1 binds the b-series gangliosides GD2, GD3, GD1b, and GT1b [74], as well as MuPyV, which recognizes GD1a, GT1b [71], and GT1a [75]. In contrast to this, SV40 is a rigorously examined polyomavirus that exclusively binds the GM1 ganglioside [71]. Its affinity is increased for simian GM1, which carries both *N*-glycolyl neuraminic acid (Neu5Gc) and *N*-acetyl neuraminic acid (Neu5Ac), while human GM1 only bears Neu5Ac [76].

The glycan binding site is usually formed by the BC1-, BC2-, DE-, and HI-loops on top of a VP1 monomer. In many cases, residues from the neighboring, clockwise (cw) monomer also contribute to binding. Prominent interactions between VP1 and engaged glycans are formed with sialic acid and, more specifically, its carboxylate group [77]. The only known exception so far is TSPyV, which not only features a sialic acid binding site at a different location, but also engages it in a different binding mode [78] (Fig 4B, C).

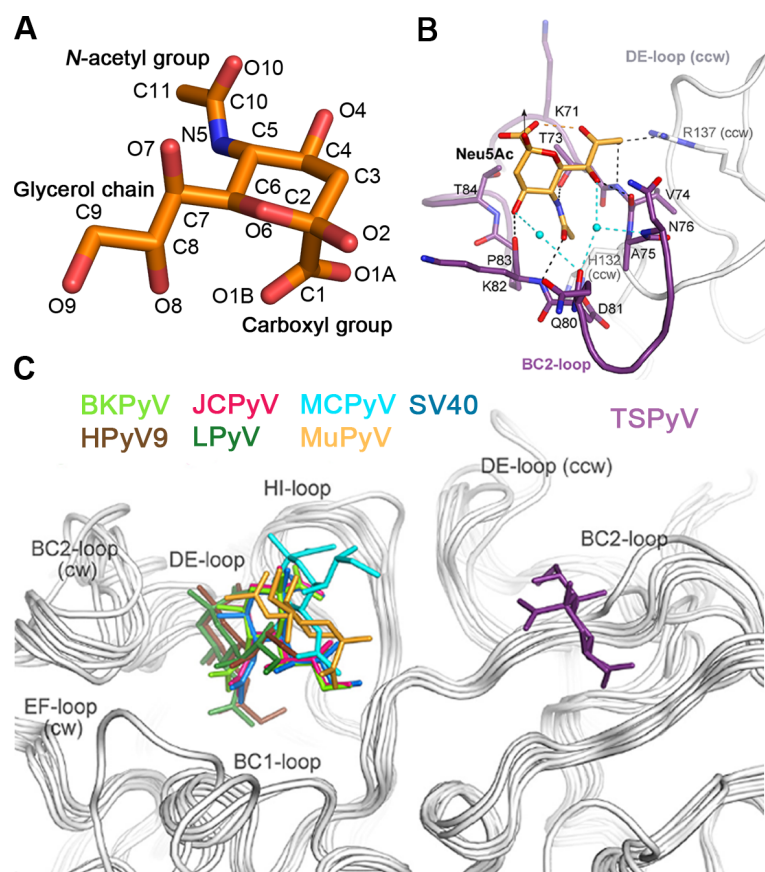


Fig. 4: Engagement of sialylated carbohydrates by polyomaviruses. **A** The structure of *N*-acetyl neuraminic acid, which caps glycans recognized by various polyomaviruses. Panel adapted from [79]. **B** Engagement of sialic acid by TSPyV VP1. While in all other complexed PyV VP1 structures elucidated so far, Neu5Ac is primarily engaged via its carboxyl group, TSPyV forms no contacts with these atoms

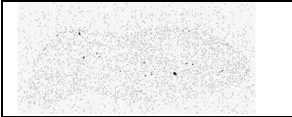



and rather binds to the O4 hydroxyl group, the glycerol chain, and the *N*-acetyl group. **C** Comparison of the binding sites of polyomaviruses. BKPyV, HPyV9, JCPyV, B-lymphotropic polyomavirus (LPyV), MCPyV, MuPyV, and SV40 engage sialic acid in different orientations, but at the same site on VP1, while the binding site in TSPyV is shifted towards the counterclockwise (ccw) VP1 monomer. Panel B taken from, and panel C modified from [78], © 2014, Ströh *et al.*

Even though the core structure of the VP1 pentamer is highly conserved across the different species and glycans are engaged by only few amino acids, the interactions are surprisingly specific. In the past it has been shown that even exchanges of single amino acids in the binding pocket can have profound effects on ligand recognition and pathogenicity in various PyV species. Two examples, JCPyV and MuPyV, will be explained in more detail.

The JCPyV Mad-1 strain binds the glycan lactoseries tetrasaccharide c (LSTc) [80]. In healthy individuals, JCPyV, like many polyomaviruses, persists in the kidney [81] and remains asymptomatic. However, in immunocompromised individuals, such as those infected with HIV or patients suffering from AIDS or autoimmune diseases, the virus spreads to the central nervous system (CNS) and infects glial cells, including oligodendrocytes and astrocytes [82, 83], which causes the fatal demyelination disease progressive multifocal leukoencephalopathy (PML). Isolates from PML patients show mutations in the VP1 binding site, namely L54F, S266F, S268F, S268Y, and others. Interestingly, these residues all form contacts with sialic acid [80]. It was shown that JC viruses with these mutations are not infectious, cannot bind to glial cells, and can no longer bind LSTc due to steric clashes [84]. It has been hypothesized that these mutations shift the specificity of JCPyV from sialic acid to another receptor [85] and also increase its resistance against antibody neutralization [86, 87].

MuPyV was the first polyomavirus to be identified, and it induces tumors in mice. The frequency and tissue tropism depend greatly on the strain [88], and three strains have been in focus of ongoing research: (1) the laboratory-derived strain RA [89], which only induces singular tumors and only in mesenchymal tissues; (2) the naturally occurring PTA strain [90, 91], which is found in feral mice [92], induces a large number of tumors, which occur in mesenchymal as well as in epithelial tissues; (3) LID, a virulent laboratory isolate strain [93, 94], which causes death in young mice within six weeks due to kidney failures and brain hemorrhages [95]. Sequencing of these strains mapped the crucial differences to two amino acids in the receptor binding pocket of MuPyV VP1, namely a glycine-to-glutamate exchange and a valine-to-alanine exchange [88, 96-99] (Tab. 1). These exchanges are sufficient to switch one strain's pathogenicity profile to that of another [97], and similar exchanges with the same consequences have also been found in MuPyV strains A2 and A3 [98]. The receptor binding properties and their consequences for the pathogenicity profiles of the RA, PTA, and LID strains will be discussed in-depth in various parts of this dissertation.

Table 1. MuPyV strains and their pathogenicity. The three strains RA, PTA, and LID, which display profoundly different disease profiles, are described in terms of their key features. Also, tumorigenicity is demonstrated in infected mice, which were sacrificed and afterwards treated with an anti-VP1 antibody from rabbits to visualize tumor development. Mouse blots taken from [97], © 1999, American Society for Microbiology.

Control	RA	PTA	LID
			
Distinctive amino acids	G91, V296	E91, V296	E91, A296
Pathogenicity	Singular tumors in mesenchymal tissues.	High tumor density of mesenchymal and epithelial origin.	Virulent. Early death due to kidney failure and brain hemorrhages.
Latency period	Long	Short	Very short

After polyomaviruses attach to their receptors on the cell surface, they are endocytosed into the cell. Interestingly, even though their overall sequence and structure is relatively similar, PyVs enter cells through different pathways before genome replication in the nucleus and, ultimately, release of progeny virus occur (Fig. 5).

MuPyV [100, 101] and SV40 [102] can enter cells through smooth, clathrin-independent vesicles. However, it was shown that MuPyV can also utilize an additional entry mechanism through a caveolae-dependent pathway [100, 103], while SV40, contrary to earlier studies, is limited to 'classical', lipid raft-mediated endocytosis [104]. SV40 and MuPyV are transported from the early to the maturing and then the late endosome, accompanied by a lowering of the pH, and subsequently transported to the ER via microtubules [100, 105]. BKPyV is also thought to enter the cell via caveolae-mediated endocytosis [106, 107], and low pH has been reported to coincide with conformational changes of the virion [108]. However, the dependency of PyV entry on pH has been a topic of debate, with some publications showing sensitivity [104] and others showing independence [105, 109].

JCPyV is known to enter cells through a different pathway. This is possibly due to the fact that it uses the pentasaccharide LSTc [80] in combination with the serotonin receptor [110] for entry, rather than gangliosides. Upon receptor binding, JCPyV is endocytosed via a clathrin-dependent mechanism [111], followed by transport of the internalized virus to early, maturing, and late endosomes, from where it is moved, like SV40, MuPyV, and BKPyV, to the ER. Here, the viruses interact with proteins related to protein disulfide isomerases (PDIs) [112-114], which expose hydrophobic loops and facilitate lipid bilayer binding. Finally, PyVs are transported to the nucleus, followed by the remaining steps of the viral life cycle. Other steps, such as the exploitation of the ER-associated degradation (ERAD) pathway [114-118] (which allows the virus to exit the ER in order to enter the nucleus via the nuclear pore complex), genome replication, progeny virus assembly, and release (Fig. 5) will not be discussed in this work.

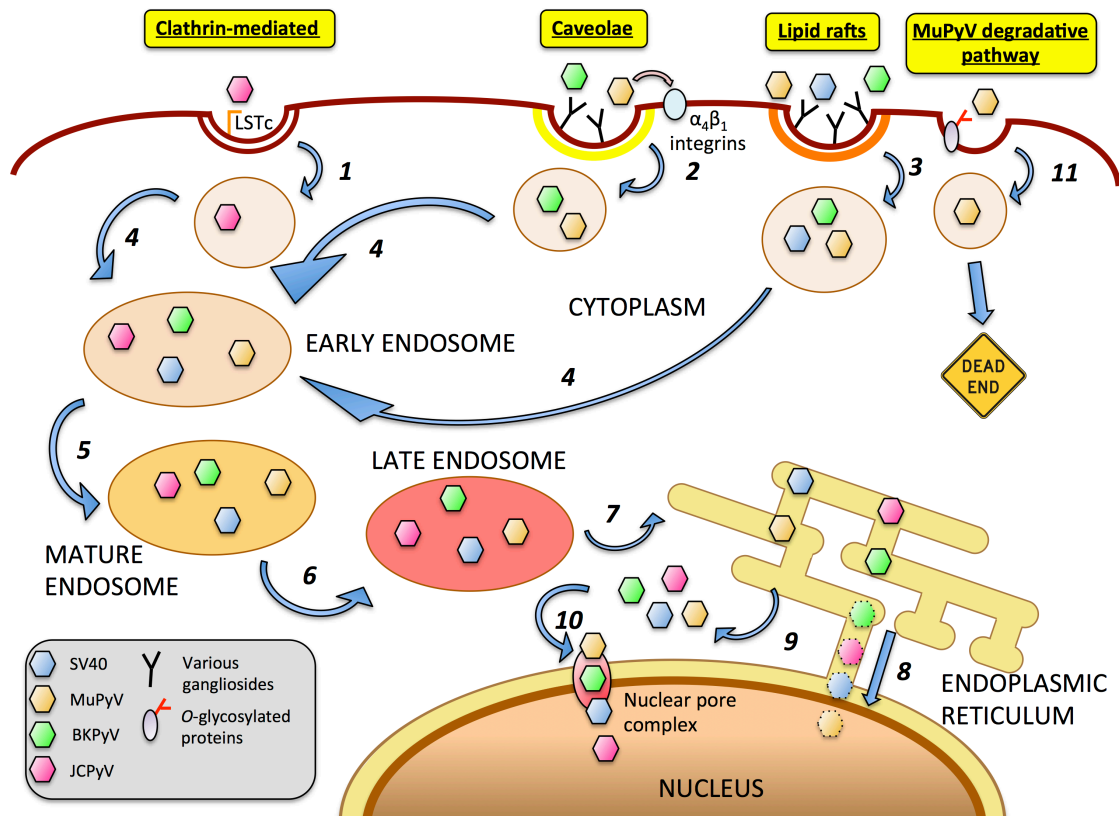


Fig. 5: Cell entry pathways of polyomaviruses. Only selected interactions are shown here. JCPyV binds LSTc and is unique in its usage of a clathrin-dependent entry pathway (1); MuPyV and BKPyV can take either caveolae (2)- or lipid raft (3)-mediated entry pathways, while SV40 is restricted to lipid rafts (3). After entry, PyVs are routed towards early endosomes (4), mature endosomes (5), and then late endosomes (6), from which they are transported to the ER by microtubules (7). This is the place where uncoating occurs (represented by dashed hexagons) and the virus enters the nucleus (8). Alternatively, PyVs can exit the ER into the cytosol via the ERAD pathway (9). From there, they can enter the nucleus via the nuclear pore complex (10). In the course of this work, hints surfaced that MuPyV can be captured by O-glycosylated proteins, internalized (11), and then be degraded without ever initiating infection. Figure adapted and modified from [79].

1.5 HERPES SIMPLEX VIRUSES AND HOW THEY ENTER A CELL

The order *Herpesvirales* contains three families of viruses that infect a large spectrum of animals: *Alloherpesviridae* infect fish and frogs, *Malacoherpesviridae* contain bivalve viruses, and the largest family, *Herpesviridae*, which was the original herpesvirus family before taxonomical reclassification by the ICTV, contains viruses that infect mammals, birds, and reptiles [119]. Altogether, there are nine species of herpesvirus that infect humans, which are distributed among three subfamilies (alpha-, beta-, and gammaherpesviruses) and various genera. This work focuses on human herpesvirus 1, which is commonly called herpes simplex virus type 1 (HSV-1) as it belongs to the genus simplexviruses in the subfamily alphaherpesviruses. Along with herpes simplex virus type 2 (HSV-2), HSV-1 is one of the most abundant pathogens in humans [120].

Herpesviruses are dsDNA viruses, whose virions are about 190 – 260 nm in size (reviewed in [121]). The linear 152 kbp genome is surrounded by an icosahedral nucleocapsid, which has an intermediate layer with T=4 and an outer layer with T=16

symmetry (Fig. 6A). The capsid is comprised of 162 capsomers, forming 11 pentons and 140 hexons from the major capsid protein VP5 and one portal complex, which in turn is formed by a dodecameric assembly of the UL6 protein. The capsid is equipped with additional proteins that decorate the hexons, form links between capsomers, and play a scaffolding role during DNA encapsidation. The capsid is surrounded by a dense layer of proteins called the tegument, whose thickness and density can vary; it contains roughly 20 proteins that are required for early steps in the viral life cycle (reviewed in [122]). Finally, the virus is enveloped by a lipid bilayer, in which twelve glycoproteins are embedded [123, 124] (Fig. 6B, C). The primary function of these proteins is mediation of attachment and entry of the virus into a host cell (see below).

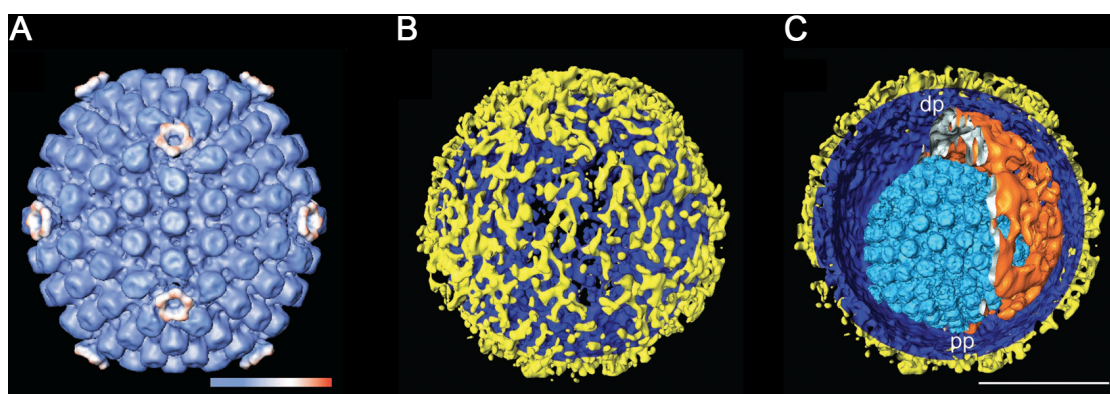


Fig. 6: Cryo-electron tomography analyses of the herpesvirus virion. **A** The nucleocapsid that contains the viral dsDNA has an outer T=16 icosahedral symmetry and is comprised of 140 hexons, 11 pentons, and one portal protein. The color bar corresponds to the radius of the capsid and ranges from 61 nm to 69 nm. **B** A whole herpesvirus virion and **C** a cutaway through half the particle. The capsid is shown in light blue and is surrounded by a layer of proteins called the tegument (orange), followed by a lipid envelope (dark blue) that is covered by glycoproteins (yellow). dp: distal pole; pp: proximal pole. The scale bar corresponds to 100 nm. Panels taken from [125], © 2003, American Association for the Advancement of Science.

Transmission of HSV-1 occurs via direct contact with body fluids or the lesions of an infected individual. HSV-1 is associated with orolabial and facial infections, while HSV-2 triggers infection of the genitals [126]. Initial infection is established in mucosal regions, where HSV-1 replicates in epithelial cells (either the eyes or the mouth region). Progeny virus is released upon cell lysis, which causes the typical sores associated with herpes infections. The liquid that fills these sores contains a very high concentration of infectious virus (over 100,000 PFU/ μ L [127]). In spite of immune responses, progeny virus manages to fuse with the termini of sensory neurons that innervate the site of primary infection, from where the capsid is moved through axons to the cell bodies of the neurons (in the case of HSV-1, this is usually the trigeminal ganglion). There, two courses of infection can be pursued. During latency, which is characterized by the absence of viral replication, the viral genome persists in a circular, chromatin-associated form in the neuronal nucleus, and only the latency-associated transcript (LTA) and microRNAs (miRNAs) are expressed. Periodically and under the influence of certain external stimuli, the latent phase is interrupted, and a lytic phase commences, which is associated with the linearization of the genome and the expression of a large number of gene products. This results in

the generation of new viruses, which are then transported back to the site of primary infection, where lesions occur (Fig. 7; reviewed in [128]). The ability to set up latency is considered a key characteristic in herpes viruses.

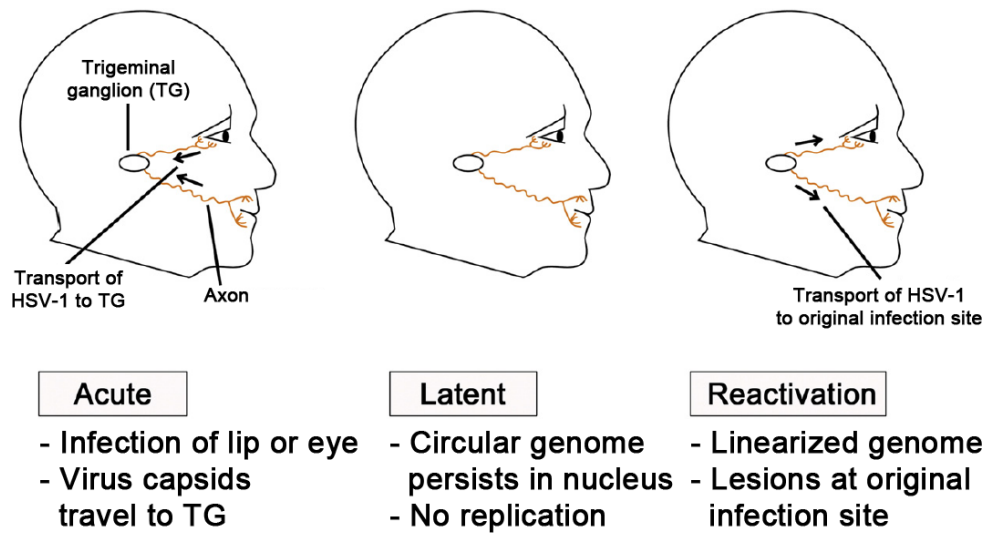


Fig. 7: Infection cycle of HSV-1. Upon first contact with the virus, acute primary infection is established, in which HSV-1 enters epithelial cells at the mouth or eyes and replicates, resulting in cell lysis and lesions. Progeny virus fuses with the sensory termini of axons, and the capsid is transported to the cell body to the trigeminal ganglion (TG). After another round of replication, the viral dsDNA genome persists in the nucleus in a circular, chromatin-associated form, marking the latent phase of infection. However, the virus can be reactivated, which causes linearization of the genome, virus replication in the nucleus, and transport of the virus back to the original infection site, where, again, epithelial cells are infected and lesions occur. Figure modified from [128], © 2009, Elsevier B.V.

While periodic reactivation of HSV-1 usually does not pose a serious health risk and is considered a troublesome aspect of the infection at worst, it is possible that the virus migrates to the brain and causes herpes simplex encephalitis (HSE; reviewed in [129]), which was first described histopathologically in 1941 [130], when inclusion bodies consistent with herpes simplex infection were identified in the brain of a newborn, and virus was isolated. Three years later, the first adult case was reported [131], and with the identification of two different herpes simplex species, this disease is now unilaterally associated with HSV-1. HSE is considered one of the most severe viral infections of the central nervous system (CNS). Investigations of the disease have shown that about one third of HSE cases are a result of primary infections, while the rest develop in the presence of antibodies. The exact trigger as well as how the virus is able to migrate to the brain are unclear; before the invention of antiviral drugs against herpes infections, such as Aciclovir, most patients died from HSE, and survivors can experience seizures, mental status changes, and motor deficits [132].

The first and most crucial step in the HSV-1 infectious cycle is attachment to a host cell. As is the case with many complex viruses (such as influenza or measles), the cell attachment and entry processes are complicated procedures that involve numerous viral proteins and a number of factors on the host cell. Some of the proteins and receptors necessary for viral entry of HSV-1 are depicted in Fig. 8 and have been reviewed in several publications [133-136].

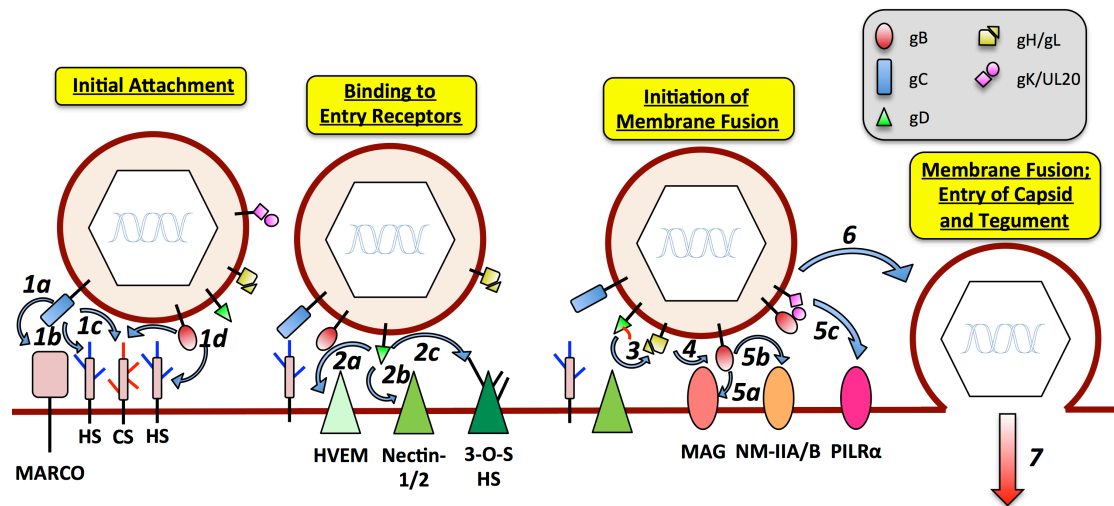


Fig. 8: Proteins and cellular receptors relevant for HSV infection. The envelope of HSV contains (among others) the glycoproteins gB, gC, gD, and the gH-gL as well as the gK/UL20 heterodimers. The virus attaches to the cell by gC binding to the scavenger protein MARCO (1a) and either HS (1b) or CS (1c). The interactions with HS or CS can also be carried out by gB (1d). Binding of gD to either HVEM (2a), nectin-1 or nectin-2 (2b), or 3-O-S HS (2c) causes the displacement of the C-terminal part of its ectodomain, which in turn recruits the gH/gL dimer (3). This activates the fusogenic activity of gB (4), which binds MAG (5a), NM-IIA or B (5b), or (with the help of the gK/UL20 heterodimer) PILRα (5c) and then triggers fusion of the viral membrane with the host cell (6). The encapsidated dsDNA genome and the tegument proteins are then released into the host cell (7). Figure adapted from [135].

Initial contact is established by gC with glycosaminoglycans (GAGs), preferably heparan sulfate (HS) [137] and the scavenger receptor MARCO [138]. The interaction with GAGs is mediated by glycoproteins B or C (gB, gC) [139, 140], and it accumulates virions on the cell surface and tethers them to the cell. In the absence of HS, chondroitin sulfate (CS) can also be used [141, 142]. Although gC is not essential for binding, its absence from the virion reduces infectivity ten-fold [139]. In addition to its attachment function, gC has also been shown to participate in immune evasion by acting as a decoy receptor for C3b of the complement system [143] (for more details on gC, see section 1.5.1). In the absence of gC, HS binding can also be carried out by gB. In some cases, HSV-1 can travel down filopodia-like membrane protrusions towards the main cell body for fusion in a movement that is aptly called “surfing” [144]. Here, gB seems to play a major regulatory role as it binds to HS in these filopodia.

The next step is cell entry, which requires four glycoproteins: gB, gD, and a gH/gL heterodimer. Entry is initiated upon binding of gD (which is essential for infection in alphaherpesviruses [145]) to one of three potential entry receptors (reviewed in [134]): (1) herpes virus entry mediator (HVEM), which is a member of the tumor necrosis factor (TNF) receptor family [146, 147]; (2) two members of the immunoglobulin superfamily named nectin-1 and nectin-2 [148, 149], which are intercellular adhesion molecules and drive the formation of tight junctions by forming highly specific and diverse dimers; (3) 3-O-S HS, which is generated in HS through reactions catalyzed by specific glucosaminyl 3-O-sulfotransferases. In the absence of other entry receptors this last receptor can enable infection by HSV-1 if cell surface HS is also present [150]. Also, a glycoprotein called B5 was shown to render resistant cells susceptible to HSV-1 infection [151], but there is no concrete

evidence yet that classifies it as a viral entry receptor for gD. Binding of any entry receptor causes conformational changes in gD, which ultimately results in the activation of the gH-gL dimer and gB. The exact mechanism is unknown, but it is suspected that the C-terminal domain of the gD ectodomain rearranges itself and allows gD to recruit the gH/gL dimer [124], which, in turn, activates gB. This activation, together with binding of gB to myelin-associated glycoprotein (MAG) [152], non-muscle myosin IIA or IIB (NM-IIA, NM-IIB) [153, 154], and, in complex with the gK/UL20 heterodimer [155], to the paired immunoglobulin-like type 2 receptor- α (PILR α) [156] triggers fusion of the viral membrane with the membrane of the host cell, which occurs either pH-independently at the cell surface or in a pH-dependent manner after endocytosis [27, 157]. The wide range of gB interaction partners provides further insight into the broad tropism of HSV-1 [153]. The majority of the fusion process is mediated by gB, although the exact details are unknown. It is interesting to note, though, that the crystal structure of gB reveals two pleckstrin-homology domains [158], which may play a role in the fusion mechanism and serve as protein interaction sites. Also, gB displays striking similarity to the fusion protein of vesicular stomatitis virus [159]. Fusion of the viral membrane with the host cell releases the viral capsid as well as the tegument proteins and initiates further stages in the viral life cycle of HSV.

1.5.1 ***Structural and Biological Properties of Glycoprotein C***

Glycoprotein C is the principal attachment protein to HS in HSV-1 [139, 160]. gC is a type 1 membrane protein that spans 511 amino acids and displays a number of noteworthy structural features (Fig. 9). It contains one transmembrane segment, followed by its C-terminus, which reaches inside the virus. One of the most notable features of gC is its glycosylation. The full-length protein contains nine consensus sites for *N*-glycosylation. In addition, the mucin-like domain spanning residues 25-127 contains numerous *O*-glycosylation sites [161, 162], which play roles in viral attachment and allow the spread from cell to cell [163, 164]. It was shown that HSV-1 mutants that lack the mucin-like region produced less progeny virus but had numerous particles trapped on the host cell surface due to weaker dissociation constants, suggesting that this region also plays a role in mediating GAG binding [165]. Mapping experiments using antibodies [166] as well as DNA sequence analyses [167] have revealed that gC also possesses two antigenic sites, termed site I (residues 307-373) and site II (residues 129-247). The amino acids responsible for attachment to HS have been mapped to antigenic site II (R143, R145, R147, T150, and G247). It was postulated that the polar arginine and threonine side chains directly interact with the sulfate and carboxylate groups of HS. Hydrophobic contacts also seem to play a role as mutations of residue I142 were shown to drastically decrease infectivity of HSV-1 [168]. Furthermore, mutation of T150 to isoleucine abolished glycosylation of N148, which also significantly hampered HS attachment [169]. The protein possesses eight cysteine residues, which are all engaged in disulfide bonds (C127-C144, C286-C347, C386-C442, and C390-C419) and help to define the antigenic sites [170]. Finally, there are four regions on gC that are reported to bind the complement protein C3b, namely residues 124-137 (I), 276-292 (II), 339-366 (III), and 223-246 (IV) [171]. In HSV-1, gC has also been reported to accelerate the decay of the C3 convertase by inhibiting properdin binding, thereby

interfering with later steps in the complement cascade [172]. These features turn gC into a multitasking protein, which not only attaches HSV-1 to its host cell but also defends it from an immune response. As such, gC has become an interesting target for the development of antibodies. One such antibody, B1C1B4 was found to be a potent inhibitor of HSV-1 infection [160] while binding to an epitope that is highly conserved among a number of different HSV-1 strains [173]. It displayed neutralizing abilities in cells with a reduced amount of glycosaminoglycans on their surface [174], which is in line with findings that this mAb binds to the GAG binding domain [168, 169, 173, 175]. The exact binding mechanism is not known, however.

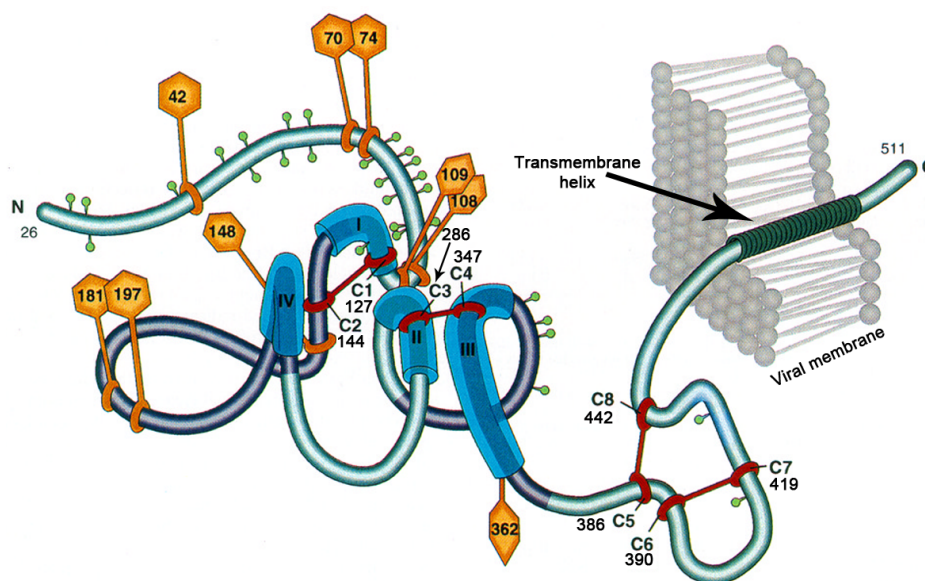


Fig. 9: Putative structure of glycoprotein C. The protein contains a large N-terminal portion facing outward and is embedded in the viral membrane by a single transmembrane helix, with a small C-terminal part reaching inside the virion. There are nine *N*-glycosylation sites (yellow hexagons) and numerous *O*-glycosylation sites (small green circles), the majority of which are located in the mucin-like region spanning residues 25-127. Antigenic sites I and II (dark blue) are established with the help of four disulfide bridges (red, connected rings), and C3 convertase is engaged by regions I-IV (light blue tubings). Figure modified from [170], © 1996, American Society for Microbiology.

1.6 THE ROLES OF GLYCANS IN VIRAL INFECTION

Carbohydrates, one of the four building blocks of life (next to lipids, proteins, and nucleic acids), undertake a large number of tasks in cells and tissues. Given their enormous complexity, difficult handling, and the complicated science underlying their structures and nomenclature, our understanding of this group of macromolecules is still lacking many details. Their functions can be divided into two general categories (reviewed in [176]): (1) structural integrity and modulatory purposes and (2) their roles as receptor molecules when they are bound by a large number of glycan-binding proteins (GBPs).

The glycocalyx on the surface of eukaryotic cells and polysaccharide structures on prokaryotes are examples of carbohydrates fulfilling protective functions. One of the most striking examples are members of the *Trypanosoma* species, which frequently alter the complex carbohydrate composition on their cell

surface to elude its host organism's immune system [177]. When attached to proteins, glycan structures can significantly enhance their solubility, and many proteins unfold due to instability and are being routed towards degradation when they are incorrectly glycosylated [178, 179].

GBP recognition can either be intrinsic or extrinsic. In intrinsic recognition, proteins bind glycans from the same organism and then fulfill their functions, for example in cell-to-cell interactions or in immune and healing responses. An example is the interaction between selectins with glycosylated ligands, which in turn mediates the interplay of blood cells and vascular cells [180, 181]. Extrinsic recognition refers to the binding of glycans by proteins of foreign organisms such as toxins, agglutinins, and viruses, which are usually highly specific, such as the binding of influenza viruses to only a subset of sialylated carbohydrates, depending on the glycosidic linkage [182-184]. In general, viruses have evolved to recognize a large number of different carbohydrates for cell attachment and entry, and this chapter will focus on gangliosides and glycosaminoglycans, which are utilized by polyomaviruses and herpesviruses, respectively.

1.6.1 *Gangliosides*

Gangliosides were first isolated from the human brain by the German chemist Ernst Klenk (1896-1971). They are ubiquitous glycosphingolipids and are most prominently expressed in the nervous system [185]. They consist of a ceramide moiety (a fatty acid attached to a sphingosine molecule, see Fig. 10A) that is embedded into the outer leaflet of the cell membrane and a carbohydrate structure, which is covalently attached to one of the ceramide's hydroxyl groups. The first two glycan moieties are glucose (Glc) and galactose (Gal), and the resulting lactosylceramide (LacCer) is the basis for most of the known gangliosides. An exception is GM4, which is based on galactosylceramide and is hence categorized as a gala-series ganglioside [186]. Almost all gangliosides are further characterized by the fact that they contain at least one moiety of Neu5Ac (see Fig. 4A and below for more information). As such, starting from LacCer, further moieties of Gal, *N*-acetyl galactosamine (GalNAc), *N*-acetyl glucosamine (GlcNAc), and Neu5Ac are added to the growing 'tree' with various linkages and by different enzymes, giving this group of glycolipids a striking amount of variability and complexity (see Fig. 10B as well as Fig. 1 in [75], which can be found at the end of this dissertation).

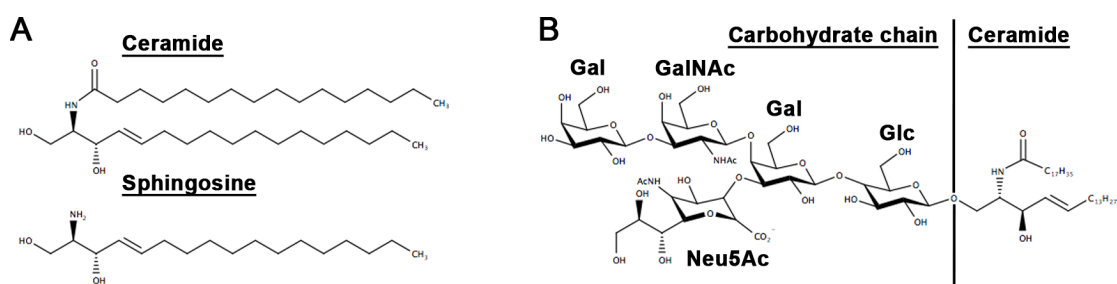


Fig. 10: General architecture of gangliosides. **A** On top, the structure of a common ceramide is shown, consisting of an 18-carbon sphingosine molecule (again shown by itself below) in an amide linkage to a 16-carbon fatty acid. Such molecules are the lipid basis for gangliosides and are the part of the glycolipid that is embedded in the membrane. Panel adapted from [187]. **B** The structure of the GM1 ganglioside is shown. Depending on the ganglioside, the carbohydrate moieties can be connected in various glycosidic linkages. Panel adapted from [186].

The biological functions of gangliosides involve cell-cell recognition and adhesion by binding to complementary molecules on opposing cell membranes as well as signal transduction. Here, they play roles in specific glycolipid-enriched microdomains [188], caveolae [189], and lipid rafts [190]. They have also been identified on nuclear membranes and are suspected to contribute to calcium homeostasis in the cell [191]. In single cells, gangliosides are not essential for survival; however, in terms of an entire organism, their absence can have profound consequences. Experiments with genetically engineered mice have demonstrated various phenotypes, such as hearing loss, myelination decreases, neurodegeneration, dysfunction of motor coordination, and death (reviewed in [186]).

1.6.1.1 *Sialic Acids as Viral Receptors*

A common component in gangliosides is sialic acid, which not only contributes to their diversity but is also a crucial interaction partner for a large number of viruses (reviewed in [192]). Sialic acids are ubiquitously expressed in all higher vertebrates and usually cap gangliosides as well as *N*- and *O*-glycosylated proteins and some GPI anchors. They fulfill a wide array of functions, such as the regulation of affinity receptors, growth, transmembrane signaling, and the protection of molecules from glycosidases and proteases (reviewed in [193]). Their remarkable diversity stems not only by the numerous modifications that can be placed on a number of carbon atoms, resulting in over 50 known variants [194], but also on the linkage they are engaged in. Neu5Ac (Fig. 4A) is usually attached to Gal, GalNAc, and GlcNAc via an $[\alpha\text{-}2,3]$ or an $[\alpha\text{-}2,6]$ linkage [195]. However, they can also be attached to each other via $[\alpha\text{-}2,8]$ or $[\alpha\text{-}2,9]$ linkages and can therefore also occupy internal positions in a glycan. Examples for such $[\alpha\text{-}2,8]$ -linked di- and trisialic motifs are found in the gangliosides GT1a and GT3 (see Fig. 1 in [75]), and $[\alpha\text{-}2,9]$ -linked polysialic acids (PSA) on glycoproteins [196]. Longer sialic acid chains are possible, and polysialic acids are found on neuronal cells and in a strikingly high concentration on tumor cells [197, 198].

Sialic acids play a critical role in the attachment processes of a number of viruses, such as reoviruses [199], coronaviruses [200], noroviruses [201], para- [202] and orthomyxoviruses (reviewed in [203]), adenoviruses [204, 205], and PyVs (reviewed in [192]), and the differing chemical groups attached to the molecule offer a large spectrum of attachment modes. This chemical diversity also helps to ensure specificity for sialic acid over other glycans (reviewed in [77]). While most viruses form contacts not only with sialic acid but also with other glycan moieties in an engaged carbohydrate, Neu5Ac-mediated interactions are often the most crucial, and receptors with more than one of these moieties can display significantly increased affinity and thereby facilitate further specificity, such as GT1a bound by MuPyV VP1 [75].

1.6.2 *Glycosaminoglycans*

Another group of carbohydrates that is exploited by number of viruses for cell attachment and entry are polyanionic GAGs. They are linear polysaccharides that are found on all animal cell surfaces as well as in the extracellular matrix. They are expressed in a tissue-dependent manner, and are generally classified into four groups: The hyaluronic acid type, the chondroitin/dermatan type, the heparan sulfate/heparin type, and the keratan type (reviewed in [206]). They are composed of repeating disaccharide units, consisting of an *N*-acetylated (or sulfated) hexosamine and a hexose or hexuronic acid and are attached to proteins via a serine residue, usually one xylose moiety, and two Gal moieties (see Fig. 11).

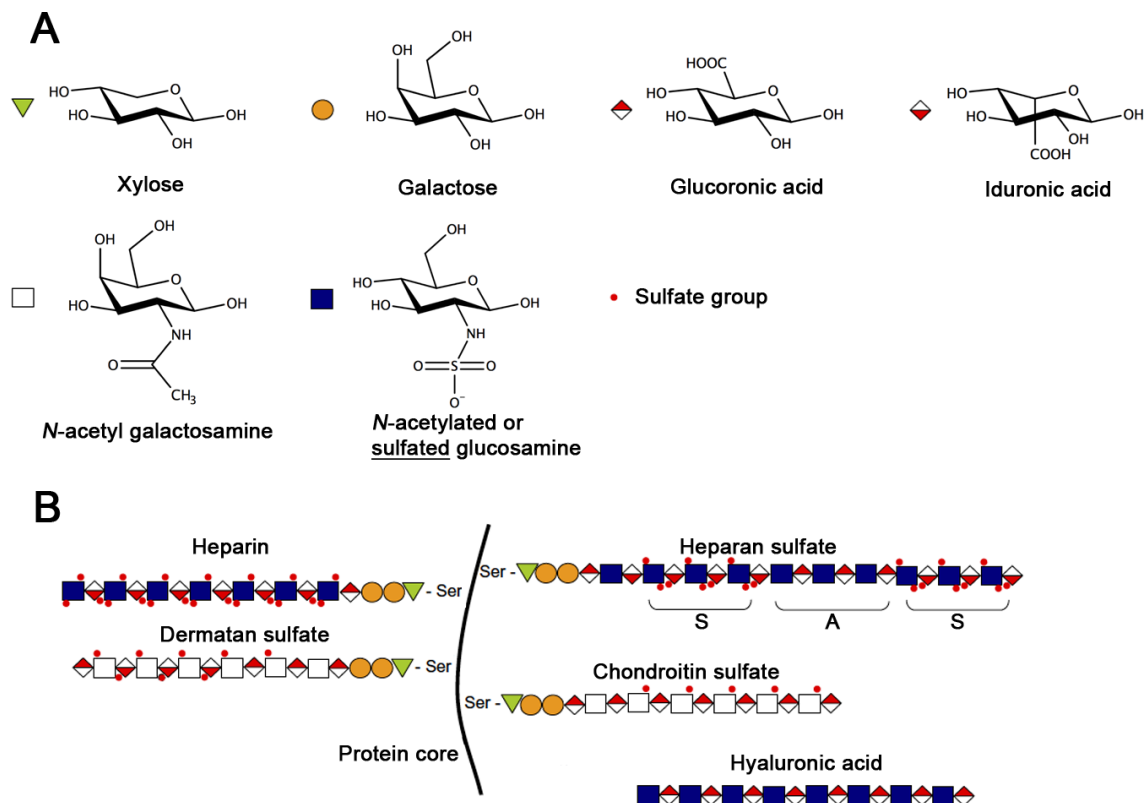


Fig. 11: The different GAG groups and their principal structure. A Building blocks of GAGs and their chemical structures are shown. **B** General GAG structure. GAGs are attached to proteins via serine residues, followed by one xylose and two Gal moieties. Their composition varies greatly, depending on the type of GAG, and so does the sulfation and acetylation state of the carbohydrates. Figure modified from [206], © 2006, Elsevier Ltd.

Depending on the group, GAGs display several different molecular properties. For example, hyaluronic acid is the only GAG that is not sulfated. Dermatan sulfate consists of alternating moieties of GalNAc and iduronic acid that is sulfated either at C4, C6, or both. C5 epimerization of iduronic acid to glucuronic acid leads to chondroitin sulfate (CS). Heparin and heparan sulfate (HS) are made up of either glucuronic or iduronic acid connected to glucosamine that is in turn either sulfated or *N*-acetylated. While heparin is homogeneously sulfated, HS is defined by short areas, termed S-domains, which show surprising variability in terms of sulfation patterns – glucosamine can, again, be *N*-acetylated or carry sulfate groups at C3 or

C6, and iduronic acid can be sulfated at C2. S-domains contain about 5-10 disaccharides and are separated by barely sulfated regions, called A-domains (Fig. 11).

GAGs in the context of proteoglycans fulfill a number of functions, such as water binding, which gives stability to cartilage structures (reviewed in [207]). Heparin is believed to chemically retain histamine molecules in the granules of mast cells [208] and is used as an anticoagulant [209]. Heparan sulfate has numerous functions such as cell adhesion; it also controls growth and proliferation and is a receptor for lipoprotein lipase and other proteins (reviewed in [208]). Keratan sulfate is mostly responsible for corneal transparency by regulating hydration levels (reviewed in [210]), while GAGs of the chondroitin/dermatan type play a role in wound repair [211] as well as brain development and the regulation of CNS regeneration [212]. Finally, hyaluronic acid is an important mediator of rheological properties in many tissues, for example in cartilage [213]. These are just some functions of GAGs, and their vast spectrum of uses as well as their ubiquitous presence makes them ideal targets for viruses. Indeed, a large number of viruses bind GAGs for cell attachment and entry, among them picornaviruses such as FMDV [214], several species of adenoviruses [215, 216], papillomaviruses [217], some polyomaviruses such as MCPyV [218], Ebola viruses [219], enteroviruses [220], coxsackieviruses [221, 222], and herpesviruses [137].

2. AIMS OF THIS DISSERTATION

This work consists of two major topics, the first of which focuses on the family *Polyomaviridae*. While the overall structure of the major capsid protein VP1, which mediates cell attachment, entry, and tropism, is conserved in the family, alterations in the connecting loops can have profound effects on the overall behavior of the virus, particularly at the cell attachment level. This, in turn, influences the rest of the viral life cycle, as was first shown for MuPyV. Various structural studies have been conducted on how polyomaviruses bind their carbohydrate receptors, but even though fundamental knowledge has been gained, many questions are left unanswered. In light of the progress in recent years in terms of protein purification, X-ray crystallography, and the availability of synthetic carbohydrates, this work aims to elucidate the molecular determinants of pathogenicity in the MuPyV strains, RA, PTA, and LID, where single amino acid exchanges in VP1 cause strikingly different behavior in tumorigenicity and spread. The specific questions to be answered were:

- Do the G91E and V296A exchanges influence the overall structures of RA, PTA, and LID VP1?
- Do the exchanges influence the binding mode for branched glycans, particularly GD1a and the recently identified receptor candidate GT1a?
- If not, what are the underlying mechanisms that result in the drastically different pathogenicity profiles?

Also, the effects of singular amino acid exchanges in PyV VP1 were to be examined in a more general context in SV40 VP1, where mutants that no longer bind the ganglioside GM1 but maintain their infectivity were analyzed. In this work, the task was to generate structural models of the mutants to explain abolishment of GM1 binding.

Finally, the recently established genus wukipolyomavirus is still relatively poorly investigated. The VP1 structures of two new human PyVs, HPyV6 and 7, were to be analyzed by X-ray crystallography in order to shed light on the receptor binding mechanisms of these new viruses. For this, expression constructs of VP1 were to be designed for protein purification, crystallization, and structure determination.

The topic of the second part of this dissertation is HSV-1. It infects cells by first attaching to GAGs on the surface with its glycoprotein C, followed by other viral proteins binding to their interaction partners and thereby mediating fusion with the cell membrane and subsequent entry. The crystal structures of these other glycoproteins have been solved in recent years, but gC has never been investigated. Nonetheless, it is an interesting protein to study, because it not only mediates the first step in infection, it also protects the virus by acting as a decoy receptor and blocking the C3b-dependent complement system. Furthermore, gC is bound by a monoclonal antibody, B1C1B4, which seems to attach to the GAG binding domain of the protein. The structure of this protein-antibody complex is not known. This work aims to lay the foundation for the structural analysis of gC by X-ray crystallography by establishing (1) a usable expression construct and (2) a reproducible purification strategy to obtain enough soluble protein for crystallization.

3. SMALL CHANGES, BIG DIFFERENCES – HOW MINOR ALTERATIONS IN THE VP1 BINDING POCKET CAN AFFECT THE ATTACHMENT AND INFECTION PROPERTIES OF POLYOMAVIRUSES

3.1 RESULTS

3.1.1 Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity

Michael H. C. Buch, A. Manuel Liaci, Samantha D. O'Hara, Robert L. Garcea, Ursula Neu, Thilo Stehle (2015) Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity. PLoS Pathogens 11: e1005104

The first step in viral infection is the attachment to a host cell. Many viruses use gangliosides or glycoproteins for this step, and the specificity and affinity of these interactions often determine tropism and pathogenicity. One established example of ganglioside-utilizing viruses is MuPyV, which infects mice and hamsters and causes tumors of varying density and origin. Glycans are engaged by a shallow groove on the major capsid protein VP1, and single amino acid exchanges in this binding groove have been demonstrated to have a profound and direct effect on the pathogenicity of the virus. The publication focuses on three MuPyV strains with remarkably different profiles: (1) a low tumorigenicity strain with limited spread, RA, that only causes few or singular tumors and only of mesenchymal origin and carries a glycine residue at position 91 of VP1 and a valine residue at position 296; (2) the naturally occurring strain PTA, which causes a high tumor density of mesenchymal and epithelial origin combined with a short latency. Like RA, it bears a valine residue at position 296, but a glutamic acid at position 91 replaces glycine; (3) the virulent LID strain causes early death in young mice by inducing brain hemorrhages and kidney failure. It also bears a glutamic acid at position 91, but V296 is replaced by alanine. These amino acid exchanges are sufficient to introduce the pathogenicity profile of one strain into another.

Earlier publications have established that oligosaccharides must feature at least one Neu5Ac moiety linked [α -2,3] to galactose in order to be bound by RA. Also, structural investigations elucidated the binding mode for a linear sialylated glycan as well as a branched glycan containing an [α -2,6] linkage together with RA VP1. In light of these findings it was hypothesized that a glutamate side chain at position 91 would partially occlude the binding pocket, which would prevent branched carbohydrates that act as pseudoreceptors from binding PTA and LID and would thereby facilitate a more effective spread within the host. Consequently, the gangliosides GD1a and GT1b were identified as infectious receptors for MuPyV, but no structural data on the binding mode had been available.

This publication presents the ganglioside, GT1a, a new infectious receptor for all three investigated strains of MuPyV. The research laboratory of Prof. Robert Garcea (University of Colorado Boulder) used mouse embryonic fibroblasts (MEFs),

which are blocked in a [β -1,4] synthase (GM2 synthase) and an [α -2,8] sialyltransferase (GD3 synthase) and therefore can only synthesize the ganglioside GM3. These ganglioside-deficient (gang^{-/-}) MEFs were supplemented with a specific ganglioside (either GD1a, GD1b, GM1, GT1a, or GT1b) that inserted itself into the cell membrane, and they were then infected with virus. Efficiency of infection was assessed by quantifying the expression of Large T-antigen in the MEF nucleus. These experiments confirmed the findings by Tsai *et al.* from 2003, who identified GD1a and GT1b as infectious receptors and excluded GM1 [71]. Also, it was shown that GD1b does not promote infection by MuPyV, either. Interestingly, GT1a was not only able to facilitate infection by all three investigated strains, but it did so at a significantly higher level than GD1a or GT1b. It was shown that virus binds to gang^{-/-} MEFs even in the absence of any ganglioside. Interestingly, the companion manuscript by You *et al.* [223] shows that virus is internalized even in the absence of gangliosides. This demonstrates that, while gangliosides are not essential for binding of virus to the cell as well as entry, they are essential for infection, with GT1a being more efficient than GD1a or GT1b.

To elucidate the mechanism of GT1a binding to MuPyV VP1, the structure of the complex was solved at high resolution. The longer arm of GT1a contains two sialic acid moieties connected by an [α -2,8] linkage (termed Neu5Ac_b and Neu5Ac_a in the publication), followed by [α -2,3]-linked galactose (Gal_a). While the Neu5Ac_a-Gal_a portion of GT1a binds identically to the structures solved by Stehle and colleagues in 1996 and 1997 [22, 224], the additional, [α -2,8]-linked Neu5Ac_b reaches out of the binding pocket, forms numerous additional hydrogen bonds with key amino acids of VP1 and causes GT1a to adopt a horseshoe-like conformation. The shorter arm of the ganglioside also contains a Neu5Ac_d-[α -2,3]-Gal_b motif, but for steric reasons it cannot be engaged to VP1. Strikingly, upon the comparison of the GT1a complexes of RA, PTA, and LID VP1, it was found that all three strains bind the glycan in an identical manner; neither the G91E exchange in PTA and LID nor the V296A exchange in LID cause any conformational changes, although the former establishes van-der-Waals contacts with Gal_a and the latter causes the loss of one van-der-Waals interaction with Neu5Ac_a.

Since the different pathogenicity profiles of RA, PTA, and LID could not be explained by variations in GT1a binding, it was reasoned that specificity for other gangliosides may play a role. Therefore, the interactions between VP1 and two other gangliosides that display distinctive Neu5Ac linkages were investigated: (1) GD1a, which is an established receptor for MuPyV and only carries [α -2,3]-linked sialic acid moieties, thereby resembling a truncated version of GT1a lacking the [α -2,8]-linked Neu5Ac_b; (2) DSLNT, which is the glycan portion of the ganglioside, 3'-6'-isoLD1 and carries an [α -2,6]-linked Neu5Ac_c. GD1a is engaged by all three strains in the same manner as was reported by Stehle and colleagues [22, 224]. Again, like GT1a, GD1a features two sialylated arms, but due to steric hindrances, only the longer arm is engaged by VP1. In contrast, some slight differences for DSLNT were observed. Like the other two investigated glycans, it features the Neu5Ac_a-[α -2,3]-Gal_a motif, which is engaged by all strains as was described before. The [α -2,6]-linked Neu5Ac_c, on the other hand, engages RA VP1 with few contacts in one binding pocket. Contrary to conclusions drawn earlier, PTA and LID are indeed able to bind DSLNT, but no interactions with Neu5Ac_c were formed. This moiety seems conformationally

flexible, which also results in moderate rearrangements and a sideward twist, both of which propagate throughout the sugar.

Since the differences in pathogenicity could not be explained by specificity for certain gangliosides, it was decided to investigate whether the strains display different affinities for their interaction partners. Unfortunately, PyV VP1 is known to have an affinity in the low millimolar range to oligosaccharides, and, coupled with the high costs and limited availability of ligands, this excluded classical means of measurement. Therefore, a crystallographic approach was used. VP1 of all three strains was crystallized in the exact same condition, and these highly isomorphous crystals were soaked with defined amounts of either GT1a, GD1a, or DSLNT. Data sets for every crystal were collected, the data were processed with isomorphous settings for the unit cell parameters as well as the same resolution (2.66 Å), subjected to simulated annealing, and the resulting $F_{\text{obs}} - F_{\text{calc}}$ density for the Neu5Ac_a-[α -2,3]-Gal_a moiety, which is present in all three glycans, was integrated. By plotting the electron density against the used carbohydrate concentration comparative binding levels for the strains could be obtained. As was found in the cell culture experiments described above, significantly better binding of GT1a to all three strains with very little difference between the strains was detected. GD1a bound weaker than GT1a, and PTA and LID binding was observed to be slightly better than in RA. Finally, binding of DSLNT was found to be the overall weakest at a level similar to RA VP1 binding to GD1a. These results were consistent with the findings obtained in cell culture and provide new insights into the molecular workings of MuPyV pathogenesis.

3.1.2 ***Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site***

Luisa J. Ströh, Ursula Neu, Bärbel S. Blaum, **Michael H. C. Buch**, Robert L. Garcea, Thilo Stehle (2014) Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site. *J Virol.* 88(18): 10,831-10,839

HPyV6 and HPyV7 have been identified on human skin and are shed together with the oncogenic MCPyV. Interestingly, while these two viruses share the same tropism with MCPyV, they are more closely related in sequence to WUPyV and KIPyV and have therefore been grouped into the genus wukipolyomavirus.

Because the details of the viral life cycle of wukipolyomaviruses remain largely unknown, the structures of HPyV6 and HPyV7 VP1, which commonly mediates cell attachment and entry, were solved by X-ray crystallography at high resolutions and refined with excellent statistics. Both structures reveal the typical jelly-roll motif found in all PyV VP1 so far, and while they both show the typically poor electron density for the CD-loop, it is interesting to note that the BC1-loop in HPyV7 VP1 has higher mobility than in HPyV6. The two VP1 structures display high similarity when superimposed onto one another – the pentamers superpose with a root mean square deviation (RMSD) of 0.7 Å. This reflects their high sequence

homology; only the EF-loop in HPyV6 is noticeably longer than in HPyV7. The two structures were also compared to VP1 of the other two wukipolyomaviruses, KIPyV and WUPyV [59]. As expected for VP1 structures in the same genus, the root mean square deviations were low (about 1.2 Å). When monomeric VP1 structures of HPyV6 and HPyV7 were compared to SV40 VP1 and MCPyV VP1 (which both belong to the orthopolyomavirus genus), the RMSD values increased slightly to about 1.4 – 1.7 Å. When pentamers were compared, the values decreased, which reflects the conserved jelly-roll motif present in all PyV VP1 structures elucidated so far.

While the core structure of VP1 is conserved, viral specificity is determined by the organization of the connecting loops on the surface of the protein. Typically, polyomaviruses engage sialylated glycans, and the binding site for Neu5Ac is mostly conserved among VP1 structures. In order to assess the binding capacities of HPyV6 and HPyV7 VP1, the BC1-, BC2-, DE-, and HI-loops were compared with the surface loops of the more closely related KIPyV and WuPyV as well as the more distantly related SV40 and MCPyV. The loop organizations of HPyV6 and HPyV7 are very similar – only the EF-loop in HPyV6 shows a slight extension. However, it is striking that the HI-loops in both viruses are significantly longer than in KIPyV or SV40. While this loop is usually just a small hairpin that closes off the glycan binding site in one direction, it folds on top of the VP1 pentamer in HPyV6 and HPyV7 and displaces the counter-clockwise (ccw) DE-loop. In contrast to the long HI-loop, the BC2-loop is significantly truncated compared to SV40 and KIPyV. This leads to a drastically altered binding surface, and it seems that the elongated HI-loops would sterically interfere with Neu5Ac binding.

To test this hypothesis, binding of HPyV6 and HPyV7 VP1 pentamers to two human cell lines was analyzed by flow cytometry. The cells were treated with neuraminidase beforehand, which abolished binding of the positive controls, JCPyV VP1 and MuPyV RA VP1; in contrast, neither the binding of HPyV6 and HPyV7 VP, nor of the negative control, JCPyV VP1 mutant L54F was in any way affected. These findings were confirmed *in vitro* by saturation transfer difference NMR spectroscopy. Again, no binding of HPyV6 or HPyV7 VP1 to either 3'-sialyllactose or 6'-sialyllactose could be detected, which is in line with results from crystal soaking and glycan array experiments.

In order to elucidate the cell attachment and antigenic properties of HPyV6 and HPyV7, the electrostatic surface potential of VP1 was analyzed and compared with those of other polyomaviruses. HPyV6 VP1 and HPyV7 VP1 are very similar to one another, which suggests similar strategies for cell attachment and perhaps also entry. Also, while they both possess concentric rings of positively and negatively charged patches around the five-fold axis of VP1, KIPyV VP1 is predominantly negatively charged and WUPyV shows a mostly positive surface potential. This demonstrates that there are differences even among members from the same genus. In line with this, it is interesting to note that HPyV6 and HPyV7 are more similar in terms of surface potential to MCPyV, considering that all three are shed from skin.

3.1.3 Mutations in the GM1 Binding Site of Simian Virus 40 Alter Receptor Usage and Cell Tropism

Thomas G. Magaldi, **Michael H. C. Buch**, Haruhiko Murata, Kimberly D. Erickson, Ursula Neu, Robert L. Garcea, Keith Peden, Thilo Stehle, Daniel DiMaio (2012) Mutations in the GM1 Binding Site of Simian Virus 40 Alter Receptor Usage and Cell Tropism. J Virol. 86(13): 7,028-7,042

This study utilizes SV40, one of the most intensely studied polyomaviruses, as a platform to investigate the effects of singular amino acid exchanges on the receptor binding properties of the virus and how switched receptor usage affects tropism. A library containing viable virus with mutations in VP1 was used in an infection screen with CV-1 cells to isolate and identify mutants that were resistant to neutralization with the classical SV40 receptor GM1. Viruses (mutant library virus and WT virus) were treated with GM1 prior to infection, and progeny virus was harvested 3 days p.i. After three rounds of infection and virus harvest, successful infection was analyzed by flow cytometry and immunostaining for large T-antigen, one of the early gene products of SV40. As expected, WT SV40 was successfully neutralized by incubation with GM1, while mutated virus was not inhibited. Harvesting and sequencing of viral genomes identified the mutants A70V, A70T, and H136Y. When equal numbers of these viruses were incubated with GM1 prior to infection, A70T and H136Y were somewhat neutralized, while A70V was actually stimulated. Introducing other hydrophobic side chains at position 70, i.e. leucine and isoleucine, could reproduce this stimulation effect. A larger hydrophobic side chain at position 70 would introduce unfavorable contacts with the galactose of GM1, which would significantly hamper binding of the ganglioside.

It was next analyzed whether the mutant viruses could use GM1 for infection. HeLa S3 cells, which naturally have a significantly lower amount of GM1 on their surface than CV-1 cells, were treated overnight with GM1, followed by infection with either WT SV40 virions or virions with hydrophobic mutations at position 70 of VP1. As expected, when compared to untreated cells, WT SV40 infection increased 4.5-fold, but the infection rates of A70V, A70L, and A70I only increased about 2- to 3-fold, suggesting that these mutants can still utilize GM1 for infection, but not as effectively as WT.

To quantify this deficiency, CV-1 cells were infected with the same number of virions of WT, A70V, A70L, A70I, A70T, and H136Y; after 24 hours p.i., the efficiency of infection was quantified by flow cytometry for large T-antigen. This showed that, while A70T and H136Y showed only a moderate decrease of infection, the hydrophobic mutants had a 3- to 4-fold reduction in infectivity compared to WT SV40. This suggests that the mutant viruses have a defect in their ability to use GM1 for the infection of cells.

Because the mutants identified so far can still utilize GM1, albeit at a lower level, a second screen with the background of the A70L VP1 mutation was carried out in 293TT cells to isolate mutants that were completely blind to GM1. These cells have low levels of GM1 on their surface and constitutively express SV40 large T-antigen, which enables high-level DNA replication and virus production. As described above, viral genomes were harvested after three rounds of screening and were

sequenced. All of the mutants carried the A70L substitution, and several mutants featured additional exchanges of amino acids that either directly or indirectly interact with GM1 moieties in WT VP1. Of the mutants identified in the screen, the double mutant A70L/N138Y and the triple mutant A70L/F75L/H129Q were investigated further.

By analyzing infection of HeLa S3 cells it was investigated whether these mutant viruses could be neutralized by GM1 incubation. A70L/N138Y and A70L/F75L/H129Q were not neutralized by GM1, and they also showed no stimulation effect like A70L. When the cells were supplemented with GM1, WT and A70L SV40 showed stimulation in infection, while the double and triple mutant did not. This showed that the new mutant viruses no longer used GM1 for infection.

Due to their resistance to GM1 neutralization and their decreased stimulation by GM1 addition to cells, it was hypothesized that mutants that carry the A70L substitution could no longer bind to GM1 on the cell surface. To validate this theory, recombinant VP1 pentamers were added to GM1-supplemented, trypsinated HeLa S3 cells, and their binding was analyzed by immunostaining and subsequent flow cytometry. While WT SV40 pentamers bound to the cells upon GM1 supplementation, the three mutant pentamers displayed no binding at all – entire capsids showed the same results. This demonstrates that the SV40 mutants are incapable of binding to GM1, which is in line with the biological findings.

The mutants identified so far were unable to utilize GM1, but were still perfectly capable of infecting cells. To analyze whether any of them displays altered tropism, infection of a wider array of cells was analyzed. The mutants infected CV-1 cells more poorly than WT virus, but the infection rate of the triple mutant was significantly increased compared to WT and the other mutants in HeLa S3 and 293TT cells, which display a lower amount of GM1 on their surface compared to CV-1 cells. The infection rate of the double mutant was restored to WT levels in these cells. This suggested that the tropism of the triple mutant was altered and allowed it to infect cells with low GM1 levels more efficiently than WT does. This seems more cell type- than species-specific, because HeLa S3 cells could be infected, while human foreskin fibroblasts could not. The double and triple mutant also showed elevated infection rates in LLC-MK2 cells from rhesus macaques. By generating A70L/F75L and A70L/H129Q double mutants and analyzing their neutralization and infection patterns it was established that A70L/F75L and A70L/H129Q infect HeLa S3 cells as well as WT virus does (and thereby affect the viral tropism), but only the triple mutant exceeded the WT infection levels and therefore seems to be best suited for interaction with a novel receptor and optimized infection.

It was unclear what molecule is used as a receptor by these viruses. Since most polyomaviruses attach to gangliosides to enter a cell, GM3 synthase, a critical enzyme for the synthesis of higher gangliosides, was repressed by RNA interference in HeLa S3 cells, which drastically reduced the abundance of gangliosides of the cell surface. Strikingly, not only WT SV40 but also the mutants showed virtually no infection, demonstrating that, while the mutants can infect cells in the absence of GM1, they require another, unknown ganglioside to do so.

Finally, it was shown before that an H136Y mutation in SV40 reduced the numbers of vacuoles in infected cells. This effect persists in the GM1-resistant mutants established in this study; however, infected cells still lysed and released at

least as much progeny virus as cells infected with WT SV40. Binding to GM1 seems to be directly linked with the virus' ability to form vacuoles in the host cell but not to establish infection.

3.2 DISCUSSION

Members of the *Polyomaviridae* family engage receptors in a binding pocket on the major capsid protein VP1. This work investigated the molecular determinants of how alterations in these pockets affect the viral life cycle.

One of the most prominent examples for such a case is MuPyV, where a relatively benign strain (RA) can adapt a high pathogenicity profile, when G91 is exchanged for E (PTA strain), and an additional V296A mutation renders the virus highly virulent (LID strain). It has been reasoned before that the G91E exchange may influence binding of sialylated carbohydrate receptors, but in 1997, structural information was only beginning to surface. With the increased commercial availability of carbohydrate compounds, a novel ganglioside with a new linkage, GT1a, has been identified, and its crystal structure in complex with MuPyV VP1 of all of the three different strains has been solved. This ligand engages the protein in a novel binding mode, with the additional, [α -2,8]-linked Neu5Ac reaching out of the binding pocket and establishing further hydrogen bonds. Even though no structures are available, it is conceivable that trisialylated compounds, such as GT3, can be accommodated in the MuPyV VP1 binding pocket as well. Furthermore, in contrast to earlier conclusions, the G91E exchange in PTA and LID neither introduces any conformational changes nor occludes the binding pocket for branched compounds. In fact, the increased contact area of E91 benefits the affinity for the classical receptor, GD1a, which may even be further amplified by avidity effects. These findings were confirmed by crystallographic soaking experiments, which have shown higher affinity for GD1a in PTA and LID VP1. These results as well as those presented in the companion manuscript [75, 223] have shown that, in general, attachment to a host cell does not necessarily equate entry, and entry does not equate infection. MuPyV particles bound to gang-/- MEFs, and this even lead to viral entry into the cells [223]; however, no expression of large T-antigen was detected in these cases, which indicates that no productive infection occurred, in line with findings that gangliosides are critical for infection. These findings suggest that promiscuous binding of MuPyV leaves it vulnerable to sorting onto a degradative entry pathway, when it attaches to an unfavorable receptor. Indeed, the carbohydrate motifs found in GD1a, GT1a, and GT1b or similar ones can also be found on *N*- and *O*-glycosylated proteins, but even though binding to these interaction partners can induce an immune reaction in animals [223], they do not confer infection, which only happens upon ganglioside binding. In light of these findings, an additional pathway has been added to Fig. 5 that sorts MuPyV towards a 'Dead End'. Also, it has been suggested in the past that integrins, particularly $\alpha_4\beta_1$, play a role in the attachment and entry process of MuPyV [225], and integrin recognition motifs do indeed exist in the VP1 structure. It was suggested that glycan binding would trigger a conformational change in VP1, which would facilitate integrin binding; however, these motifs are buried deep within the protein, and only significant unfolding would make it accessible for interaction with an integrin molecule, which seems unfavorable at this

early stage of infection, although it was shown that abolition of the interactions with this integrin reduces MuPyV infectivity. Taken together, these results clearly show that the straightforward thought model “Attachment, Entry, Infection” needs to be reviewed, as MuPyV binding of gangliosides appears to bear far more intricate consequences than previously thought. On a technical note, it should be mentioned that affinity measurements via X-ray crystallography can be very useful, when traditional methods such as surface plasmon resonance (SPR) cannot be used, and these techniques have already been utilized for adenoviruses [226], PyVs [227], and noroviruses [228]. They are, however, not without complications: numerous isomorphous crystals are needed (the unit cell parameters cannot deviate by more than 1%), which requires a considerable amount of beam time, even at synchrotron radiation sources; also, depending on the crystallization condition, high amounts of ligand are necessary, which can couple with high costs. The reproducibility of such soaking experiments can also be problematic. It was nicely demonstrated for MuPyV as well as for JCPyV [227], though, that even though results may not be directly physiological, they are consistent with findings obtained by *in vivo* experiments and also overlap with findings for isothermal calorimetry experiments conducted with SV40 and GM1 [23].

Another example of amino acid exchanges with far-reaching consequences was investigated in mutational studies with SV40 VP1. Here, a number of mutants were identified that are still viable and infectious but no longer use GM1 for infection. While it was shown that they still require gangliosides for infection, the nature of these new interaction partners is still elusive. The structural models of the mutants demonstrate that amino exchanges at positions 70, 75, 129, and 138 can have profound effects on glycan binding as they abolish the interaction with GM1 due to steric clashes.

These cases demonstrate that the surface of polyomavirus pentamers is very sensitive to changes. In MuPyV, the exchange of one amino acid can change the pathogenicity profile without directly excluding a specific receptor. In the case of SV40, these exchanges can abolish binding to one receptor, GM1, altogether, while the general receptor class, gangliosides, is still required; another interesting case is BKPyV, where the substitution of K63 to serine switches the specificity of the virus from GD3 *in vitro* to GM1 and thereby introduces a tropism change [74].

An extreme case of changes in PyV VP1 architecture is the genus wukipolyomavirus. Here, the topology of the binding pocket is changed to such a degree that ‘classical’ engagement of gangliosides is no longer required for infection. This has been shown for WUPyV and KIPyV [229] and also for the recently discovered HPyV6 and HPyV7. It is unknown what these viruses bind, and they have been implicated in respiratory infections, but their precise role in disease is still not completely clear [230, 231], although HPyV7 has been detected in thymic epithelial cancers [232].

Recombinant pentamers present an excellent platform to study cell attachment processes in polyomaviruses using structural biology methods, particularly X-ray crystallography. A portion of each of their N- and C-termini are missing, but the overall structure of the pentamers, especially the receptor-binding pocket, is not compromised. They are easily expressed and purified with high yields, and the resulting VP1 crystals usually diffract to resolutions in the range of at least

2–3 Å or better, which is more than sufficient for studies of protein-carbohydrate interactions. A large number of structures have been published in recent years, which has provided us with insights into the life cycles of PyVs. However, it should be noted that this platform has limits. The discovery that the MCPyV binding site for GAGs [218] does not coincide with the sialic acid binding site and no such conserved sites can be found on known PyV pentamers [233] requires investigations on a larger scale by using virus-like particles (VLPs). In this setting, binding that occurs in-between capsomers can be investigated, and avidity effects can better be taken into account than with single pentamers. However, as of now, the production of PyV VLPs is rather troublesome, results in poor yields, and crystallography with such large structures has cumbersome complications, such as a large unit cell.

Furthermore, it should be stated that, while investigations of attachment provide fundamental clues about viral pathogenicity, tropism, and the overall course of infection, this first step certainly is not the only factor that regulates the life cycle of PyVs. The studies conducted and presented in this work do not focus on VP1 interactions in the ER or on processes that occur upon virus egression. Many viruses make use of proteins that cleave off their original attachment receptors in order to prevent reinfection of an already infected cell, such as influenza viruses (neuraminidase [234]) and herpesviruses (heparanase [235]). PyVs do not use such enzymes, so it remains unclear how ganglioside binding of progeny virus influences the overall infection and spread efficiency. In light of this, it is conceivable that glycan affinity of VP1 is “fine-tuned” in such a way that it is high enough to mediate attachment to a host cell but at the same time low enough to facilitate efficient release of progeny virus from a parent cell.

Finally, it must be noted that even though our knowledge of protein-carbohydrate interaction has grown steadily over the last years, glycans are, in comparison to other macromolecules, overall still poorly understood. As was stated in a review by Ströh and Stehle [192], investigations of carbohydrates suffer from their heterogeneity, the high effort of synthesizing them, and the fact that they are not genetically encoded. Moreover, while we can identify interactions of PyVs with gangliosides, we cannot say with certainty on which cell types, under what circumstances, in what cell stage, and at what concentrations they are presented and distributed. The work presented in Buch *et al.* and You *et al.* increases our understanding of pathogenicity regulation in MuPyV, but a direct link between a higher tumorigenicity and a single amino acid exchange can, unfortunately, not be given. We must gain broader insight into the fundamental workings of glycobiology to obtain a better understanding of what exact roles these macromolecules play in viral infections.

4. DEVELOPMENT OF A PURIFICATION STRATEGY FOR HSV-1 GLYCOPROTEIN C

4.1 MATERIAL AND METHODS

4.1.1 Materials

4.1.1.1 Chemicals

β-Mercaptoethanol, 99%	Sigma (Munich)
Acrylamide-bisacrylamide solution	
Rotiphorese® Gel 30	Roth (Karlsruhe)
Ampicillin sodium salt (Amp)	Sigma (Munich)
Acetic acid, 100% p.a.	Roth (Karlsruhe)
Agarose NEEQ, Ultra Qualität	Roth (Karlsruhe)
Ammonium persulfate (APS)	Roth (Karlsruhe)
Bromophenol Blue, sodium salt	Serva (Heidelberg)
Coomassie Brilliant Blue R-250	Serva (Heidelberg)
Dimethyl sulfoxide (DMSO) Hybri-Max®	Sigma (Munich)
Ethylene diamino tetraacetate (EDTA) > 99%	Roth (Karlsruhe)
Ethanol, abs.	Sigma (Munich)
GelRed® 10,000x in water	Genaxxon Bioscience (Ulm)
Glycerol ROTIPURAN® ≥99.5%, p.a.	Roth (Karlsruhe)
Glycine PUFFERAN® ≥ 99%, p.a.	Roth (Karlsruhe)
Guanidine hydrochloride, ≥ 99%	Sigma (Munich)
HCl conc.	Sigma (Munich)
HEPES PUFFERAN® ≥ 99%, Buffer Grade	Roth (Karlsruhe)
Imidazole ≥ 99%	Sigma (Munich)
Methanol analytical reagent grade	Fisher Scientific (Leicester, UK)
Milk powder blotting grade	Roth (Karlsruhe)
NaOH	Roth (Karlsruhe)
Nickel sulfate hexahydrate, puriss., p.a., Reag. ACS	Riedel-de Haën (Seelze)
Ponceau S	Roth (Karlsruhe)
Powdered milk, blotting grade	Roth (Karlsruhe)
Rotiphorese® 10x Running Buffer	Roth (Karlsruhe)
Sodium dodecyl sulfate (SDS) ≥ 99%	Roth (Karlsruhe)
Sodium acetate	Sigma (Munich)
Sodium chloride ≥ 99.8%	Roth (Karlsruhe)
Sodium phosphate dibasic ReagentPlus® ≥ 99.0%	Sigma (Munich)
TEMED, puriss.	Roth (Karlsruhe)
Trizma base, ≥ 99%	Sigma (Munich)
Tween® 20	Sigma (Munich)

4.1.1.2 Reagents, Enzymes, and Kits

10x Buffer Tango	Fermentas (St. Leon-Rot)
10x Ligase Buffer	Fermentas (St. Leon-Rot)
BamHI	Fermentas (St. Leon-Rot)
Deoxyribonucleotides (100 mM each)	Promega (Mannheim)
ECL Plus Western Blotting Detection System	GE Healthcare (Uppsala, Sweden)

Endo H _f	New England Biolabs (Ipswich, USA)
GeneRuler® 1 kb Plus DNA ladder	Fermentas (St. Leon-Rot)
Goat Anti-Mouse IgG HRP Conjugate	Novagen (Billerica, USA)
His-Tag® Monoclonal Antibody	Novagen (Billerica, USA)
Instant Blue™	Expedion (Cambridge, UK)
Lipofectamine® 2000 Transfection reagent	Thermo Fisher Scientific (Waltham, USA)
Monoclonal Anti-Human IgG1	Sigma (Munich)
<i>NotI</i>	Fermentas (St. Leon-Rot)
PageRuler™ Prestained Protein Ladder	Fermentas (St. Leon-Rot)
PageRuler™ Unstained Protein Ladder	Fermentas (St. Leon-Rot)
<i>PvuI</i>	New England Biolabs (Ipswich, USA)
QIAquick® PCR purification Kit	Qiagen GmbH (Hilden)
ReproFast DNA polymerase	Genaxxon Biosciences (Ulm)
Shrimp alkaline phosphatase (SAP)	Fermentas (St. Leon-Rot)
T4 DNA ligase	Fermentas (St. Leon-Rot)
<i>Taq</i> DNA polymerase	Fermentas (St. Leon-Rot)
Trypan Blue Stain 0.4%	Gibco (Carlsbad, USA)
Wizard® Plus Miniprep SV Kit	Promega (Mannheim)
<i>XhoI</i>	Fermentas (St. Leon-Rot)

4.1.1.3 **Solutions and Buffers**

Bacteria Culture

Ampicillin (Amp) Stock	500 mg Amp dissolved in 10 mL H ₂ O, filtered and stored at -20°C
LB	25 g Luria Broth dissolved in 1 L H ₂ O
LB-Amp	LB with 50 µg/mL Amp
LB-Amp Agar plates	7.5 g Agar-Agar dissolved in 500 mL LB and, after autoclaving, supplemented with 50 µg/mL Amp and poured into Petri dishes
SOC medium	5 g/L yeast extract 20 g/L casein peptone 8.5 mM NaCl 2.5 mM KCl 10 mM MgCl ₂ 20 mM glucose after autoclaving

Molecular Biology

50x TAE	60.6 g Tris base 14.3 mL glacial acetic acid 25 mL 0.5 M EDTA pH 8.0 H ₂ O ad 250 mL
Agarose Gel	600 mg agarose in 40 mL 1x TAE buffer
6x DNA sample buffer	4 g sucrose 2 mL 0.25% (w/v) Bromophenol Blue 100 µL 10% (w/v) SDS stock 20 µL 0.5 M EDTA stock H ₂ O ad 10 mL
Sodium acetate stock	3 M sodium acetate pH 5.2, titrated with glacial acetic acid

Cell Culture

DPBS w/o Calcium, w/o Magnesium	Genaxxon Bioscience (Ulm)
FBS	Gibco (Carlsbad, USA)
Geneticin® Selective Antibiotic (G418 Sulfate)	Gibco (Carlsbad, USA)
L-Glutamine 200 mM (100x)	Gibco (Carlsbad, USA)
MEM Alpha Medium 1x	Gibco (Carlsbad, USA)
Penicillin/Streptomycin (100x)	PAA (Pasching, Austria)
Sodium pyruvate 100 mM solution	Sigma (Munich)
Trypsin/EDTA (1x Solution)	Genaxxon Bioscience (Ulm)
Ultra-Low IgG FBS	Gibco (Carlsbad, USA)
G418 stock	10 g G418 333 mL MEM Alpha Medium 1x
Non-selection medium (NSM)	500 mL MEM Alpha Medium 1x 6 mL Penicillin/Streptomycin 6 mL 100 mM sodium pyruvate 6 mL 200 mM L-glutamine 58 mL FBS Stored at 4°C
Minimal medium	Same as NSM, but omitting FBS Stored at 4°C

Selection medium (SEL)	492 mL MEM Alpha Medium 1x 6 mL Penicillin/Streptomycin 6 mL 100 mM sodium pyruvate 6 mL 200 mM L-glutamine 30 mL 30 mg/mL G418 stock (1.5 mg/mL final concentration, based on a kill curve for CHO Lec 3.2.8.1. cells, established in [236]) 60 mL FBS; stored at 4°C
Expression medium (NSM-LOW)	500 mL MEM Alpha Medium 1x 6 mL Penicillin/Streptomycin 6 mL 100 mM sodium pyruvate 6 mL 200 mM L-glutamine 58 mL Ultra-low IgG FBS Stored at 4°C
Protein Electrophoresis	
4x sample buffer	1.6 mL 10% (w/v) SDS 2.8 mL glycerol 0.5 mL 1.5 M Tris, pH 6.8 0.4 mL β-mercaptoethanol 2.5 mL H ₂ O Spatula tip Bromophenol Blue
Stacking gel buffer	1.5 M Tris pH 6.8 titrated with HCl at room temperature.
Separation gel buffer	1.5 M Tris pH 8.8 titrated with HCl at room temperature.
Buffers	
All buffers were made with MilliQ water cooled to 4°C.	
Size exclusion buffer	20 mM HEPES 150 mM NaCl Titrated to pH 7.5 at 4°C with NaOH, filtered, degassed, stored at 4°C
Milk solution	1 x TBS (see below) 5% (w/v) powdered milk Stored at 4°C
Ponceau stain	0.1% (w/v) Ponceau S 5% (v/v) acetic acid

Protein A binding buffer	173 mM glycine 500 mM NaCl Titrated to pH 9 at 4°C with NaOH, Filtered, degassed, stored at 4°C
Protein G binding buffer	20 mM Na ₂ HPO ₄ 150 mM NaCl 10 mM EDTA Titrated to pH 7 with HCl Filtered, degassed, stored at 4°C
Protein A/G elution buffer	100 mM glycine Titrated to pH 2.8 with HCl, Filtered, degassed, stored at 4°C
Protein A neutralization buffer	1 M Tris 1 M NaCl Titrated to pH 8 with HCl, Filtered, stored at 4°C
Protein G neutralization buffer	Same as Protein A neutralization buffer, but omitting 1 M NaCl
10 x TBS	200 mM Tris 1.5 M NaCl Titrated to pH 7.5, stored at RT
TBS-T	1 x TBS 0.5% (v/v) Tween-20
Western Blot transfer buffer	3.025 g Trizma Base 14.4 g glycine 200 mL methanol 800 mL H ₂ O; stored at 4°C

4.1.1.4 **Bacterial Strains and Vectors**

E. coli XL10 Gold is a propagated stock in the laboratory of Prof. Dr. Thilo Stehle. It was used as a cloning host and for plasmid propagation.

Table 2. Vectors used and generated in this study.

Vector	Source
gC II-5-His in pFastBac1	Dr. Beata Adamiak, laboratory of Prof. Dr. Tomas Bergström.
pcDNA 3.1(-)	Stock of Dr. Michaela Renner-Schneck, laboratory of Prof. Dr. Thilo Stehle.
gC II-5-His in pcDNA 3.1(-)	Generated in this study.
MuPyV PTA His ₆ -VP1* in pET15b	Generated in [237].
MuPyV RA His ₆ -VP1* in pET15b	
Nectin-4 V in pcDNA 3.1(-)	Generated in the laboratory of Prof. Dr. Thilo Stehle in an unrelated project.
gC II-5-Fc in pcDNA 3.1(-)	Generated in this study.

4.1.1.5 **Consumables**

Amersham Hyperfilm™ ECL	GE Healthcare (Uppsala, Sweden)
Amicon® Ultra Centrifugal Devices 50 kDa MWCO	Millipore (Schwabach)
CELLMASTER polystyrol (PS) Filter Roller Bottles, 2,125 cm ²	greiner bio-one (Frickenhausen)
CELLSTAR® 96 well cell culture plate	greiner bio-one (Frickenhausen)
CELLSTAR® filter top cell culture flasks 25 cm ² , 75 cm ² , 175 cm ²	greiner bio-one (Frickenhausen)
CELLSTAR® serological pipets 1 mL, 2 mL, 5 mL, 10 mL, 25 mL, 50 mL	greiner bio-one (Frickenhausen)
Centrifuge filters (0.22 µm)	Corning (Corning, USA)
coliRollers™ plating beads	Merck (Darmstadt)
Cryo.s™, PP, 1mL tubes	greiner bio-one (Frickenhausen)
Gel loading tips	greiner bio-one (Frickenhausen)
KimWipes	Kimberly-Clark (Roswell, USA)
Latex gloves, powder-free	Kimberly-Clark (Roswell, USA)
Membrane filters 0.22 µm, 0.45 µm, 5 µm	Millipore (Schwabach)
Microplate 96 well	greiner bio-one (Frickenhausen)
Nitrile gloves	Kimberly-Clark (Roswell, USA)
nProtein A Sepharose™ 4 Fast Flow	GE Healthcare (Uppsala, Sweden)
Parafilm “M”	Picheny (Chicago, USA)
PCR Micro Amp reaction tubes	Applied Biosystems (Foster City, USA)
Petri Dishes Cellstar®	greiner bio-one (Frickenhausen)
Protein A/G UltraLink® Resin	Thermo Scientific (Langenselbold)
Roti®-PVDF, 0.45 µm pore size	Roth (Karlsruhe)
Tissue culture dishes	greiner bio-one (Frickenhausen)
Tissue culture plate 6 well	Sarstedt (Numbrecht)
Pipet tips 10 µL, 200 µL, 1,250 µL	nerbe plus (Winsen/Luhe)

Plastic Cuvettes	Sarstedt (Numbrecht)
Reaction tubes, 1.5 mL, 2 mL	greiner bio-one (Frickenhausen)
Reaction tubes, 15 mL, 50 mL	greiner bio-one (Frickenhausen)
Syringe-top filters 0.2 µm	VWR International (Vienna, Austria)
Syringes, with Luer lock	BD (Franklin Lakes, USA)
Whatman Blotting Paper	VWR International (Vienna, Austria)

4.1.1.6 *Instruments*

Cell Culture

Bench top roller	Wheaton (Millville, USA)
HERAcell® 150 CO ₂ incubator	Kendro Laboratory Products (Langenselbold)
HERAcell® 240 CO ₂ incubator	Kendro Laboratory Products (Langenselbold)
HERAsafe KS12 safety cabinet	Kendro Laboratory Products (Langenselbold)
Leica DM IL microscope	Leica (Wetzlar)
Nalgene® Mr. Frosty freezing container	Sigma (Munich)
Neubauer cell counting chamber, 0.1 mm depth	Superior Marienfeld (Lauda Königshofen)

Protein Purification

ÄKTA Ettan LC	GE Healthcare (Uppsala, Sweden)
ÄKTA Prime Plus	GE Healthcare (Uppsala, Sweden)
BioLogic DuoFlow™	Bio-Rad (Munich)
HiTrap™ Protein A HP, 5 mL	GE Healthcare (Uppsala, Sweden)
HiTrap™ Protein G HP, 1 mL	GE Healthcare (Uppsala, Sweden)
HisTrap™ column, 1 mL, 5 mL	GE Healthcare (Uppsala, Sweden)
Hypercassette™	Amersham Biosciences (Amersham, UK)
Peristaltic Pump EconoPump	Bio-Rad (Munich)
Superdex 200 INCREASE 10/300 Tricorn	GE Healthcare (Uppsala, Sweden)
Superdex 200 PC 3.2/30	GE Healthcare (Uppsala, Sweden)
Hamilton® syringe, 50 µL volume	Sigma (Munich)

Miscellaneous

Analytical balances ME	Sartorius (Göttingen)
Autoclave VX95	Systec (Wettenberg)
basic 2 illuminator	Kaiser (Buchen)
Energia microwave	Panasonic (Hamburg)
Heidolph MR Hei-Mix	neoLab® (Heidelberg)
Heraeus Multifuge 1 L-R	Thermo Scientific (Langenselbold)
Ice machine AF-80	Scotsman (Milano, Italy)
iCycler	Bio-Rad (Munich)
Intelli-Mixer	neoLab® (Heidelberg)
Mini Protean 3 cell	Bio-Rad (Munich)

NanoDrop® ND-1000 spectrophotometer	Peqlab (Erlangen)
neoAccupette 3-9905	neoLab® (Heidelberg)
neoBlock 1 2-2503	neoLab® (Heidelberg)
pH electrode PB-11	Sartorius (Göttingen)
PowerPack HC™	Bio-Rad (Munich)
Pipetman pipets	Gilson (Middleton, USA)
Precision balance	Sartorius (Göttingen)
RCT Basic heating plate	IKA Werke (Staufen)
Shaking incubator Multitron 2	Infors (Bottmingen, Switzerland)
Shaker DOS 10-L	neoLab® (Heidelberg)
Table centrifuge 5415D	Eppendorf (Hamburg)
Thermomixer comfort, 1.5 mL	Eppendorf (Hamburg)
Trans-Blot® SD Semi-dry transfer cell	Bio-Rad (Munich)
Transilluminator Universal Hood II	Bio-Rad (Munich)
Transsonic T460	Elma (Singen)
Ultrapure water Biocell	Millipore (Schwabach)
Vacuum centrifuge SpeedVac Heto	Heto (Allerød, Denmark)
Vortex Genie 2™	Bender & Hobein (Zurich, Switzerland)
X-ray film processor SRX-101A	Konica Minolta (Tokyo, Japan)

4.1.1.7 **Software and Online Tools**

DNA Calculator	Sigma Genosys (The Woodlands, USA) www.sigma-genosys.com
Fermentas Double Digest	www.fermentas.com/en/tools/doubledigest
LALIGN	Pearson WR, based on [238] www.ch.embnet.org/software/LALIGN_form.html
NEBcutter, V2.0	New England Biolaboratorys®, Inc (Ipswich, USA) www.neb.com [239]
Peptide Properties Calculator	www.basic.northwestern.edu/biotools/ProteinCalc.html
PROTPARAM	www.expasy.org [240]
REVERSE COMPLEMENT DNA	www.bioinformatics.org/sms/rev_comp.html [241]

4.1.2 **Methods**

4.1.2.1 **Microbiological Methods**

Bacteria Culture

All bacteria were handled only in a sterile working environment. All work involving bacteria was performed close to a flame, and only sterile, autoclaved pipet tips were used. Serological pipets were passed through the flame before medium was pipetted. Before starting the work, the tabletop work space as well as gloves and pipets were sterilized with 70% (v/v) ethanol. All necks of bottles and flasks containing medium were sterilized by passing the necks and lids through the flame before and after work.

Overnight cultures were used for small-scale plasmid preparation (Miniprep, see below). 10 mL of LB-Amp in a Falcon tube (50 mL) were inoculated with a glycerol stock or 200 μ L transformed bacteria and incubated at 37°C overnight, shaking at 150 rpm (rotation radius 5 cm).

For disposal purposes, all media that had come into contact with bacteria were treated with several mL of sodium hypochlorite and incubated before being disposed. Any consumable materials that had come into contact with bacteria were autoclaved prior to disposal.

Preparation and handling of glycerol stocks

Glycerol stocks were set up by mixing 900 μ L of confluent bacteria from an overnight culture with 300 μ L 50% (v/v) glycerol. For the inoculation of overnight cultures, glycerol stocks were handled in accordance with standard procedures of the laboratory of Prof. Dr. Thilo Stehle.

Transformation of Bacteria

50 μ L of chemically competent cells (*E. coli* XL 10 Gold) were thawed on ice. Then, 1 μ L of plasmid DNA or 5 μ L of a ligation reaction were added to the bacteria and mixed by gentle stirring with the pipet tip. The bacteria were then incubated on ice for 5 minutes. To induce transformation, the cells were exposed to a heat shock of 42°C for 30 seconds and placed on ice again for 2 minutes. In order to generate a resistance to ampicillin, 500 μ L of SOC medium without antibiotics were added, and the cells were incubated at 37°C on a Thermomixer rotating at 300 rpm for one hour. After that, 200 μ L of grown bacteria were plated on LB-Agar plates containing 50 μ g/mL ampicillin. The plates were incubated at 37°C overnight.

4.1.2.2 Methods in Molecular Biology

Isolation of Plasmid DNA and Assessment of DNA Concentration

Small-scale plasmid DNA preparations were performed using the Wizard® Plus Miniprep SV Kit (Promega). The kit was used according to the manufacturer's recommendations with one exception: in the last step of the preparation procedure, plasmid DNA was eluted from the column with 50 μ L water instead of 100 μ L water.

DNA concentration was assessed using a NanoDrop® against a water blank by measuring the absorption at 260 nm, A_{260} . The DNA concentration was calculated as follows:

$$c_{\text{DNA}} \left[\frac{\text{ng}}{\mu\text{L}} \right] = A_{260} \cdot 50 \frac{\text{ng}}{\mu\text{L}} \quad (2)$$

Polymerase Chain Reaction

DNA amplification by polymerase chain reaction (PCR) [242] was used for three cloning purposes in this study: (1) introduction of a mutation by using two different sets of preparative PCR reactions (overlap extension PCR, Fig. 12), (2) amplification of the desired construct using preparative PCR, and (3) screening for successful insertion of constructs into a vector by analytical PCR.

Table 3. Primers used in this study. Overhangs required for restriction enzymes are underlined, restriction sites are in bold face. In VM_fw2 and VM_rv2, the base triplets introducing a valine-to-methionine mutation are shaded yellow. pcDNA3.1(-)_seq_fw and pcDNA3.1(-)_seq_rv are specific sequencing primers for pcDNA 3.1(-) taken from [236]. Since the sequencing method employed by MWG only reliably covers about 600 bp, gCII-5_opt-seq1 was designed and aligns to bp 592-620 of gC II-5-Fc.

Name	Sequence (5'-3')	T _m [°C]	Purpose
gC_fw_Xho2	<u>AT</u> CTC GAG ATG GCA CCT GGA CGT GTA GGC CTT GCT GTG GTC CTG	82	Preparative PCR of gC II-5-His.
gCHis_rv_Bam3	<u>GT</u> GGA TCC TTA ATG ATG ATG ATG ATG ATG CTC GAT TGC CTC GAT AAC CTG CCG CTC GGT	86	
VM_fw2	AGT ACG <u>TGG</u> CAC CGC GAC TCC GTG ATG TTC TCC CGA CG AAT	74	Overlap extension PCR for site-directed mutagenesis.
VM_rv2	CCG GTG <u>GCA</u> TTG CGT CCG GAG AAC ATC ACG GAG TC	68	
gCII-5_opt_Xho_fw1	<u>AT</u> CTC GAG ATG GCT CCA GGT CGT GTT GGA CTT G	72	Preparative PCR of gC II-5-Fc.
gCII-5_opt_Bam_rv1	<u>GT</u> GGA TCC TCA TTT GCC AGG GGA CAG GCT CAA	74	
pcDNA3.1(-)_seq_fw	TAA TAC GAC TCA CTA TAG GG	39	Sequencing of gC II-5-Fc.
gCII-5_opt-seq1	ACG AGT ATG GCA CGT GGG TCC GAG TGC GG	75	
pcDNA3.1(-)_seq_rv	TAG AAG GCA CAG TCG AGG	45	

Primer Design

Primers for construct amplification were designed to anneal with the template with at least 2/3 of their length. In order to prevent sliding, stretches of the same nucleotide were avoided. Also, equal amounts of A/T and G/C were preferred. The restriction sites necessary for cloning were added to the 5' ends of the primers, including specific overhangs for the respective restriction endonuclease (Tab. 3).

Primers for site-directed mutagenesis were designed by modifying the intended base triplet and adding bases in both the 5' and 3' direction to ensure specificity and to avoid self-annealing and the formation of secondary structures. The primers were checked for annealing to other sequences within the insert and the vector (LALIGN, www.expasy.org) and for the formation of any secondary structures or primer dimers (DNA CALCULATOR, www.sigma-genosys.com).

Primers were produced by biomers.net GmbH (Ulm), who also provided melting temperatures, T_M, of the primers by using the "Nearest Neighbor" approach (Tab. 3, [243]):

$$T_M = \frac{1000 \cdot d_H}{S + d_S + R \cdot \ln\left(\frac{C}{4}\right)} - 273.15 \cdot 16.6 \cdot \log[K^+] \quad (3)$$

d_H denotes the enthalpy sum of nucleotide pairs, S the entropy of helix formation (-10.8 cal), d_S the entropy sum of nucleotide pairs, R the gas constant

($1.984 \text{ cal} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$), C the oligonucleotide concentration (250 pM), and $[\text{K}^+]$ the potassium ion concentration (50 mM). The values for d_H and d_S were taken from [244] and [245], respectively.

From this, the annealing temperature, T_A , was calculated as follows:

$$T_A = T_M - 5^\circ\text{C} \quad (4)$$

Site-directed Mutagenesis

Overlap extension PCR was used for the removal of an unwanted valine-to-methionine mutation in the gC II-5-His construct provided by Dr. Beata Adamiak (Tab. 2). First, two separate sets of PCRs were carried out with the original construct as the template and the primer sets 1/3 and 2/4 (Reactions A1 and A2 in Fig. 12). The two generated PCR products are insert fragments that overlap at the intended mutation site. In the second PCR reaction, both fragments were used as templates with primers 1 and 4 (Reaction B), which yielded the final DNA fragment that no longer carried the mutation (Fig. 12).

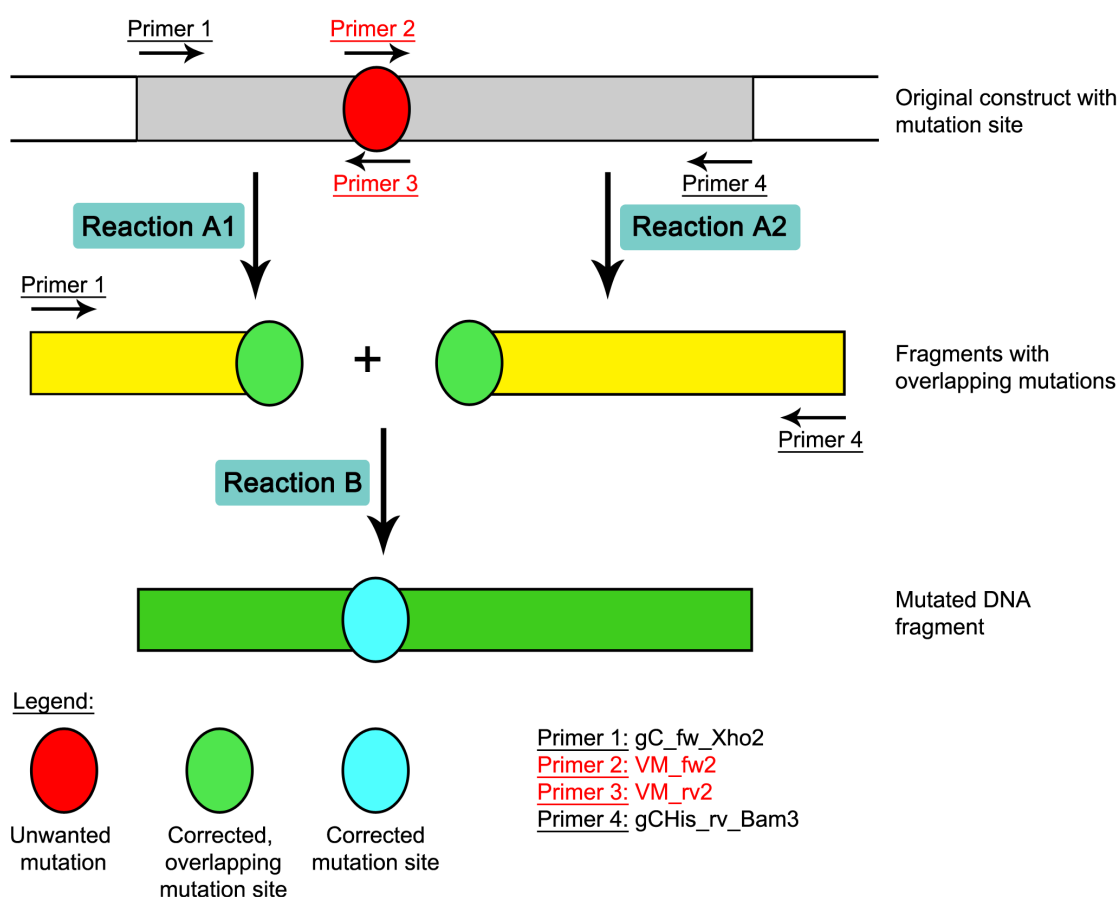


Fig. 12: General scheme of an overlap extension PCR. The first set of PCR reactions (Reactions A1 and A2) utilizes the original construct (gray) carrying the mutation site (red circle) as well as primer pairs 1/3 and 2/4, respectively. These reactions yield two fragments (yellow), which both possess the overlapping, corrected mutation site (green circles). In a second PCR reaction (Reaction B), both fragments are used as templates together with primers 1 and 4. This yields the final fragment (green), in which the unwanted mutation is corrected (blue circle). This fragment can then be treated with restriction endonucleases and ligated into the desired target vector.

The pipetting schemes for the reaction are identical to those used for regular PCR reactions. Products from Reactions A1 and A2 were purified using a QIAquick® PCR purification Kit (Qiagen). For Reaction B, 100 ng of each product were used as template. PCR products from both reactions were purified as described below. After purification, the mutated inserts were treated with restriction enzymes and then ligated into pcDNA 3.1(-) (see below).

Preparative PCR

In order to prevent mutations in PCR reactions used preparatively for cloning, a high-fidelity DNA polymerase, ReproFast (Genaxxon Biosciences), was used. PCR reactions were set up and carried out in accordance to standard procedures in the laboratory of Prof. Dr. Thilo Stehle.

Analytical PCR

Analytical PCR was used to confirm the presence of an insert inside an expression vector. Since the products of this PCR were not used for further cloning, no high-fidelity DNA polymerase was necessary. For budget reasons, the less precise *Taq* polymerase (Fermentas) was used. The reactions were carried out identically to preparative PCR reactions, except that only 30 amplification cycles were used instead of 35. Each PCR reaction contained 1 µL of template.

Agarose Gel Electrophoresis

To analyze PCR reactions and restriction digests, electrophoresis in 1.5% (w/v) agarose gels was performed. Gels were prepared and electrophoresis was performed in accordance to standard laboratory procedures; 2 µL GelRed® were used as a DNA-detecting dye.

DNA Purification

DNA fragments amplified with preparative PCR were purified using the QIAquick® PCR purification kit (Qiagen) according to the manufacturer's recommendations. The DNA was eluted in 50 µL H₂O.

Restriction digests of vectors and inserts were purified by ethanol precipitation according to standard procedures. After purification, the DNA concentration was assessed with a NanoDrop®.

Restriction Digests of DNA

Inserts and vectors were treated with the restriction enzymes *Xho*I and *Bam*HI (both Fermentas) to create single-stranded nucleotide overhangs before ligation. The double digests were carried out in twofold concentrated Buffer Tango (DOUBLE DIGEST, www.fermentas.com). To avoid star activity and to ensure complete digestion, 1 µg insert was digested with 0.5 µL *Xho*I and 0.75 µL *Bam*HI, while 2 µg pcDNA 3.1(-) were digested with 0.75 µL *Xho*I and 1 µL *Bam*HI. The reactions were incubated overnight at 37°C. Before heat inactivation for 20 minutes at 80°C, the vector was dephosphorylated with 3 µL SAP at 37°C for 30 minutes.

After a ligation reaction, the ligase was heat-inactivated, and the reaction mixtures were treated with *Not*I (Fermentas) in order to linearize uncut vector. *Not*I cleaves pcDNA 3.1(-) only between the restriction sites of *Xho*I and *Bam*HI and does

not digest the inserts (NEBcutter, www.neb.com). 1 μL *NotI* were added to a ligation reaction and incubated at 37°C for 30 minutes. The enzyme was then heat-inactivated at 80°C for 20 minutes.

Ligation of DNA

Ligation reactions were set up according to standard laboratory practice and were carried out using a stepwise gradient starting at 25°C with the temperature being reduced by 1°C every 60 minutes until it reached 10°C. After 60 minutes at 10°C, the ligase was heat-inactivated at 65°C for 20 minutes. Then, any uncut vector was linearized with *NotI* (see above). The ligation products were then used for the transformation of bacteria.

Sequencing of DNA

15 μL samples (50-100 ng/ μL DNA together with a sequencing primer [see Tab. 3]) in H_2O were sent to MWG (Ebersberg) for sequencing. The resulting sequences were compared to the expected ones using LALIGN (www.expasy.org) and, if necessary, by visual inspection of the ferograms. The resulting protein sequences were also compared to the expected ones using LALIGN.

4.1.2.3 Methods in Cell Culture

All work involving Chinese hamster ovary (CHO) Lec 3.2.8.1. cells was performed using a clean bench under sterile conditions. Only autoclaved pipet tips and sterile serological pipets were used. All liquid containers as well as gloves were disinfected with 70% (v/v) ethanol before being opened under the bench. For incubation, all cells were stored at 37°C and 5% CO_2 in a water-vapor-saturated atmosphere. Protocols and the general work flow of protein expression in eukaryotic cells in this study are based on the dissertation of Steffen Müller [236].

Thawing of Frozen Cells

Frozen cells were taken from the liquid nitrogen tank and thawed in a 37°C water bath. Once thawed, they were transferred to the clean bench and pipetted into a 50 mL Falcon tube. Then, 10 mL NSM were added slowly to the cells, followed by another 10 mL at normal speed. The cells were centrifuged for 8 minutes at 1,200 rpm in a Heraeus centrifuge, after which the supernatant was discarded. The cell pellet was resuspended in 15 mL NSM and seeded into a 75 cm^2 tissue culture flask, which was incubated at 37°C.

Cultivation and Passaging of Cells

Cells were monitored every other day under a light microscope. Once they reached about 90% confluency, the medium was aspirated, and the cell layer washed twice with 10 mL DPBS. Then, the cells were covered with a thin layer of trypsin/EDTA solution and incubated for about 5 minutes. They were detached by softly beating the flask, mixed with medium (which inhibited further trypsin activity), transferred to a 50 mL Falcon tube, and centrifuged for 8 minutes at 1,200 rpm in a Heraeus centrifuge. Afterwards, the trypsin-containing supernatant was aspirated, and the cells were resuspended in fresh medium and seeded into new flasks at a 3-10 fold dilution. The fresh flasks were incubated at 37°C.

Cryoconservation of Cells

For setting up cryostocks, cells at about 90% confluency were trypsinated and centrifuged as described above. However, the cell pellet was not resuspended in normal medium, but in 5 mL NSM containing 10% (v/v) DMSO. Five cryostocks were set up and transferred into a Mr. Frosty freezing container, which was then incubated at -80°C for 48-72 hours. This allowed for a slow cooling rate of about 1°C per hour. Finally, the frozen cells were stored in liquid nitrogen.

Transfection

Lipofectamine® 2000 was used to transfer the pcDNA 3.1(-) vector carrying its respective gC II-5 insert as well as a resistance marker into CHO Lec 3.2.8.1. cells. This reagent packages plasmid DNA into liposomes, whose surface is positively charged, allowing them to cross the negatively charged cell membrane and deliver their cargo to the nucleus [246].

On the day before transfection, cells that had grown to confluency were trypsinated and seeded into 6 well tissue culture dishes. Also, 5 µg of the recombinant vector were digested over night at 37°C with *PvuI* (New England Biolabs) in order to linearize the DNA. On the day of transfection, 2 mL NSM were replaced in the wells with confluent cells. Two solutions were set up: (1) 5 µg DNA mixed with 250 µL minimal medium and (2) 10 µL Lipofectamine® 2000 and 240 µL Minimal medium. After incubation for 5 minutes at RT, the two solutions were combined, gently mixed, incubated for another 20 minutes at RT, and then added to the cells, which were incubated at 37°C for 24 hours. On the next day, the cells were trypsinated and transferred to a 75 cm² flask with SEL. The G418 selection marker in the SEL medium, which was replaced every two days, killed the cells that had not taken up the plasmid; after about 10 days, only successfully transfected cells had survived in the flask and were grown to a confluency of about 70%, at which point they were transferred to a 175 cm² flask and subsequently frozen.

Subcloning

Successfully transfected cell colonies are called polyclonal cells, meaning that, while they carry the plasmid containing the gene of interest, individual cells in the colony have varying expression levels concerning the desired protein. To ensure consistent levels of protein expression, individual cells had to be isolated to give rise to new, monoclonal cells.

Polyclonal cells that were grown to a confluency of about 80-90% were trypsinated. Upon resuspension in SEL, the cells were counted, and the resuspension was diluted to 5 cells/mL. Of this dilution, 200 µL were pipetted into each well of a 96 well cell culture tray. After a week, the trays were visually inspected, and colonies that arose from a single cell were trypsinated and sequentially transferred in to 6 well tissue culture dishes. When the cells had reached 90% confluency, the medium was exchanged, and after two more days, samples from the supernatant were taken and tested for protein expression. The clone with the highest protein expression level was then chosen for large-scale protein expression.

Cell Counting

For the generation of monoclonal cells, it was necessary to dilute and isolate single cells and grow colonies from them. This required the assessment of the cell concentration. Cells were counted in a Neubauer counting chamber.

Trypsinated and resuspended cells were mixed 1:1 with 0.4% Trypan Blue Stain solution, which stains dead cells dark blue. 10 μ L of the mix was pipetted onto the counting chamber, and the living cells in the outer 4x4 squares of the chamber were counted, averaged, and the cell concentration was calculated as follows:

$$c = \frac{N_{av} \cdot F_d}{A \cdot d} \cdot 10^4 \quad (5)$$

c denotes the cell concentration (cells/mL), N_{av} the averaged number of counted cells, F_d the dilution factor of Trypan Blue (equals 2), A the area of a large square (1 mm^2), and d the depth of the chamber (0.1 mm).

Roller Bottle Setup and Harvest of Supernatant

For large-scale production of gC II-5-Fc, cells were grown and cultivated in roller bottles. This setup allows for a large cultivation area and a high volume of medium the protein can be secreted into.

The cell content of 4 confluent grown 175 cm^2 flasks was trypsinated and centrifuged as described above. The pellets were resuspended in 10 mL NSM-LOW each and transferred into a roller bottle containing 210 mL NSM-LOW. After 7 days, the protein-containing medium was collected and replaced with 250 mL fresh NSM-LOW. After four weeks, roller bottles were discarded and replaced with new roller bottles containing fresh cells. Roller bottles were incubated at 37°C and 5% CO_2 in a water-saturated atmosphere.

4.1.2.4 **Methods in Protein Analysis**

Protein A/G Pull-down

Protein A and Protein G are cell wall proteins isolated from *Staphylococcus aureus* and type G streptococci, respectively, which exhibit nanomolar affinity to the Fc part of IgG1-type antibodies [247]. This feature is exploited in the purification of Fc-tagged proteins and antibodies via affinity chromatography as well as immunoprecipitation techniques that were employed in pull-down assays to capture any gC II-5-Fc from singular CHO Lec 3.2.8.1. cell clones.

Each singular CHO Lec 3.2.8.1. clone was grown in a separate well in 6-well trays. When a confluency of 100% was reached, 2 mL NSM-LOW were exchanged, and two days later, 750 μ L of this supernatant were taken from the wells. The supernatant was filtered through a centrifuge tube filter (0.22 μm pore size). 10 μ L Protein A/G UltraLink[®] resin were washed three times by mixing with 1 mL Protein A binding buffer and centrifugation at 3,000 g for 3 minutes. The filtered supernatant was incubated with the washed resin in an orbital shaker at 4°C for 2 hours. Afterwards, the mixture was centrifuged again as described above, and the resin was washed three times as before. 20 μ L 4x SDS sample buffer was added to every

sample, they were incubated at 95°C for 5 minutes and subjected to SDS-PAGE. The presence of Fc-tagged gC II-5 was further confirmed by Western Blot.

Discontinuous SDS-PAGE

Proteins were separated according to their molecular weight by discontinuous SDS-PAGE [248]. Gels were cast and electrophoresis was carried out according to standard laboratory procedures. Gels were stained with Instant Blue™ solution by microwaving them for 30 seconds after removal of excess SDS with water.

Western Blot

Western Blotting [249] was used to confirm the presence of the Fc-tag or of gC II-5 in purified protein solutions. First, samples were subjected to SDS-PAGE using a prestained protein marker. The gel was afterwards equilibrated in Western Blot transfer buffer for 15 minutes, while a slightly larger PVDF membrane was equilibrated for 10 minutes in methanol, then for 5 minutes in water, and finally for another 10 minutes in transfer buffer. Whatman Papers were briefly moistened with transfer buffer as well, and the blot was assembled as follows from top to bottom:

- 4 Whatman Papers
- Equilibrated gel from SDS-PAGE
- Equilibrated PVDF membrane
- 4 Whatman Papers

The blot was run at 20 V for one hour at RT. Afterwards, the membrane was incubated in Ponceau staining solution for about one minute, and the marker bands as well as the lanes were marked with a pen. The stain was washed off with water. To avoid unspecific interactions, the membrane was incubated in milk solution at 4°C over night. On the next day, it was washed with TBS-T and TBS for 10 minutes each at RT. Afterwards, the membrane was incubated with the primary antibody used for detection (either an α -His-tag antibody (Merck) at a 1:10,000 dilution, an α -Fc antibody (Sigma) at 1:7,500 dilution, or the recombinantly produced B1C1B4 α -gC antibody at a dilution of 1:2,000) in milk solution over night at 4°C. On the final day, the membrane was again washed with TBS-T (1 minute, 20 minutes, 20 minutes) and then with TBS for 5 minutes. The secondary antibody (goat-anti-mouse 1:5,000, Novagen) was applied to the membrane for roughly 1.5 hours at RT, before the blot was washed 5 times with TBS-T for altogether 40 minutes. During the last wash, the ECL detector solution was prepared by mixing 1 mL of Solution A from the ECL kit with 25 μ L of Solution B. In a dark development chamber, the ECL detector solution was pipetted over the membrane, and medicinal X-ray films were exposed to the membrane for 30 seconds, 1 minute, 2 minutes, and 5 minutes. The films were developed in an X-ray processor.

NanoDrop™

For protein concentration assessment, UV measurements using a NanoDrop™ were carried out. For this, the instrument was initialized with water and blanked with 2 μ L of the appropriate buffer. Then, 2 μ L of the protein solution were pipetted on the

wiped pedestals. The expected molecular weight and extinction coefficient were provided, and the protein concentration was assessed.

4.1.2.5 **Protein Purification**

All steps during protein purification were carried out at 4°C or in a cooling cabinet, unless stated otherwise. Protein samples were kept on ice and stored at -20°C. All columns were stored in water. The purification strategy for gC II-5 can be found in section 4.2.4. For the composition of the used buffers refer to section 4.1.1.3.

Preparation of Cleared Cell Culture Supernatant

gC II-5-Fc-containing supernatant from transfected, monoclonal CHO Lec 3.2.8.1. cells was thawed at 4°C. Dead cells and other debris were removed by centrifugation at 3,500 rpm for 5 minutes in a Heraeus centrifuge. This cleared supernatant was passed sequentially through 5 µm, 0.45 µm, and 0.22 µm-sized membrane filters. Finally, the pH was adjusted to 9 with 5 M NaOH.

Protein A Affinity Chromatography

The high affinity of Protein A to IgG1-type antibodies was exploited in the purification of gC II-5-Fc, which is equipped with the Fc portion of human IgG1 at its C-terminus. Protein A chromatography was used for two purposes in the purification of gC II-5: (1) The isolation of gC II-5-Fc from cleared cell culture supernatant and (2) the separation of gC II-5 from the Fc-tag after TEV protease cleavage.

Isolation of gC II-5-Fc from Cleared Cell Culture Supernatant

The cleared cell culture supernatant was loaded onto a HiTrap™ Protein A HP column over night at a loading speed of 3.5 to 4 mL/min. This high loading speed prevented column clogging by unspecific binding of BSA. The column was then subsequently washed with 20 mL Protein A binding buffer and 20 mL Protein A binding buffer supplemented with 20% Protein A elution buffer. For elution of gC II-5-Fc, 25 fraction collector tubes were supplemented with 150 µL neutralization buffer each, and the protein was eluted with Protein A elution buffer. 1 mL fractions were collected. Peak fractions were analyzed by SDS-PAGE, pooled and frozen at -20°C after their protein concentration was assessed using a NanoDrop®.

To maximize the yield of gC II-5-Fc, the cell culture supernatant flowthrough was once more filtered through a 0.22 µm membrane filter and again loaded over night onto the HiTrap™ Protein A HP column as described above. After affinity chromatography and testing that no more gC II-5-Fc was present in the flowthrough, it was discarded.

Isolation of gC II-5 after TEV Protease Cleavage

After proteolytic cleavage on-column, gC II-5 and the Fc-tag needed to be separated from one another. After successful cleavage, gC II-5 no longer carried its purification tag, could no longer bind to the column, and was collected by running 15 mL Protein A binding buffer over the column. It was further processed by concentration and preparative size exclusion chromatography. The Fc-tag was eluted with about 20 mL

Protein A elution buffer, and the column was either washed with water and stored, or regenerated.

Column Regeneration

After two complete purification runs, the HiTrap™ Protein A HP column resin was regenerated to remove any non-specifically bound proteins, particularly BSA, which may clog and damage the column. In accordance with the manufacturer's instructions, the column was sequentially washed with 5 column volumes water, 5 column volumes 6 M guanidine hydrochloride, and 5 column volumes water.

Protein G Affinity Chromatography

The murine α -gC monoclonal antibody B1C1B4 was purified by exploiting the high affinity of its Fc-part to Protein G. B1C1B4-containing mouse hybridoma cell culture supernatant was provided by Dr. Beata Adamiak; it was centrifuged for 5 minutes at 1,500 g to clear away cell debris. The supernatant was then diluted 1:1 with Protein G binding buffer and loaded continuously over night onto a HiTrap™ Protein G HP column that had been equilibrated with binding buffer. The next day, the column was washed with binding buffer, and the antibody was eluted by running a gradient of 0-100% Protein G elution buffer. 1 mL fractions were collected, and each fraction was supplemented with 150 μ L neutralization buffer to counteract the low pH of the elution buffer. Peak fractions were analyzed by SDS-PAGE, pooled, and subjected to analytical SEC. Purified antibody was stored at -80°C.

Proteolytic cleavage

To remove the C-terminal tag from gC II-5-Fc, the TEV protease cleavage site was utilized. Digests were carried out in solution as well as on-column. For cleavage in solution, about 1 mg of gC II-5-Fc was incubated with 100 μ g TEV protease at 4°C over night. For cleavage on-column, gC II-5-Fc was loaded onto a HiTrap™ Protein A HP column, followed by 100 μ g TEV protease per mg gC II-5-Fc in a final volume of 2 mL with Protein A binding buffer. The column was incubated at 4°C for 48 hours, after which cleaved gC II-5 was eluted with binding buffer. TEV protease was expressed recombinantly and purified in the laboratory of Prof. Dr. Thilo Stehle.

Protein concentration

gC II-5 was concentrated after TEV cleavage for preparative size exclusion chromatography. For this, Amicon® Ultra Centrifugal Devices with a molecular weight cutoff of 50 kDa were used. They were first equilibrated with about 3 mL Protein A binding buffer and centrifuged for 2 minutes at 3,500 rpm in a Heraeus centrifuge. Then, the protein sample was loaded and centrifuged in 5-minute steps at the same speed. After each centrifugation step, the protein solution was resuspended with a pipet to prevent protein from getting stuck on the filter membrane. In order to minimize protein loss, pipet tips were not exchanged during centrifugation. This was repeated until the protein solution had the desired volume, at which point it was subjected to preparative size exclusion chromatography.

Size exclusion chromatography

Size exclusion chromatography (SEC) is employed to separate proteins in accordance to their Stokes radius. In this study, two types of this method, also commonly referred to as gel filtration, were used: (1) analytical size exclusion chromatography with a small protein sample on columns mounted on an ÄKTA Ettan system and (2) preparative size exclusion chromatography with a larger protein sample on a column mounted on a Bio-Rad BioLogic DuoFlow™ system. In both cases, size exclusion buffer was used, and the columns were first equilibrated with this buffer using standard wash programs. The injection loops were washed with water and size exclusion buffer, and every sample that was injected into the chromatography systems was first filtered through a 0.22 µm centrifugation filter.

Analytical SEC

This method of size exclusion chromatography was used to analyze the size, homogeneity, and purity of a protein sample. It utilizes a Superdex 200 PC column mounted on an ÄKTA Ettan LC chromatography system. A 23 µL protein sample was injected into the system and run over the column. The elution volumes of the peaks were then assigned to their protein size using a standard equilibration curve established in the laboratory of Prof. Dr. Thilo Stehle. 50 µL fractions were collected in a 96 well collector plate.

Preparative SEC

After removal of the Fc-tag and any TEV protease, gC II-5 was subjected to preparative size exclusion chromatography to separate aggregated protein from monomeric gC II-5. The protein sample was concentrated to a final loading volume of 500 µL; the fractions (0.5 mL each) that corresponded to the peak were analyzed by SDS-PAGE. Fractions containing pure gC II-5 were pooled, their concentration was assessed, and they were stored at -20°C.

4.1 RESULTS

4.1.1 Expression Constructs of gC II-5

The aggregation-prone mucin-like region, which also tends to increase dissociation rates from HSV-1 to GAGs [165], as well as the highly hydrophobic transmembrane helix near its C-terminus render full-length gC 1 unsuitable for purification and structure determination by X-ray crystallography. Therefore, protein variants lacking these problematic portions and with varying glycosylation sites have been designed in the past ([174], Dr. Beata Adamiak, personal communication). Based on this, two constructs have been prepared for this study (Fig. 13); they, too, do not contain either the mucin-like region (amino acids 33-123) or the C-terminal transmembrane helix (amino acids 478-511) but retain the N-terminal signaling sequence (amino acids 1-32). Also, they both contain all four *N*-glycosylation sites of gC. As such, both constructs span amino acids 1-32 and 124-477 and contain either a non-cleavable, C-terminal His-tag (gC II-5-His) or a cleavable, C-terminal Fc-tag (gC II-5-Fc). For cleavage of the Fc-tag, gC II-5-Fc was equipped with a linker that contains a cleavage site for TEV protease between gC II-5 and the tag. TEV protease was chosen because it cleaves very specifically at the sequence ENLYFQ/S, where / denotes the cleavage point [250]. Also, by virtue of an expression construct provided by the research group of Prof. Dr. Andreas Peschel, the protease can be easily produced in the laboratory. gC II-5-Fc with its Fc-tag removed is termed gC II-5 in this work.

gC II-5-His contains four glycosylation sites, while gC II-5-Fc contains five, one of which is located on the Fc-tag (see also Tab. 4). Such posttranslational modifications render the proteins unsuitable for expression in standard prokaryotic systems such as *E. coli*. Instead, a mammalian expression system was chosen. CHO cells carrying the lectin phenotypes 3, 2, 8, and 1 are adherent cells that were first isolated in 1989 [251], and they have been particularly useful in producing glycoproteins for crystallographic studies [252, 253]. Their mutations render them deficient in a number of transport proteins that are required in the synthesis of complex, sialylated glycans. As such, the glycoproteins produced by these cells only carry two GlcNAc and five to nine mannose residues, which can be cleaved off with glycosidases such as endoglycosidase H (EndoH_f).

All constructs were cloned into pcDNA 3.1(-), a mammalian expression vector that has been successfully used in the laboratory of Prof. Dr. Thilo Stehle before [236]. This plasmid has numerous advantages, such as a high expression level in eukaryotic cells and a high copy number in *E. coli*. It codes for two resistance genes; the neomycin resistance allows for selection with G418 in eukaryotic cells, and the ampicillin resistance allows for selection with Amp in *E. coli*.

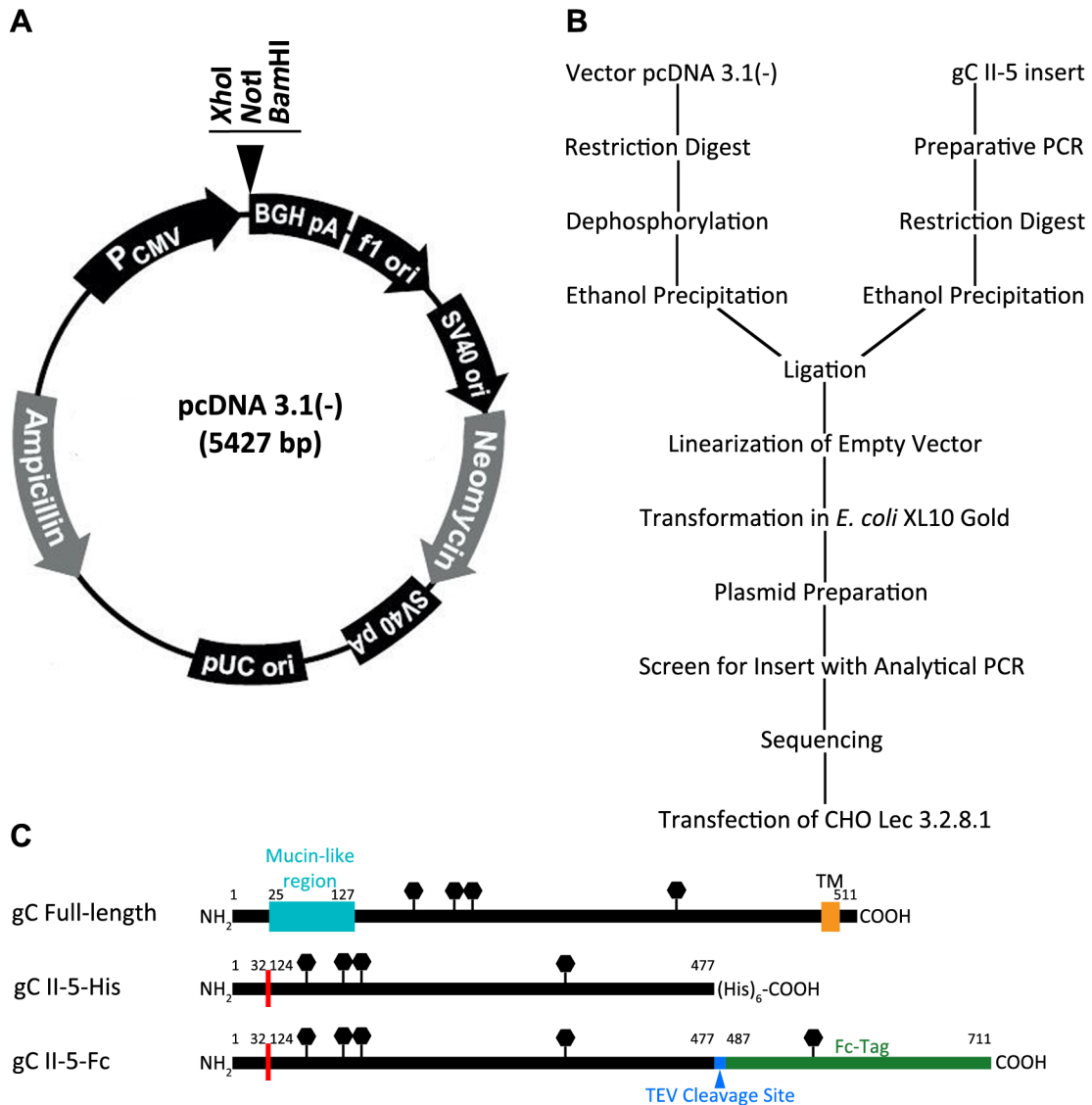


Fig. 13: Cloning of gC expression constructs. **A** Simplified vector map of pcDNA 3.1(-). The multiple cloning site (MCS) is indicated by a black triangle, along with the used restriction endonucleases. Panel adapted from Invitrogen (http://tools.lifetechnologies.com/content/sfs/manuals/pcdna3_1_man.pdf). **B** Cloning strategy for both gC constructs. Figure adapted from U. Neu [254]. **C** Schematics of the constructs relevant in this study. In both gC II-5-His and gC II-5-Fc, the mucin-like region (teal bar) and the C-terminal transmembrane helix (orange bar) have been removed. The excision of the mucin-like region is represented by a red line. N-linked glycosylation sites are symbolized by black hexagons. In gC II-5-Fc the linker and the TEV cleavage site are shown in blue, and the C-terminal Fc-tag is shown in green.

Table 4: Properties of the expression constructs. The C-terminal His-tag of gC II-5-Fc is not cleavable. The TEV cleavage site in gC II-5-Fc is shown in bold. For the full sequence of the Fc-tag (Fc, second row), see 5.3.

Construct	Non-native amino acids at the C-terminus	Length [aa] ¹	MW [kDa] ^{1,2}	pI ¹
gC II-5-His	HHHHHH	392	43.22	6.17
gC II-5-Fc	GTENLYF QG -Fc	620	68.73	6.15
gC II-5	GTENLYFQ	394	43.35	5.77

¹Calculations include the N-terminal signaling sequence.

²The molecular weight does not include any glycosylation.

4.2.1 *Cloning of gC II-5-His into pcDNA 3.1(-)*

A construct coding for gC II-5-His in the baculovirus expression vector pFastBac1 was provided by Dr. Beata Adamiak (Tab. 2). Sequencing showed a methionine-to-valine mutation at position 357, which required correction. Therefore, overlap extension PCR was used. Primers 1 and 4 (gC_fw_Xho2 and gCHis_rv_Bam3, respectively; see Tab. 3) used for this approach were equipped with the restriction sites for *Xho*I and *Bam*HI as well as their specific required overhangs. The PCR products (Fig. 14A) were purified using a QIAquick® purification kit (Qiagen).

Both the insert and the vector, which was dephosphorylated with SAP to prevent religation, were digested with *Xho*I and *Bam*HI. The digests were purified by ethanol precipitation, and the gC II-5-His insert was ligated into the pcDNA 3.1(-) vector. Before transformation into *E. coli* XL10 Gold, the ligation reactions were digested with *Not*I to linearize empty vector and minimize false positives. Clones were used for overnight cultures, and plasmid DNA was purified and screened for inserts by analytical PCR (Fig. 14B). Positive clones were sent to MWG (Ebersberg) for sequencing. Clone # 10 contained the correct sequence. An *E. coli* XL10 Gold glycerol stock was created from this clone, and the purified plasmid DNA was used for the transfection of CHO Lec 3.2.8.1. cells for protein expression.

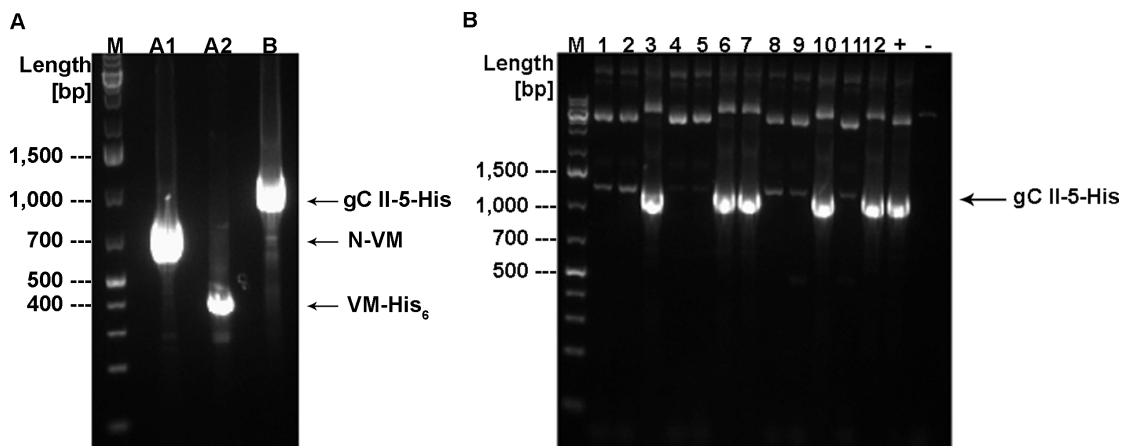


Fig. 14: Cloning of gC II-5-His into pcDNA 3.1(-). **A** Agarose gel electrophoresis of the overlap extension PCR to generate the gC II-5-His insert and remove the M357V mutation. M: Marker; A1: Reaction A1; A2: Reaction A2; B: Preparative PCR with templates from reactions A1 and A2; N-VM: Fragment 1 covering the template from the N-terminus to the unwanted M357V mutation; VM-His₆: Fragment 2 covering the template from the M357V mutation to the C-terminus. **B** Agarose gel electrophoresis of the analytical PCR to screen for plasmids containing the gC II-5-His insert. Clone # 10 proved a match upon sequencing. M: Marker; 1-12: Screened clones 1 to 12; +: Positive control (gC II-5-His in pcDNA 3.1(-) containing a sequence error); -: Negative control (MuPyV RA His₆-VP1* in pET15b).

4.2.1.2 *Cloning of gC II-5-Fc into pcDNA 3.1(-)*

After the gC II-5-His construct proved unsuitable for protein expression (see section 4.2.2.1), a new construct containing a cleavable C-terminal Fc-tag was designed. The insert was generated by MWG (Ebersberg) and provided in a pUC57 vector (see 5.4 for vector map). The insert is 1,875 bp long and was amplified by preparative PCR (Fig. 15A) using primers that introduced *Xho*I and *Bam*HI endonuclease restriction sites at the 3' and 5' ends, respectively (gCII-5_opt_Xho_fw1 and

gCII-5_opt_Bam_rv1; see Tab. 3). Restriction digests, ligation, analytical PCR (Fig. 15B), and sequencing were carried out as for the gC II-5-His construct. Sequencing showed clone # 11 to carry the correct sequence, and this clone was used for protein expression in CHO Lec 3.2.8.1. cells.

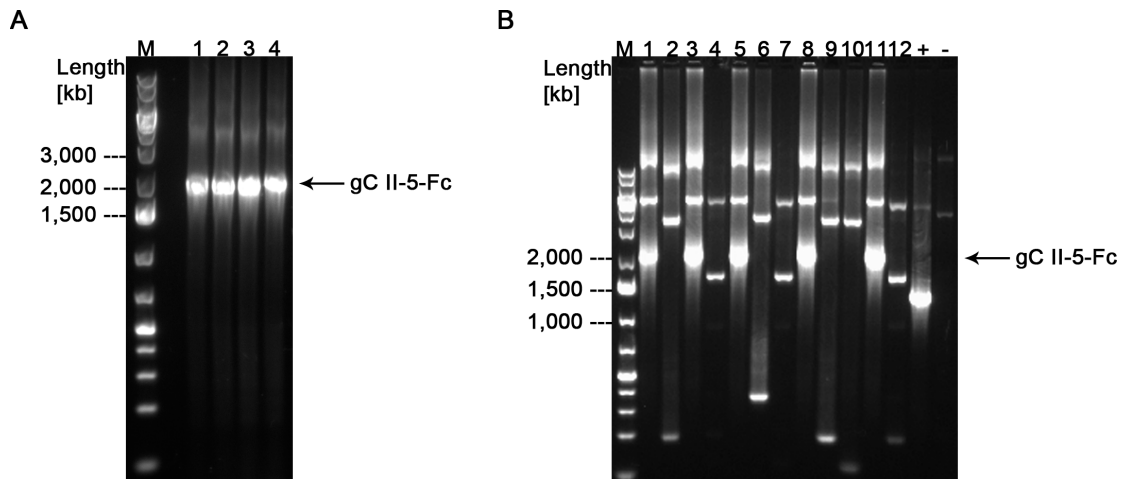


Fig. 15: Cloning of gC II-5-Fc into pcDNA 3.1(-). **A** Agarose gel electrophoresis of preparative PCR of gC II-5-Fc in pUC57. M: Marker; 1-4: PCR reactions 1 to 4. **B** Agarose gel electrophoresis of analytical PCR of gC II-5-Fc in pcDNA 3.1(-). Clones 1, 3, 5, 8, and 11 show amplicates of the desired size. Upon successful sequencing, clone #11 was used for protein expression. M: Marker; 1-12: Clones 1-12; +: positive control (Nectin4-V-Fc in pcDNA 3.1(-); -: negative control (MuPyV RA His₆-VP1* in pET15b)

4.2.2 ***Expression of gC II-5 in CHO Lec 3.2.8.1.***

In order for them to express the desired protein, CHO Lec 3.2.8.1. cells have to be transfected with the expression vector carrying the desired insert. In this study, transfections were carried out by lipofection using Lipofectamine® 2000. All transfections were carried out with circular plasmid DNA as well as with linearized DNA. For linearization, gC II-5-containing pcDNA 3.1(-) vectors were digested with *PvuI*. After transfection, cells were treated with medium containing G418 as a selection marker. Once a G418-resistant, polyclonal cell line was established, single cells were isolated to generate monoclonal colonies that display a more consistent protein expression profile. The presence of protein in the cell culture supernatant was monitored by Western Blot for gC II-5-His and by Protein A/G pull-down followed by Western Blot for gC II-5-Fc.

4.2.2.1 ***Expression of gC II-5-His***

During the isolation of single cells, only four colonies, S1, S2, L1, and L2, could be obtained. These four clones were grown in 25 cm² flasks using G418-free medium, because G418 inhibits the elongation step during protein expression. The presence of gC II-5-His in the cell culture supernatant was analyzed by Western Blot. Two blots with different primary antibodies were carried out: (1) the mouse hybridoma cell-derived α -gC monoclonal antibody (B1C1B4) that was purified in the course of this study (see section 4.2.3). When not even a positive signal with a commercial gC sample could be obtained, a murine α -His antibody (2) was used.

Unfortunately, no protein expression could be detected in any of the clones (Fig. 16). Therefore, the gC II-5-His construct was not pursued any further.

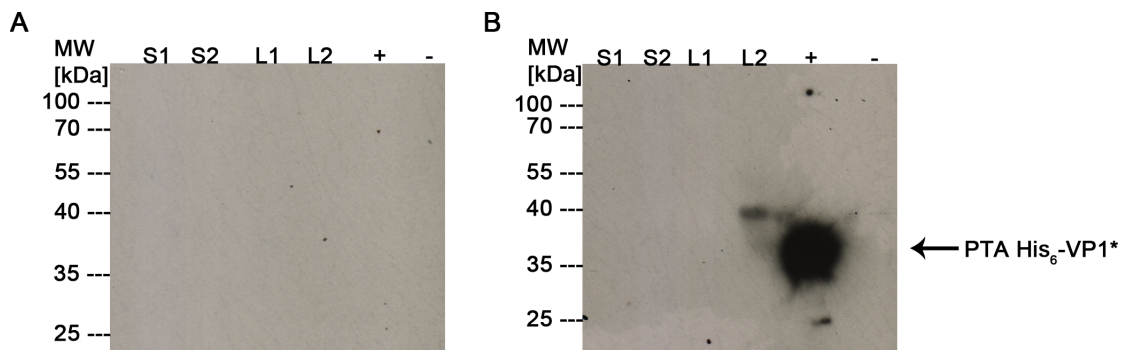


Fig. 16: Expression of gC II-5-His in CHO Lec 3.2.8.1. cells. Neither an antibody against gC itself or against the C-terminal His-tag of gC II-5-His could detect any protein expression. MW: Molecular weight; S1, S2, L1, L2: Clones from transfections with supercoiled and linearized vector, respectively; +: positive control; -: negative control. **A** Western Blot stained with B1C1B4. A commercially produced gC sample was used as a positive control, MuPyV RA VP1* as a negative control. Strangely, even the positive control did not give a signal. **B** Western Blot stained with a murine α -His-tag monoclonal antibody. MuPyV PTA His₆-VP1* was used as a positive control, RA VP1* as a negative control.

4.2.2.2 Expression of gC II-5-Fc

gC II-5-Fc was utilized after gC II-5-His showed no protein expression in all transfection attempts. While transfections with supercoiled plasmid resulted in four clones, two single colonies could be obtained from the linearized transfection. Protein expression levels from these four clones as well as from the polyclonal transfections were analyzed by a Protein A/G pull-down assay followed by a Western Blot. The primary antibody used for the blot was a murine monoclonal antibody specific for the Fc part of human IgG-1 (α -Fc). Clones S3, S4, L1, and L2 displayed no detectable protein expression (Fig. 17), but clones S1 and S2 as well as the polyclonal transfection with supercoiled plasmid showed protein bands around 85 kDa, which correspond to glycosylated gC II-5-Fc. The identity of the bands was confirmed by Bushra Amin and Sophie Stotz (laboratory of Dr. Hubert Kalbacher) using MALDI-MS (see 5.5). Protein production and purification were attempted with cell culture supernatants from both S1 and S2; ultimately, S1 was shown to express gC II-5-Fc at a higher level than S2.

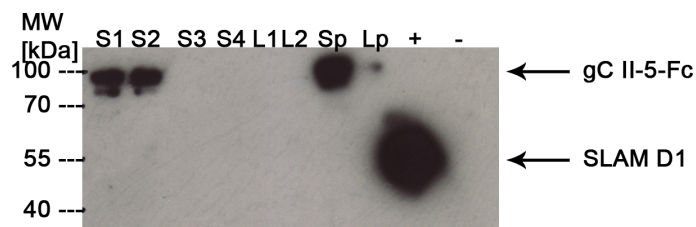


Fig. 17: Expression of gC II-5-Fc in CHO Lec 3.2.8.1. cells. Clones S1 and S2 display bands around 85 kDa, which correspond to gC II-5-Fc. The multiple bands seen here are typical for various glycosylation states of glycoproteins. MW: Molecular weight; S1-S4: Four singular clones transfected with supercoiled plasmid; L1, L2: Two singular clones transfected with linearized plasmid; Sp: polyclonal cells transfected with supercoiled plasmid; Lp: polyclonal cells transfected with linearized

plasmid; +: positive control (Fc-tagged SLAM D1 produced from CHO Lec 3.2.8.1. cells, provided by Dr. Georg Zocher); -: negative control (MuPyV RA VP1*).

4.2.3 Purification of the B1C1B4 Monoclonal α -gC Antibody

B1C1B4 is a monoclonal antibody developed by the laboratory of Prof. Dr. Tomas Bergström that selectively reacts with HSV-1 gC. It was expressed in mouse hybridoma cells grown in dialysis tubing. The antibody was purified from cell culture supernatant by Protein G affinity chromatography and analyzed by size exclusion chromatography.

4.2.3.1 Protein G Affinity Chromatography

Like Protein A from *S. aureus*, which was used for the purification of gC II-5-Fc, Protein G binds immunoglobulins with nanomolar affinity [255], but with different specificities. It displays a higher affinity for murine IgG1, which makes it highly suitable for the purification of B1C1B4.

The antibody-containing cell culture medium was centrifuged to remove any cell debris. The supernatant was diluted 1:1 in B1C1B4 binding buffer and loaded onto a 1 mL HiTrap™ Protein G HP column (GE Healthcare) over night. Then, the column was transferred to an ÄKTA Prime system and washed. The antibody was eluted using a B1C1B4 elution buffer gradient. Fig. 18 shows a typical SDS-PAGE and chromatogram of a purification run. The pooled fractions display a high amount of purity. They were stored at -80°C.

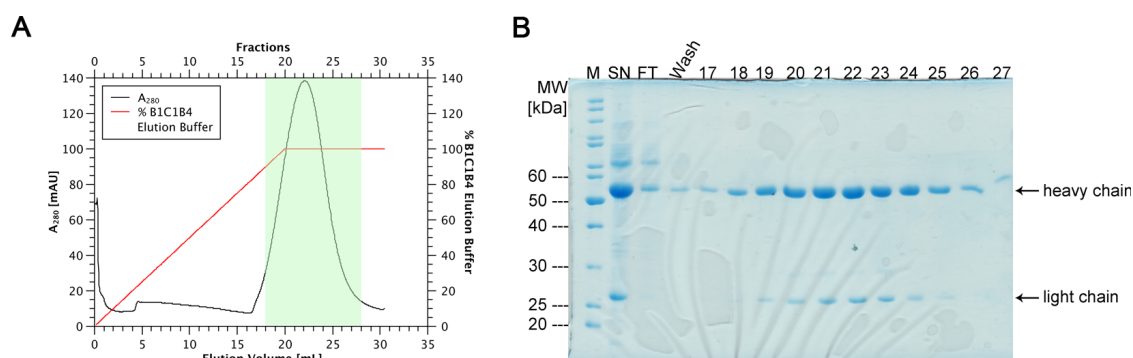


Fig. 18: B1C1B4 purification by Protein G affinity chromatography. **A** Chromatogram from a typical chromatography run. Fractions 18-27 (shaded green) were analyzed by SDS-PAGE (**B**), pooled, and stored at -80°C. **B** SDS-PAGE of a Protein G purification run. The analyzed fractions (18-27) show a high purity of heavy chain and light chain of the B1C1B4 antibody. M: Marker; SN: mouse hybridoma cell culture supernatant that was loaded onto the Protein G column; FT: loading flowthrough.

4.2.3.2 Analytical Size Exclusion Chromatography

After B1C1B4 was purified by Protein G affinity chromatography, its purity and assembly state were further analyzed by analytical size exclusion chromatography. For this, 23 μ L of the pooled antibody were filtered and loaded on an equilibrated Superdex 200 column installed in an Ettan system. The pristine elution peak at 1.27 mL corresponds to 180 kDa (Fig. 19), which is about the molecular weight of two heavy chains and two light chains (about 165 kDa). This demonstrates high purity of the antibody.

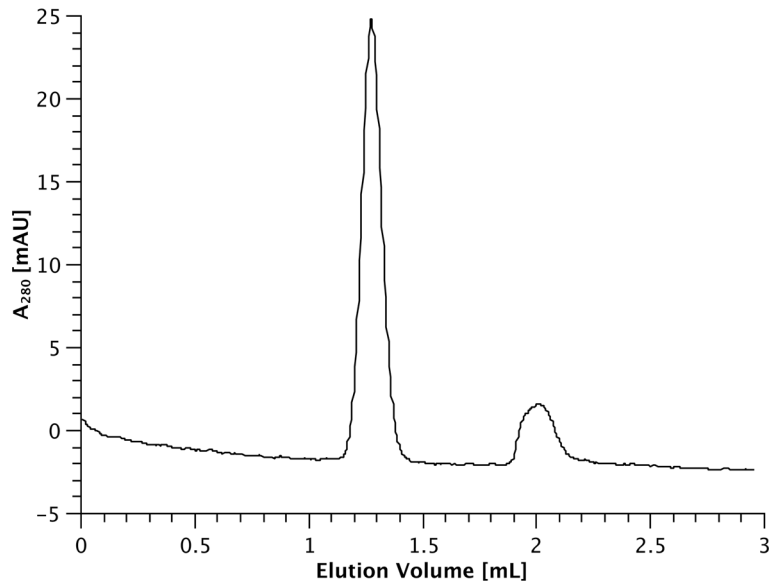


Fig. 19: B1C1B4 analytical size exclusion chromatography. The B1C1B4 antibody was purified by Protein G affinity chromatography and analyzed with an Superdex 200 column. The peak at 1.27 mL corresponds to 180 kDa. The second peak at around 2 mL can be attributed to tris in the neutralization buffer.

4.2.4 ***Purification of gC II-5-Fc***

gC II-5 was expressed for structural analysis using X-ray crystallography. Due to the signaling sequence spanning residues 1-32, the protein was secreted into the cell culture medium and then purified by Protein A affinity chromatography. Since the Fc-tag introduces non-native dimerization, it was cleaved off with TEV protease. The Fc-tag and the protease were removed by applying the solution onto a Protein A column and a HisTrap column, respectively, and the resulting gC II-5 solution was subjected to size exclusion chromatography (Fig. 20). Fully purified protein was stored at -20°C. The purification strategy was adapted from one established by Steffen Müller to purify UL16 [236].

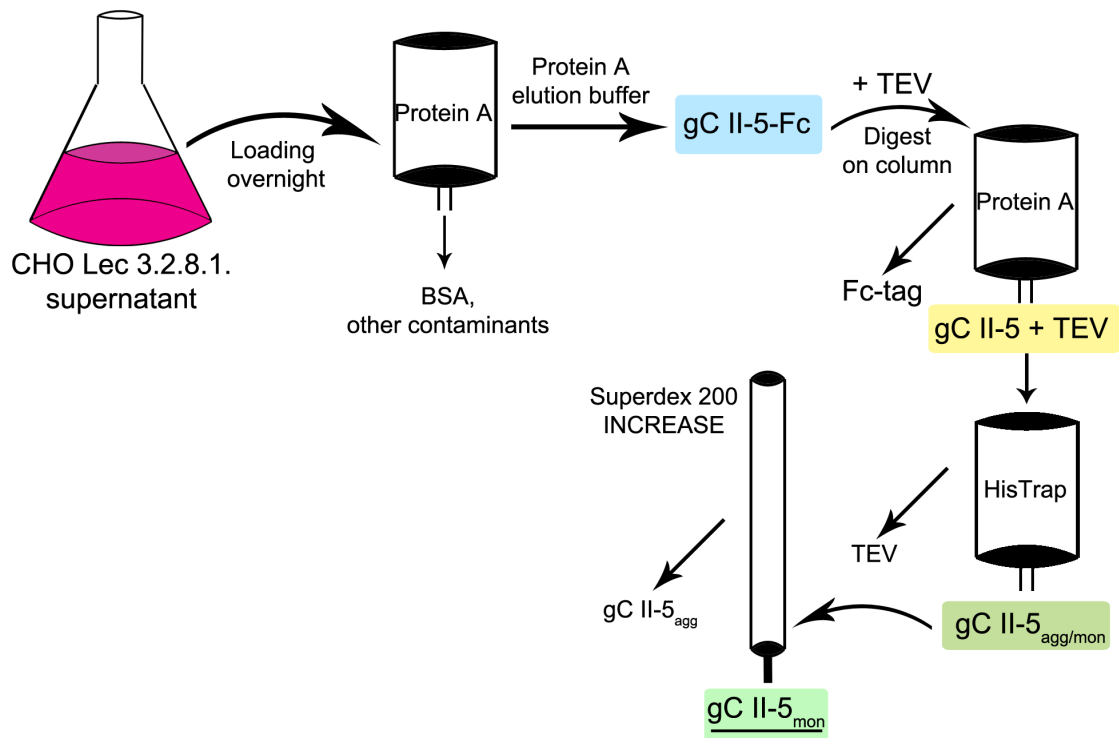


Fig. 20: General purification strategy for gC II-5. Harvested cell culture supernatant from CHO Lec 3.2.8.1. cells was loaded over night onto a Protein A column, and gC II-5-Fc was eluted with elution buffer at pH 2.8. The gC II-5-Fc containing eluate was again loaded onto the Protein A column together with TEV protease and incubated at 4°C. The cleaved gC II-5 was then loaded onto a HisTrap column, which removed any remaining TEV protease. The resulting mixture of monomeric gC II-5 (gC II-5_{mon}) and aggregated gC II-5 (gC II-5_{agg}) was further purified by size exclusion chromatography on a Superdex 200 INCREASE column.

4.2.4.1 *Protein A Affinity Chromatography*

The CHO Lec 3.2.8.1. supernatant used for purification also contained contaminants such as dead cells or precipitated BSA. To obtain feasible amounts of protein, purification was only carried out when the supernatant contained medium from cells that were at least three weeks old. Before loading, the supernatant was briefly centrifuged and subsequently passed through 5 µm, 0.45 µm, and 0.22 µm filters to remove BSA and cell debris, followed by a pH adjustment to 9 using NaOH. The filtered supernatant was loaded onto a HiTrap™ Protein A HP column over night, and the column was afterwards washed sequentially with binding buffer and binding buffer supplemented with 20% elution buffer. gC II-5-Fc was eluted with 100% elution buffer and collected in 1 mL fractions. To counteract the low pH of the elution buffer, 150 µL neutralization buffer were added to each fractionation tube beforehand. Peak fractions were analyzed by SDS-PAGE (Fig. 21), the protein concentration was assessed using a NanoDrop™, and the protein was frozen at -20°C. Each purification run usually yielded about 3-4 mg gC II-5-Fc.

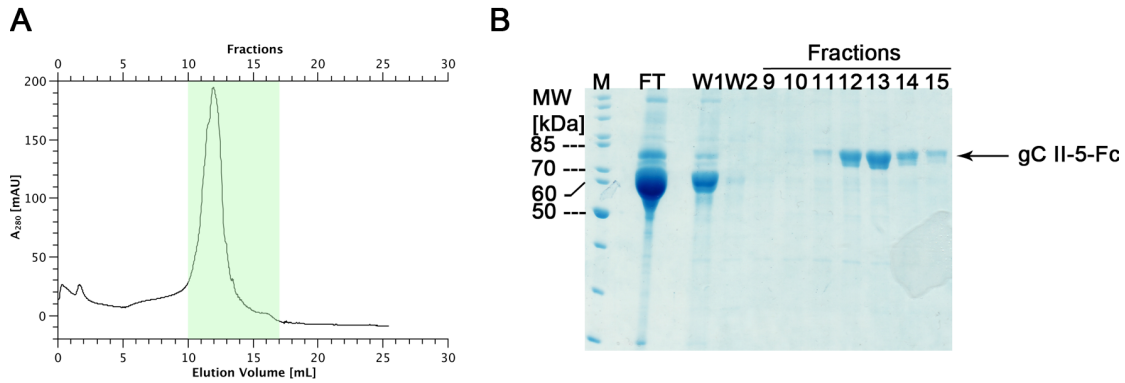


Fig. 21: Protein A affinity chromatography of gC II-5-Fc. **A** Chromatogram of a typical purification run. gC II-5-Fc was eluted with 100% elution buffer. The peak fractions (shaded light green) were analyzed by SDS-PAGE (**B**), pooled, and frozen at -20°C . **B** SDS-PAGE. gC II-5-Fc bands appear in the peak fractions around 85 kDa, which is consistent with the bands observed during Western Blot analysis (Fig. 17). The multiple bands demonstrate the various glycosylation states of the protein. M: Marker; FT: Loading Flowthrough; W1: Wash #1 with binding buffer; W2: Wash #2 with 20% elution buffer; 9-15: Collected fractions.

4.2.4.2 TEV Protease Cleavage

The Fc-tag needed to be proteolytically removed to obtain protein suitable for crystallization. Cleavage of the tag was attempted in solution as well as on-column. While incubation with TEV protease at 4°C for 44 hours yielded cleaved gC II-5 for both strategies (Fig. 22A), gC II-5 swiftly precipitated in solution. On the other hand, cleavage on-column yielded completely cleaved protein, and digested gC II-5 could be eluted from the column with binding buffer. Western Blot analysis only showed the Fc-tag in the eluate with elution buffer, but not in the flowthrough, which demonstrates complete cleavage (Fig. 22B).

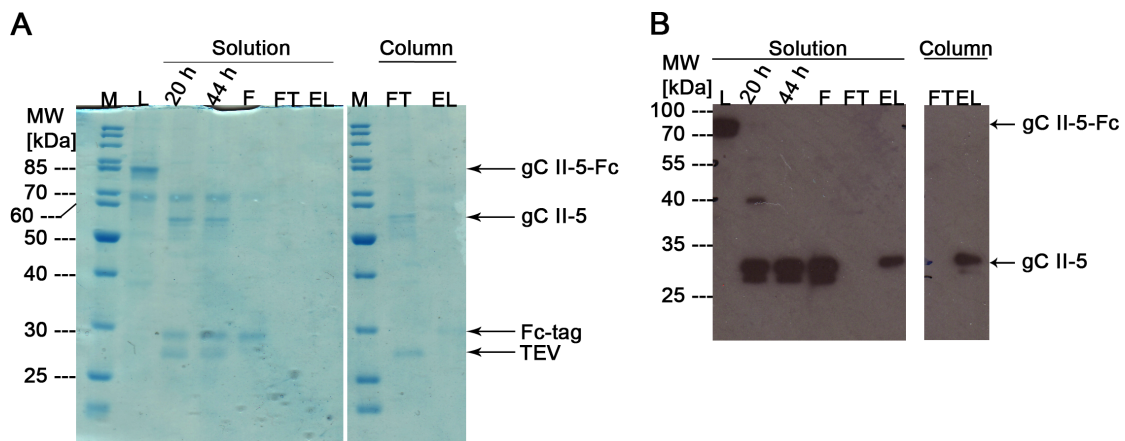


Fig. 22: TEV protease cleavage of gC II-5-Fc. **A** SDS-PAGE of digestion trials in solution and on-column. After 20 hours, no gC II-5-Fc could be detected anymore, while untagged gC II-5 at about 55 kDa as well as the Fc-tag at 29 kDa appeared. The band at about 27 kDa corresponds to TEV protease. Unfortunately, when incubated with TEV protease in solution, gC precipitated. Accordingly, almost no gC II-5 could be detected when the digest was filtered (F). When the filtrate was run over a Protein A column to separate gC II-5 and the Fc-tag, no gC II-5 could be detected in the flowthrough (FT), and very little Fc-tag could be detected in the eluate (EL). When incubated with TEV protease on-column, however, cleaved gC II-5 could be obtained when the column was washed with binding buffer (FT).

Very little Fc-tag could be detected in the eluate with elution buffer (EL), which is most likely due to a higher dilution. M: Marker; L: Load, i.e. uncleaved gC II-5-Fc. **B** The same samples shown in **A** were analyzed by Western Blot in combination with the α -Fc antibody. While the antibody shows a band at about 85 kDa for gC II-5-Fc in solution, a band for the Fc-tag is visible after TEV protease cleavage at about 29 kDa. The same holds true for on-column cleavage, where the Fc-tag is only visible upon elution with elution buffer. Multiple bands reflect the different glycosylation stages of gC II-5-Fc and the Fc-tag.

When subjected to analytical size exclusion chromatography, gC II-5 showed an aggregate peak at 0.9 mL. Most of the protein eluted at 1.44 mL, which corresponds to a molecular weight of about 85 kDa, consistent with an elongated monomer (Fig. 23). Digested and eluted gC II-5 was further subjected to preparative size exclusion chromatography.

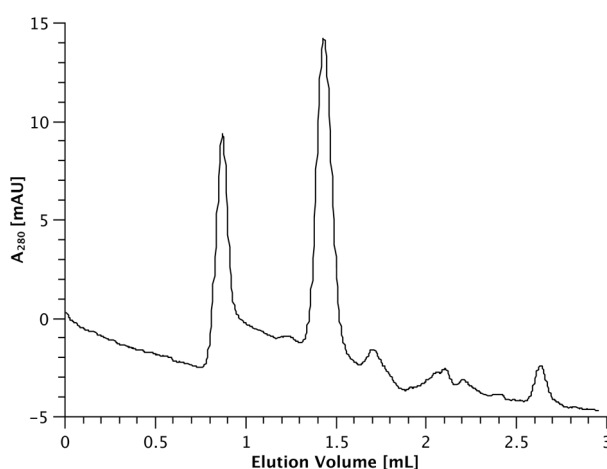


Fig. 23: Analytical size exclusion chromatography after TEV protease cleavage. gC II-5 with its Fc-tag removed was eluted from the Protein A column with binding buffer and applied to a Superdex 200 column. The peak at 0.9 mL corresponds to aggregated protein, the one at 1.44 mL to a molecular weight of about 85 kDa, which is consistent with an elongated gC II-5 monomer.

4.2.4.3 *Size Exclusion Chromatography*

After the Fc-tag was cleaved off from gC II-5-Fc with TEV protease, the untagged protein was subjected to size exclusion chromatography to separate different species and remove any aggregate. Due to the size of gC II-5, a Superdex 200 INCREASE column mounted on a BioLogic DuoFlow™ system was chosen. Before the chromatography run, the cleaved protein was concentrated in an Amicon protein concentrator. Fig. 24A depicts a typical chromatogram. Peak fractions were analyzed by SDS-PAGE (Fig. 24B), and fractions containing pure gC II-5 were pooled, their concentration was measured with a NanoDrop, and they were frozen away at -20°C. The elution peak at 13.4 mL corresponds to a molecular weight of about 75 kDa and therefore an elongated monomer of gC II-5. This finding was confirmed by analytical size exclusion chromatography of the pooled samples, which shows a pristine gC II-5 monomer peak at an elution volume corresponding to approximately 80 kDa, with little to no aggregation (Fig. 24C).

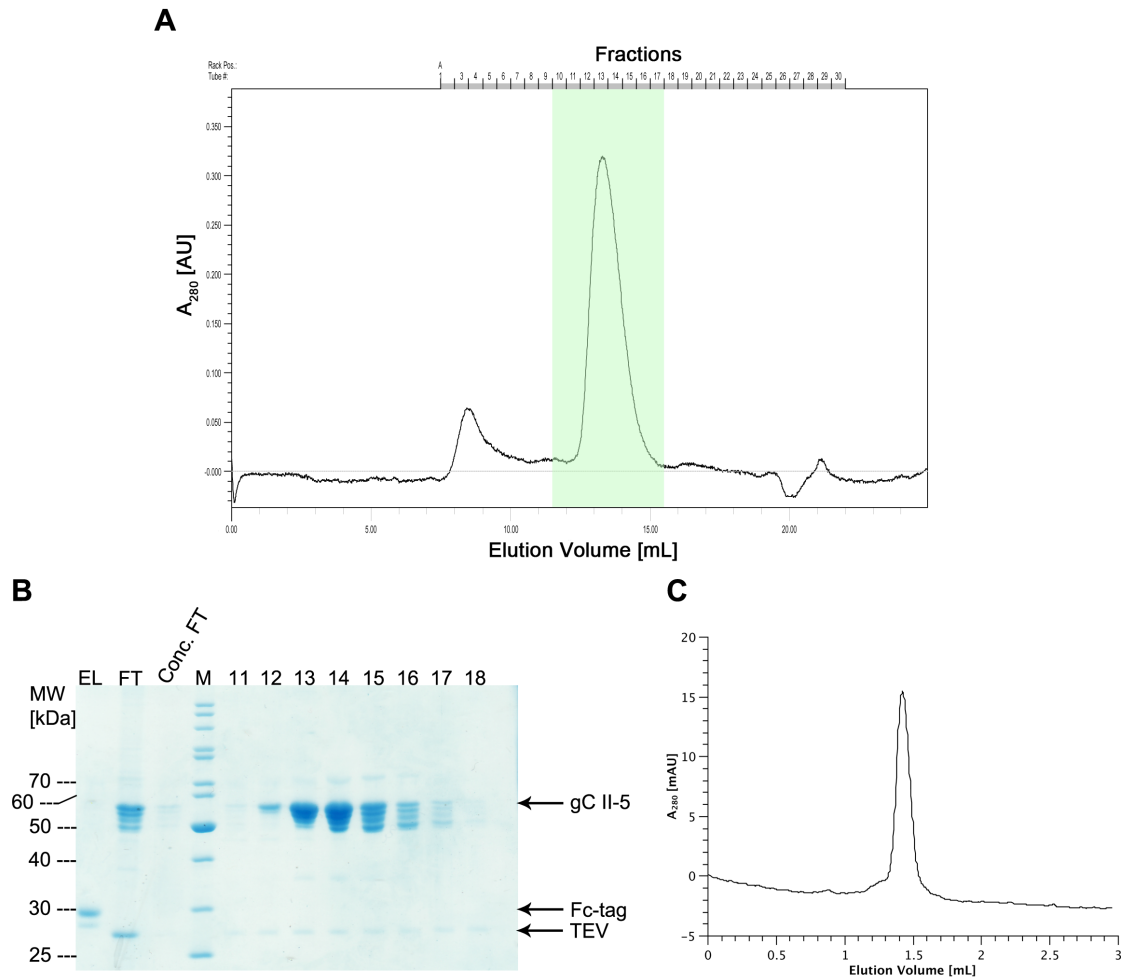


Fig. 24: Preparative size exclusion chromatography after TEV protease cleavage. **A** After removal of the Fc-tag with TEV protease, gC II-5 was subjected to size exclusion chromatography on a Superdex 200 INCREASE column. The small peak at about 8.5 mL (fraction 3) corresponds to aggregated protein, while the larger peak at 13.4 mL corresponds to a 72 kDa entity on this column, likely an elongated monomer of gC II-5. The analyzed and later pooled and stored fractions are highlighted green. **B** SDS-PAGE of TEV protease cleavage, protein concentration, and subsequent size exclusion chromatography. The cleaved Fc-tag could be eluted from the Protein A column with elution buffer (EL), while the cleaved protein, along with TEV protease could be washed off with binding buffer (FT) beforehand. The flowthrough of the used concentrator shows a bit of protein (Conc. FT), indicating that it was somewhat leaky. The analyzed size exclusion chromatography fractions show the different glycosylation states of gC II-5 and some TEV protease (which was not removed by nickel affinity chromatography at the time) still present. M: Marker. **C** Analytical size exclusion chromatography of fully purified gC II-5. The peak at 1.45 mL corresponds to about 80 kDa, an elongated gC II-5 monomer.

4.3 DISCUSSION AND OUTLOOK

There are nine species of herpesviruses, distributed across various families and subfamilies, that infect humans as their most common pathogens, and once they gain access to an organism, they establish life-long persistency with periodically occurring infections. The most common of these viruses is herpes simplex type 1 (HSV-1), which is known for the cold sores around the mouth and eyes that accompany its outbreaks. This large dsDNA virus has a very complicated life cycle, and its attachment and entry into a host cell are mainly mediated by at least six of the twelve glycoproteins embedded in the viral membrane.

This dissertation focuses on glycoprotein C, which is responsible for the initial attachment to GAGs, preferentially HS or CS. Once the virus has attached, binding of gD to either HVEM, Nectin-1 or 2, or a specially sulfated HS species recruits a gH-gL heterodimer, resulting in the activation of the fusogenic activity of gB when it binds to PILR α , MAG, or NM-IIA/B. In subsequent steps, the tegument proteins and the viral capsid that hosts the genome are released into the cell, and replication commences. gC is a multi-function protein; it not only mediates attachment (a task that can also be performed by gB), but it also helps control release of progeny virus from a lysed cell via its mucin-like domain. Also, it protects the virus by interfering with the complement system by binding to C3b and by inhibiting properdin binding to the C3 convertase. Finally, a monoclonal antibody, B1C1B4, was shown to inhibit HSV-1 infection by binding to gC. For these reasons the structure of gC was to be solved. This work shows the first steps towards achieving this goal.

A construct provided by Dr. Beata Adamiak and Prof. Dr. Tomas Bergström was equipped with a C-terminal Fc-tag that could be cleaved off with TEV protease and expressed in CHO Lec 3.2.8.1. cells. gC II-5-Fc was purified from the cell culture supernatant by Protein A affinity chromatography, which yielded about 1 – 1.5 mg of already relatively pure protein per liter of supernatant. However, the high serum content of the cell culture medium caused significant contamination with BSA during purification and could even damage Protein A columns at low loading speed by unspecific binding and subsequent clogging. Even though this problem could be avoided by increasing loading speeds, serum-free medium should be tried out in future attempts. The Fc-tag could be cleaved on-column with TEV protease, and the protein was successfully further purified by preparative size exclusion chromatography, although the fractions shown in this work still display minor contaminations with TEV protease; however, the protease is equipped with a His-tag, so these contaminations should be removable by nickel affinity chromatography.

The biggest challenge in X-ray crystallography is obtaining high-quality crystals that yield high resolution for structure determination. Since one cannot predict the conditions under which any given protein crystallizes, the community is still rather dependent on a trial-and-error principle. This method, in turn, requires relatively large amounts of protein; depending on the protein in question, this prerequisite can already be troublesome by itself. The results presented in this work provide important groundwork for the production of a suitable amount of gC II-5 for crystallization trials. It is not clear at this point whether the four glycosylation sites should be removed or not. Glycosylation, especially at residue N148, has been reported to be critical for HS binding [169], and it is reasonable to assume that the

presence of carbohydrates increase the solubility and stability of gC II-5. On the other hand, even homogenous, long carbohydrate chains are quite flexible, which can interfere with the formation of a regular lattice and thereby, crystal growth. During his BSc thesis, Felix Klostermann has attempted deglycosylation of gC II-5-Fc with the enzyme EndoH_f, which proved successful (Fig. 25).

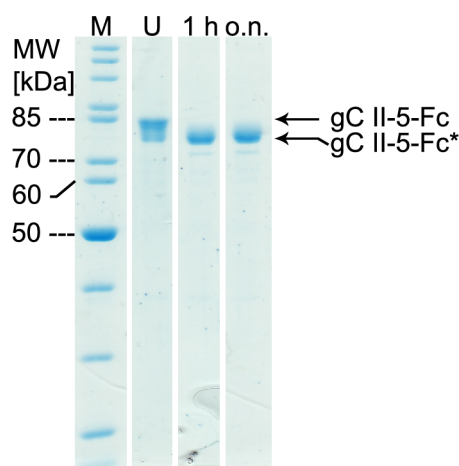


Fig. 25: Deglycosylation of gC II-5-Fc. Even after one hour, the EndoH_f digestion was complete, indicated by the band shift from roughly 85 kDa to about 80 kDa. M: Marker; U: untreated gC II-5-Fc; 1 h: digest after one hour; o.n.: digest over night; gC II-5-Fc: glycosylated gC II-5-Fc; gC II-5-Fc*: deglycosylated gC II-5-Fc. Figure adapted and modified, with permission, from Felix Klostermann [256].

X-ray crystallographic studies of gC should focus on a number of different interactions. As was discussed by Altgärde and colleagues [165], the deletion of the mucin-like region in gC II-5 should decrease the overall affinity for HS, but the binding should be less reversible than in naturally occurring glycoprotein C. Perhaps this structure will also to some degree explain how the absence of the mucin-like region hampers efficient release of progeny virus from the parent cell. It should be kept in mind, though, that the studies of the HS interactions will most likely not be without complications, because (1) the isolation and synthesis of homogenous GAGs is still a somewhat problematic issue and (2) GAGs usually bind to proteins through ionic interactions, which may result in non-specific binding modes, generating heterogeneity [206]. However, this endeavor should be pursued – binding of GAGs is a strategy employed by a number of virus families (it also emerges for polyomaviruses [218]; see the first half of this dissertation), and a general understanding of the mechanisms behind it will greatly further our understanding of them. The X-ray structure of gC will also be helpful in understanding how it interferes with the C3b-mediated complement system. The established, successful, and reproducible purification strategy presented here is an important step towards achieving these goals and furthering our understanding of this multifunctional protein.

5. APPENDIX

5.1 ACKNOWLEDGEMENTS

This work could not have been done without the help of a large number of people I would like to thank here, so this will take a bit. Pardon the mix of English and German.

Thilo Stehle – I remember a discussion I had with you in the beginning of 2011, when I asked you if I could be your PhD student. At the time I was a bit worried whether it might be a better idea to visit another lab for a change. Today I can say I am very happy that I stayed! These last years have been the most interesting, fun, and adventurous I have had yet. I am very grateful for your competent and positive supervision, great feedback, and advice. Thank you!

Jack Johnson – In 2009 I was very lucky to be able to spend three months in your lab in La Jolla, and these months have stuck with me to this day. I was thrilled in 2011 when you agreed to be the second advisor for this thesis. I have always enjoyed our annual discussions, especially when they would happen in California, and I am truly humbled by your generosity when I invited myself over. Finally, I greatly appreciate your always optimistic and motivating advice that far exceeds the scope of this work. I have made a wise choice in 2011!

Ursula Neu – You have taken me under your wing in early 2008 and have taught me many of the things I know today. I consider myself very lucky to have been able to work together with you and wish you all the best from the bottom of my heart for the paths that lie ahead of you!

Manuel Liaci – Ich hatte Ende 2011 schon ein wenig Bammel, gleich einen Diplomanden zu betreuen. Allerdings konnte ich auch nicht absehen, was mir da für ein talentierter und fleißiger Mensch an die Seite gestellt wird. Ich denke, wir haben aus etwas, das ja eigentlich nur noch als Nebenprojekt laufen sollte, eine tolle Story gemacht, und ich bin Dir für Deine Hilfe, Deinen Input sowie Deine Freundschaft und Ehrlichkeit über die Jahre sehr dankbar!

Luisa Ströh – You have been a constant in the polyomavirus world for me for almost as long as Ulla. Not only do I thank you for your supervision way, way, way back during a portion of my Diploma thesis, but also for letting me help out in the HPyV 6/7 project and for allowing me to constantly pester you with questions about PyVs. I also thank you for your valuable input during the density integration portion of the MuPyV project and, of course, for your advice concerning the PyV part of this work. I wish you all the best for the future!

Irmgard Hähnlein-Schick – Ich danke Dir von Herzen für die angenehme Arbeitsatmosphäre an Roboter und Zellkultur und darüber hinaus sowie die freundschaftlichen Gespräche beim Kaffee (sofern ich es so früh geschafft habe, ins Labor zu kommen). Außerdem hast Du mir de facto alles beigebracht, was ich heute

über Zellkultur im Allgemeinen und das Hegen und Pflegen von CHO Zellen im Speziellen weiß. Das nächste Konzert kommt bestimmt!

Georg Zocher – Du bist ein würdiger Gegner! Ich danke Dir für die Hilfe in allem, was mit Linux zu tun hat (auch wenn andere auch einmal helfend zur Seite standen) und dafür, dass mir in diesen Jahren meine Schlagfertigkeit nicht eingerostet ist. ☺

Im gleichen Atemzug bedanke ich mich auch noch bei **Melanie Dietrich**, die mich mit Georg zusammen motiviert hat, endlich mal wieder etwas für meine Fitness zu tun. Ohne Euch zwei hätte ich mich dazu, glaube ich, nie wieder aufgerafft, und das weiß ich mehr zu schätzen, als ich es hier ausdrücken kann. Nebenher war es schön, am Wochenende immer wieder jemanden zu haben, mit dem man ab und zu einen Kaffee trinken kann.

Sonja Rakette – Auch bei Dir bedanke ich mich dafür, dass Du mich nach meinem Austauschsemester in La Jolla bei Dir als HiWi hast arbeiten lassen; aber auch für die Zeit danach und die oft wahrgenommenen Gelegenheiten, bei einem Bier oder zwei sich auch einmal den Frust von der Seele reden zu können, haben mir sehr viel bedeutet. Ich freue mich aufs nächste Wiedersehen!

Felix Klostermann – Auch wenn's am Ende leider nicht so ganz geklappt hat, wie wir uns das vorgestellt haben, danke ich Dir herzlich für Deine Arbeit an gC im Rahmen Deiner BSc-Arbeit und natürlich ganz besonders dafür, dass ich die Rohdaten des EndoH_f Verdaus hier zeigen darf.

I would like to thank the entire **Stehle group** for the awesome working environment and the countless activities that come with being a group member here. The 'Herrenabende' with Micha, Manu, Felix, Volker, Erik, and Nici come to mind; also, the legendary ASV meeting at Penn State with Kerstin, Melanie, and Georg. Great times! Finally, a shout-out to Joana, Felix, and Manu as my office mates.

I would like to thank **Bushra Amin, Sophie Stotz, and Hubert Kalbacher** for the MALDI-MS analyses of gC II-5-Fc. Special thanks to Sophie for allowing me to show her analysis in this thesis.

Xang a cappella – Sine musica nulla vita! Meinen lieben Freunden und Gesangskollegen **Andreas Bürker, Gerd Esche, Andreas Gut, Thorsten Kindermann, Michael Kölle** und **Hansjörg Wagner** ergeht ein besonderer Dank für die vielen tollen Stunden, die wir zusammen mit Musik gefüllt haben! Irgendwann wird die Tübinger Zeit für mich vorbei sein, und der Abschied von Xang wird mir mit ziemlicher Sicherheit mit am schwersten fallen.

Thomas Shiozawa – es macht saumäßigen Spaß, mit Dir, lieber Freund, Musik zu machen! Aber auch das gelegentliche Treffen am Wochenende ohne Musik, der Kampf gegen das Böse (am Computer, versteht sich) und auch die jährliche Völlerei am Umbrisch-Provencialischen Markt möchte ich ungern missen und ich bin froh, diese Sachen mit Dir machen zu können! Ich wünsche Dir und natürlich auch **Maggie** und **Leonard** von Herzen alles Gute!

Ich bedanke mich bei **meinen lieben Bundesbrüdern meiner Sängerschaften Hohentübingen zu Tübingen und Fridericiana zu Halle** für alles, was sie mir in den letzten 11 Jahren für mein Leben mit auf den Weg gegeben haben. Ohne Euch würde mir definitiv was fehlen!

Ein besonderer Dank geht an meine Freunde, die mich durch das Studium und auch schon vorher begleitet haben: **Daniel, Steffen, Nadine, Toni, Micha, Svenja, Simon, Benedikt, Kevin, Anja, Albrecht, Seyfe, Korbinian, Katha und Baschdl, Pinky, Rolf, Vollmer, Haak, Vera und Heiko, Gunnar und Laetitia**, sowie **Stephan und Caro**. Schön, dass es Euch alle gibt. Danke auch noch an Albrecht, dass er diese Arbeit am Ende korrekturgelesen hat und an Felix für die Hilfe beim Paper-Formatieren.

Zuletzt bedanke ich mich bei meiner **Familie** sowohl hier als auch auch in den USA dafür, dass sie immer zu mir gestanden haben und mich bei allem unterstützt haben, was ich mir vorgenommen habe. Ein besonderer Gruß an das **Samtschnäuzchen** darf natürlich auch nicht fehlen!

5.2 LICENSES FOR FIGURES FROM OTHER PUBLICATIONS

A number of figures in this dissertation were taken from original publications. The licenses for these figures or panels were obtained via the s100.copyright.com website of RightsLink® and are listed below.

5.2.1 ***Fig. 3A: Yan et al., 1996; Structure***

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Michael Buch
Customer address	Interfaculty Institute of Biochemistry D-72076 Tübingen, Baden-Württemberg
License number	3758730563836
License date	Nov 30, 2015
Licensed content publisher	Elsevier
Licensed content publication	Structure
Licensed content title	Structure determination of simian virus 40 and murine polyomavirus by a combination of 30-fold and 5-fold electron-density averaging
Licensed content author	Youwei Yan, Thilo Stehle, Robert C Liddington, Haiching Zhao, Stephen C Harrison
Licensed content date	February 1996
Order reference number	Lic1_Capsid1
Original figure numbers	Figure 1

5.2.2 ***Fig. 3B: Lee et al., 2008; Structure***

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Michael Buch
Customer address	Interfaculty Institute of Biochemistry D-72076 Tübingen, Baden-Württemberg
License number	3758731469078
License date	Nov 30, 2015
Licensed content publisher	Elsevier
Licensed content publication	Structure
Licensed content title	Virus Capsid Expansion Driven by the Capture of Mobile Surface Loops
Licensed content author	Kelly K. Lee, Lu Gan, Hiro Tsuruta, Crystal Moyer, James F. Conway, Robert L. Duda, Roger W. Hendrix, Alasdair C. Steven, John E. Johnson
Licensed content date	8 October 2008
Order reference number	Lic2_HK97
Original figure numbers	Fig. 1A, 1B

5.2.3. ***Fig. 6: Grünewald et al., 2003; Science***

License Number	3758741257902
License date	Nov 30, 2015
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	Three-Dimensional Structure of Herpes Simplex Virus from Cryo-Electron Tomography
Licensed content author	Kay Grünewald, Prashant Desai, Dennis C. Winkler, J. Bernard Heymann, David M. Belnap,

Licensed content date Wolfgang Baumeister, Alasdair C. Steven
Nov 21, 2003
Order reference number Lic4_HSVCryoTM

5.2.4 **Fig. 7: Bloom et al., 2010; Biochimica et Biophysica Acta – Gene Regulatory Mechanisms**

Supplier Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK
Registered Company Number 1982084
Customer name Michael Buch
Customer address Interfaculty Institute of Biochemistry
D-72076 Tübingen, Baden-Württemberg
License number 3758750084748
License date Nov 30, 2015
Licensed content publisher Elsevier
Licensed content publication Biochimica et Biophysica Acta (BBA) - Gene
Regulatory Mechanisms
Licensed content title Epigenetic regulation of latent HSV-1 gene
expression
Licensed content author David C. Bloom, Nicole V. Giordani, Dacia L.
Kwiatkowski
Licensed content date March–April 2010
Order reference number Lic5_HSVspread
Original figure numbers Fig. 1

5.2.5 **Fig. 11: Imberty et al., 2007; Carbohydrate Research**

Supplier Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK
Registered Company Number 1982084
Customer name Michael Buch
Customer address Interfaculty Institute of Biochemistry
D-72076 Tübingen, Baden-Württemberg
License number 3758750503054
License date Nov 30, 2015
Licensed content publisher Elsevier
Licensed content publication Carbohydrate Research
Licensed content title Structural view of glycosaminoglycan–protein
interactions
Licensed content author Anne Imberty, Hugues Lortat-Jacob, Serge Pérez
Licensed content date 26 February 2007
Order reference number Lic6_GAG
Original figure numbers Fig. 1

5.3 SEQUENCES

5.3.1 *gC II-5-His*

DNA sequence, which has been cloned into pcDNA 3.1(-). The base triplets for amino acids 32 and 124, which flank the omitted mucin-like region, are in bold face. The base numbering reflects this deletion. The first 90 nucleotides coding the signaling sequence are colored orange.

10	20	30	40	50	60
ATGGCCCCGG	GGCGGGTGGG	CCTTGCCGTG	GTCCTGTGGA	GCCTGTTGTG	GCTCGGGGCG
70	80	90	373	383	393
GGGGTGGCCG	GGGGCTCGGA	AACTGCCTCC	ACC GGGCCCCG	TGTGGTGCGA	CCGCCGCGAC
403	413	423	433	443	453
CCATTGGCCC	GGTACGGCTC	GCGGGTGCAG	ATCCGATGCC	GGTTTCGGAA	TTCCACCCGC
463	473	483	493	503	513
ATGGAGTTCC	GCCTCCAGAT	ATGGCGTTAC	TCCATGGGTC	CGTCCCCCCC	AATCGCTCCG
523	533	543	553	563	573
GCTCCCGACC	TAGAGGAGGT	CCTGACGAAC	ATCACCGCCC	CACCCGGGGG	ACTCCTGGTG
583	593	603	613	623	633
TACGACAGCG	CCCCAACCT	GACGGACCCC	CACGTGCTCT	GGGCGGAGGG	GGCCGGCCCCG
643	653	663	673	683	693
GGCGCCGACC	CTCCGTTGTA	TTCTGTCACC	GGGCCGCTGC	CGACCCAGCG	GCTGATTATC
703	713	723	733	743	753
GGCGAGGTGA	CGCCCGCGAC	CCAGGGAATG	TATTACTTGG	CCTGGGGCCG	GATGGACAGC
763	773	783	793	803	813
CCGCACGAGT	ACGGGACGTG	GGTGC GCGTC	CGCATGTTCC	GCCCCCGTC	TCTGACCCTC
823	833	843	853	863	873
CAGCCCCACG	CGGTGATGGA	GGGTCAGCCG	TTCAAGGCGA	CGTGCACGGC	CGCCGCCTAC
883	893	903	913	923	933
TACCCGCGTA	ACCCCGTGGA	GTTTGTCTGG	TTCGAGGACG	ACCGCCAGGT	GTTTAACCCG
943	953	963	973	983	993
GGCCAGATCG	ACACGCAGAC	GCACGAGCAC	CCCGACGGGT	TCACCACAGT	CTCTACCGTG
1003	1013	1023	1033	1043	1053
ACCTCCGAGG	CTGTCGGCGG	CCAGGTCCCC	CCGCGGACCT	TCACCTGCCA	GATGACGTGG
1063	1073	1083	1093	1103	1113
CACCGCGACT	CCGTGATGTT	CTCGCGACGC	AATGCCACCG	GGCTGGCCCT	GGTGCTGCCG
1123	1133	1143	1153	1163	1173
CGGCCAACCA	TCACCATGGA	ATTTGGGGTTC	CGGCATGTGG	TCTGCACGGC	CGGCTGCGTC
1183	1193	1203	1213	1223	1233
CCCAGGGGCG	TGACGTTTGC	CTGGTTCCCTG	GGGGACGACC	CCTCACCGGC	GGCTAAGTCG
1243	1253	1263	1273	1283	1293
GCCGTTACGG	CCCAGGAGTC	GTGCGACCAC	CCCGGGCTGG	CTACGGTCCG	GTCCACCCTG
1303	1313	1323	1333	1343	1353
CCCATTTTCGT	ACGACTACAG	CGAGTACATC	TGTCGGTTGA	CCGGATATCC	GGCCGGGATT
1363	1373	1383	1393	1403	1413
CCCGTTCTAG	AGCACCACGG	CAGTCACCAG	CCCCACCCA	GGGACCCAC	CGAGCGGCAG
1423	1433	1453	1462		
GTGATCGAGG	CGATCGAGCA	TCATCATCAT	CATCATTAA		

Amino acid sequence of gC II-5-His. The N-terminal signaling sequence is depicted in orange, the amino acids flanking the mucin-like region are in bold face.

10	20	30	131	141	151
MAPGRVGLAV	VLWSSLWLGA	GVAGGSETAS	TGPVWCDRRD	PLARYGSRVQ	IRCRFRNSTR
161	171	181	191	201	211
MEFRLQIWRY	SMGSPPIAP	APDLEEVLTN	ITAPPGGLLV	YDSAPNLTD	HVLWAEAGAP
221	231	241	251	261	271
GADPPLYSVT	GPLPTQRLII	GEVTPATQGM	YYLAWGRMDS	PHEYGTWVRV	RMFRPPSLTL
281	291	301	311	321	331
QPHAVMEGQP	FKATCTAAAY	YPRNPVEFVW	FEDDRQVFNP	GQIDTQTHEH	PDGFTTVSTV
341	351	361	371	381	391
TSEAVGGQVP	PRTFTCQMTW	HRDSVMFSRR	NATGLALVLP	RPTITMEFGV	RHVCTAGCV
401	411	421	431	441	451
PEGVTFAWFL	GDDPSPAAKS	AVTAQESCDH	PGLATVRSTL	PISYDYSEYI	CRLTGYPAGI
461	471	481	483		
PVLEHHGSHQ	PPPRDPTERQ	VIEAIEHHHH	HH		

5.3.2 gC II-5-Fc

DNA sequence, provided by Eurofins MWG Operon LLC (Ebersberg; see below), which has been cloned into pcDNA 3.1(-) in the course of this study. The base triplets flanking the mucin-like region are in bold face. The bases coding for the signaling sequence are depicted in orange, gC II-5 is colored black, the TEV cleavage site-containing linker is blue, and the Fc-tag is green.

10	20	30	40	50	60
ATGGCTCCAG	GTCGTGTTGG	ACTTGCCGTT	GTCCTGTGGA	GTCTCTTGTTG	GTTGGGAGCC
70	80	90	373	383	393
GGGGTAGCAG	GCGGATCAGA	AACCGCTTCA	ACTGGGCCAG	TGTGGTGCGA	CAGAAGAGAT
403	413	423	433	443	453
CCTCTTGCCC	GCTATGGATC	AAGGGTTCAG	ATTTCGTGTA	GGTTTCGTAA	CAGCACCAGG
463	473	483	493	503	513
ATGGAGTTCC	GTCTGCAGAT	TTGGAGATAT	TCCATGGGCC	CCAGTCCACC	CATTGCCCCT
523	533	543	553	563	573
GCTCCCGATC	TTGAGGAGGT	TCTCACGAAC	ATCACCGCTC	CACCCGGAGG	TTTGCTGGTC
583	593	603	613	623	633
TACGATAGTG	CCCCAAACCT	GACCGATCCT	CACGTA	GGCCGAAGG	GGCTGGCCCT
643	653	663	673	683	693
GGCGCAGACC	CTCCCCTGTA	CAGTGTAACC	GGACCTCTGC	CCACACAACG	ACTTATTATC
703	713	723	733	743	753
GGGGAAGTGA	CACCCGCCAC	ACAGGGAATG	TACTACCTGG	CTTGGGGCCG	CATGGATTCT
763	773	783	793	803	813
CCCCACGAGT	ATGGCACGTG	GGTCCGAGTG	CGGATGTTCC	GCCACCCAG	CCTGACACTG
823	833	843	853	863	873
CAGCCACATG	CCGTTATGGA	AGGACAGCCT	TTCAAGGCCA	CGTGTACCGC	CGCCGCCTAC
883	893	903	913	923	933
TATCCCCGGA	ATCCCGTGGA	GTTTGTTTGG	TTCGAGGACG	ATAGGCAAGT	GTTCAATCCC
943	953	963	973	983	993
GGCCAGATCG	ACACTCAGAC	CCACGAACAC	CCAGACGGAT	TCACTACTGT	GAGTACCGTG
1003	1013	1023	1033	1043	1053
ACCTCCGAAG	CTGTAGGCGG	TCAGGTGCCT	CCTCGGACAT	TCACTTGCCA	GATGACCTGG
1063	1073	1083	1093	1103	1113
CACCGAGACT	CCGTCATGTT	TTCACGCCGC	AATGCAACTG	GGCTTGCCTT	GGTACTTCCT
1123	1133	1143	1153	1163	1173
AGACCAACCA	TCACCATGGA	ATTTGGCGTG	CGACATGTTG	TCTGCACCGC	TGGTTGTGTC
1183	1193	1203	1213	1223	1233
CCCAGGGCG	TGACATTTGC	ATGGTTCCCTG	GGTGATGACC	CCAGTCCCGC	CGCTAAGAGC
1243	1253	1263	1273	1283	1293
GCTGTAACAG	CCCAGGAGTC	TTGCGATCAT	CCAGGGCTGG	CAACTGTCCG	GTCTACATTG

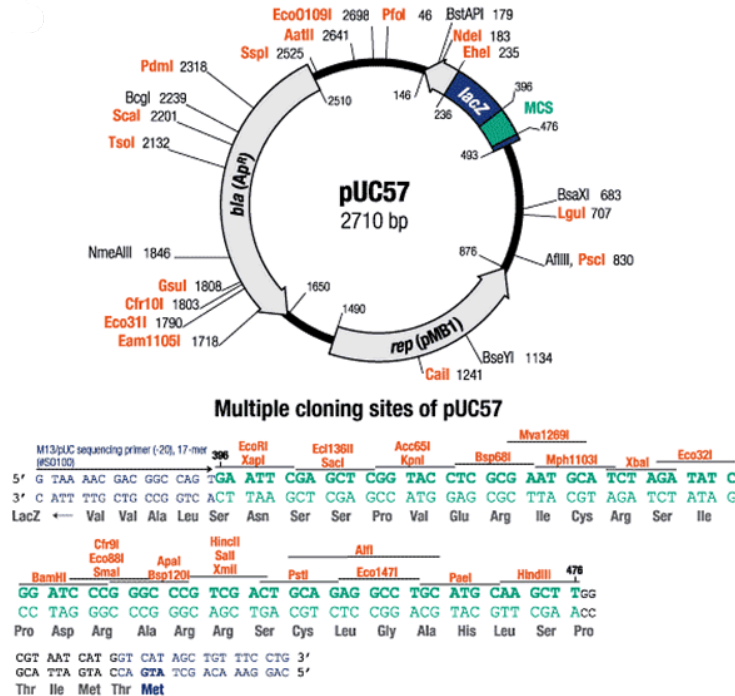
1303	1313	1323	1333	1343	1353
CCCATCTCCT	ATGACTACTC	CGAATACATT	TGCCGACTCA	CAGGCTATCC	CGCAGGGATA
1363	1373	1383	1393	1403	1413
CCTGTTCTGG	AACACCATGG	GTCACACCAA	CCTCCTCCCA	GGGACCCAC	AGAGAGGCAA
1423	1433	1443	1453	1463	1473
GTGATTGAGG	CCATCGAGGG	<u>AACTGAGAAC</u>	<u>CTGTACTTTC</u>	<u>AGGGGACCCA</u>	<u>CACTTGTCCA</u>
1483	1493	1503	1513	1523	1533
CCCTGTCCTG	CACCTGAGTT	GCTGGGTGGC	CCATCTGTGT	TCCTGTTTCC	TCCCAAGCCC
1543	1553	1563	1573	1583	1593
AAAGACACCC	TCATGATCTC	TCGGACACCA	GAGGTGACTT	GTGTGGTGGT	GGATGTGAGC
1603	1613	1623	1633	1643	1653
CATGAGGACC	CAGAGGTTAA	GTTTAACTGG	TACGTCGATG	GTGTGGAGGT	GCATAACGCT
1663	1673	1683	1693	1703	1713
AAGACGAAAC	CCAGAGAGGA	ACAGTACAAT	AGCACCTATC	GGGTGGTGTC	CGTGCTCACC
1723	1733	1743	1753	1763	1773
GTGCTGCATC	AGGATTGGCT	CAACGGTAAAG	GAATACAAAT	GCAAGGTGTC	CAACAAGGCT
1783	1793	1803	1813	1823	1833
CTGCCAGCAC	CAATCGAGAA	AACTATCAGC	AAAGCAAAAAG	GCCAACCTAG	GGAGCCTCAG
1843	1853	1863	1873	1883	1893
GTCTATACAC	TGCCACCCTC	TCGCGAAGAA	ATGACCAAGA	ATCAGGTCAG	CCTGACTTGC
1903	1913	1923	1933	1943	1953
CTCGTGAAAG	GCTTCTATCC	CTCTGACATA	GCCGTCGAAT	GGGAGTCTAA	TGGTCAGCCA
1963	1973	1983	1993	2003	2013
GAGAACAAC	ATAAGACTAC	ACCACCTGTA	CTGGACAGCG	ATGGGTCCCTT	CTTCTGTGAC
2023	2033	2043	2053	2063	2073
AGTAAGCTGA	CCGTTGACAA	GTCCAGATGG	CAACAGGGAA	ATGTCTTTTC	CTGTTCAGTC
2083	2093	2103	2113	2123	2133
ATGCATGAAG	CTCTCCACAA	TCACTACACA	CAGAAGTCTT	TGAGCCTGTC	CCCTGGCAAA
2136					
TGA					

Amino acid sequence of gC II-5-Fc, which was used for protein purification. Coloring as above. The TEV cleavage site is underlined.

10	20	30	131	141	151
MAPGRVGLAV	<u>VLWSLLWLGA</u>	<u>GVAGGSETAS</u>	<u>TGPFVWCDRRD</u>	PLARYGSRVQ	IRCRFRNSTR
161	171	181	191	201	211
MEFRLQIWRY	SMGPSPIAP	APDLEEVLTN	ITAPPGGLLV	YDSAPNLTD	HVLWAEAGAP
221	231	241	251	261	271
GADPPLYSVT	GPLPTQRLII	GEVTPATQGM	YYLAWGRMDS	PHEYGTWVRV	RMFRPPSLTL
281	291	301	311	321	331
QPHAVMEGQP	FKATCTAAAY	YPRNPVEFVW	FEDDRQVFNP	GQIDTQTHEH	PDGFTTVSTV
341	351	361	371	381	391
TSEAVGGQVP	PRTFTCQMTW	HRDSVMFSRR	NATGLALVLP	RPTITMEFGV	RHVCTTAGCV
401	411	421	431	441	451
PEGVTFAWFL	GDDPSPAAKS	AVTAQESCDH	PGLATVRSTL	PISYDYSEYI	CRLTGYPAGI
461	471	481	491	501	511
PVLEHHGSHQ	PPRPDPTERQ	VIEAIE <u>GTEN</u>	<u>LYFQG</u> <u>THTCP</u>	<u>PCPAPELLGG</u>	<u>PSVFLFPPKP</u>
521	531	541	551	561	571
KDTLMISRTP	EVTCVVVDVS	HEDPEVKFNW	YVDGVEVHNA	KTKPREEQYN	STYRVVSVLT
581	591	601	611	621	631
VLHQDWLNGK	EYKCKVSNKA	LPAPIEKTIS	KAKGQPREPQ	VYTLPPSREE	MTKNQVSLTC
641	651	661	671	681	691
LVKGFYPSDI	AVEWESNGQP	ENNYKTTPPV	LDSDGSEFFLY	SKLTVDKSRW	QQGNVFSCSV
701	711				
MHEALHNHYT	<u>QKSLSLSPGK</u>				

5.4 PROVIDED gC II-5-Fc CONSTRUCT BY EUROFINs MWG OPERON LLC

Vector map and multiple cloning site (MCS) of pUC57. gC II-5-Fc was cloned bluntly into the *Sma*I restriction site. Vector map provided by MWG Operon LLC. Copyright 2014 Eurofins MWG Operon LLC A Eurofins Genomics Company.



5.5 MALDI-MS ANALYSIS OF gC II-5-Fc:

Results of the MALDI-MS analysis are depicted below. The analysis was carried out by Bushra Amin and Sophie Stotz (laboratory of Dr. Hubert Kalbacher). Gray bars and red residues denote areas that have been covered by the analysis of the tryptic digest. Yellow highlights show glycosylation sites. Copyright of MALDI-MS analysis Sophie Stotz.

10	20	30	40	50	60	70	80	90	100
MAPGRVGLAV	VLWSSLWLGA	GVAGGSETAS	TGPVMCDRRD	PLARYGSRVQ	IRCRFRNSTR	MEFRLOIHRVY	SMGSPPIAP	APDLEEVLTN	ITAPPGLLV
110	120	130	140	150	160	170	180	190	200
YDSAPNLTDP	HVLWAEAGAP	GADPPLYSVT	GPLPTQRLII	GEVTPATQGM	YLLAWGRMDS	PHEYGTWVRV	RNFRPPSLTL	QPHAVMEGQP	FKATCTAAAY
210	220	230	240	250	260	270	280	290	300
YPRNPVEFVM	FEDDRQVFN	GOIDTQTHEH	PDGFTTVSTV	TSEAVGGQVP	PRTFTCQHTW	HRDSVMFSRR	NATGLALVLP	RPTITHEPGV	RHVVTAGCV
310	320	330	340	350	360	370	380	390	400
PEGVTFWFL	GDDPSPAAS	AVTAQESCDH	PGLATVRSSTL	PISYDYSEYI	CRLTGYPAGI	PVLEHHGSHQ	PPRPDPTERQ	VIEAIEGTEN	LYFQGTHTCP
410	420	430	440	450	460	470	480	490	500
PCPAPPELLGG	PSVFLFPFKP	KDTLMISRTP	EVTCCVVVDVS	HEDPEVKFNW	YVDGVEVHNA	KTKPREEQYN	STYRVVSVLT	VLHQDWLNGK	EYKCKVSNKA
510	520	530	540	550	560	570	580	590	600
LPAPIEKTIS	KARGQPREPQ	VYTLPPSREE	MTRKNQVSLTC	LVKGFYPSDI	AVEVESNGQP	ENNYKTTPPV	LSDSGSFFLY	SKLTVDKSRW	QQGNVFSCSV
610	620	630							
MHEALHNYT	OKSLSLSPGK								

6. REFERENCES

1. Svennerholm L, *Chromatographic separation of human brain gangliosides*. J Neurochem, 1963. **10**: p. 613-23.
2. Gross CP and Sepkowitz KA, *The myth of the medical breakthrough: smallpox, vaccination, and Jenner reconsidered*. Int J Infect Dis, 1998. **3**(1): p. 54-60.
3. Willis NJ, *Edward Jenner and the eradication of smallpox*. Scott Med J, 1997. **42**(4): p. 118-21.
4. Riedel S, *Edward Jenner and the history of smallpox and vaccination*. Proc (Bayl Univ Med Cent), 2005. **18**(1): p. 21-25.
5. Koch R, *Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des Bacillus anthracis*. Mitteilungen aus dem Kaiserlichen Gesundheitsamte, 1881(1): p. 174-206.
6. Sakula A, *Robert Koch: centenary of the discovery of the tubercle bacillus, 1882*. Can Vet J, 1983. **24**(4): p. 127-31.
7. Stanley WM, *Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus*. Science, 1935. **81**(2113): p. 644-45.
8. Kausche GA, Pfankuch E, and Ruska H, *Die Sichtbarmachung von pflanzlichem Virus im Übermikroskop*. Naturwissenschaften, 1939. **27**(18): p. 292-299.
9. Baltimore D, *Expression of animal virus genomes*. Bacteriol Rev, 1971. **35**(3): p. 235-41.
10. Caspar DL and Klug A, *Physical principles in the construction of regular viruses*. Cold Spring Harb Symp Quant Biol, 1962. **27**: p. 1-24.
11. Johnson JE and Speir JA, *Quasi-equivalent viruses: a paradigm for protein assemblies*. J Mol Biol, 1997. **269**(5): p. 665-75.
12. Caspar DL, Dulbecco R, Klug A, Lwoff A, Stoker MG, Tournier P, and Wildy P, *Proposals*. Cold Spring Harb Symp Quant Biol, 1962. **27**: p. 49-50.
13. Mannige RV and Brooks CL, III, *Periodic table of virus capsids: implications for natural selection and design*. PLoS One, 2010. **5**(3): p. e9423.
14. Griffith JP, Griffith DL, Rayment I, Murakami WT, and Caspar DLD, *Inside polyomavirus at 25-Å resolution*. Nature, 1992. **355**(6361): p. 652-54.
15. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, and Harrison SC, *Structure of simian virus 40 at 3.8-Å resolution*. Nature, 1991. **354**(6351): p. 278-84.
16. Yan Y, Stehle T, Liddington RC, Zhao H, and Harrison SC, *Structure determination of simian virus 40 and murine polyomavirus by a combination of 30-fold and 5-fold electron-density averaging*. Structure, 1996. **4**(2): p. 157-64.
17. Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, and Johnson JE, *Topologically linked protein rings in the bacteriophage HK97 capsid*. Science, 2000. **289**(5487): p. 2129-33.
18. Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, and Bustamante C, *The bacteriophage straight Φ 29 portal motor can package DNA against a large internal force*. Nature, 2001. **413**(6857): p. 748-52.
19. Baker ML, Jiang W, Rixon FJ, and Chiu W, *Common ancestry of herpesviruses and tailed DNA bacteriophages*. J Virol, 2005. **79**(23): p. 14967-70.

20. Lee KK, Gan L, Tsuruta H, Moyer C, Conway JF, Duda RL, Hendrix RW, Steven AC, and Johnson JE, *Virus capsid expansion driven by the capture of mobile surface loops*. *Structure*, 2008. **16**(10): p. 1491-502.
21. Mercer J, Schelhaas M, and Helenius A, *Virus entry by endocytosis*. *Annu Rev Biochem*, 2010. **79**: p. 803-33.
22. Stehle T and Harrison SC, *Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments*. *Structure*, 1996. **4**(2): p. 183-94.
23. Neu U, Woellner K, Gauglitz G, and Stehle T, *Structural basis of GM1 ganglioside recognition by simian virus 40*. *Proc Natl Acad Sci USA*, 2008. **105**(13): p. 5219-24.
24. Dermody TS, Kirchner E, Guglielmi KM, and Stehle T, *Immunoglobulin superfamily virus receptors and the evolution of adaptive immunity*. *PLoS Pathog*, 2009. **5**(11): p. e1000481.
25. Stewart PL and Nemerow GR, *Cell integrins: commonly used receptors for diverse viral pathogens*. *Trends Microbiol*, 2007. **15**(11): p. 500-07.
26. Kolokoltsov AA, Deniger D, Fleming EH, Roberts NJ, Jr., Karpilow JM, and Davey RA, *Small interfering RNA profiling reveals key role of clathrin-mediated endocytosis and early endosome formation for infection by respiratory syncytial virus*. *J Virol*, 2007. **81**(14): p. 7786-800.
27. Nicola AV, Hou J, Major EO, and Straus SE, *Herpes simplex virus type 1 enters human epidermal keratinocytes, but not neurons, via a pH-dependent endocytic pathway*. *J Virol*, 2005. **79**(12): p. 7609-16.
28. Pearse BM, *Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles*. *Proc Natl Acad Sci USA*, 1976. **73**(4): p. 1255-59.
29. Marsh M and Helenius A, *Virus entry into animal cells*. *Adv Virus Res*, 1989. **36**: p. 107-51.
30. Matlin KS, Reggio H, Helenius A, and Simons K, *Pathway of vesicular stomatitis virus entry leading to infection*. *J Mol Biol*, 1982. **156**(3): p. 609-31.
31. Saeed MF, Kolokoltsov AA, Albrecht T, and Davey RA, *Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes*. *PLoS Pathog*, 2010. **6**(9): p. e1001110.
32. Briand N, Dugail I, and Le Lay S, *Cavin proteins: new players in the caveolae field*. *Biochimie*, 2011. **93**(1): p. 71-77.
33. Lajoie P and Nabi IR, *Lipid rafts, caveolae, and their endocytosis*. *Int Rev Cell Mol Biol*, 2010. **282**: p. 135-63.
34. Parton RG and Simons K, *The multiple faces of caveolae*. *Nat Rev Mol Cell Biol*, 2007. **8**(3): p. 185-94.
35. Lajoie P and Nabi IR, *Regulation of raft-dependent endocytosis*. *J Cell Mol Med*, 2007. **11**(4): p. 644-53.
36. Hayer A, Stoeber M, Ritz D, Engel S, Meyer HH, and Helenius A, *Caveolin-1 is ubiquitinated and targeted to intraluminal vesicles in endolysosomes for degradation*. *J Cell Biol*, 2010. **191**(3): p. 615-29.
37. Gross L, *Neck tumors, or leukemia, developing in adult C3H mice following inoculation, in early infancy, with filtered (Berkefeld N), or centrifugated (144,000 X g), Ak-leukemic extracts*. *Cancer*, 1953. **6**(5): p. 948-58.

38. Eddy BE, Stewart SE, Young R, and Mider GB, *Neoplasms in hamsters induced by mouse tumor agent passed in tissue culture*. J Natl Cancer Inst, 1958. **20**(4): p. 747-61.
39. Stewart SE, Eddy BE, and Borgese N, *Neoplasms in mice inoculated with a tumor agent carried in tissue culture*. J Natl Cancer Inst, 1958. **20**(6): p. 1223-43.
40. Sweet BH and Hilleman MR, *The vacuolating virus, S.V. 40*. Proc Soc Exp Biol Med, 1960. **105**: p. 420-27.
41. Shah K and Nathanson N, *Human exposure to SV40: review and comment*. Am J Epidemiol, 1976. **103**(1): p. 1-12.
42. Gardner SD, Field AM, Coleman DV, and Hulme B, *New human papovavirus (B.K.) isolated from urine after renal transplantation*. Lancet, 1971. **1**(7712): p. 1253-57.
43. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, and Dessel BH, *Cultivation of papova-like virus from human brain with progressive multifocal leukoencephalopathy*. Lancet, 1971. **1**(7712): p. 1257-60.
44. Bogdanovic G, Priftakis P, Giraud G, Kuzniar M, Ferraldeschi R, Kokhaei P, Mellstedt H, Remberger M, Ljungman P, Winiarski J, and Dalianis T, *Association between a high BK virus load in urine samples of patients with graft-versus-host disease and development of hemorrhagic cystitis after hematopoietic stem cell transplantation*. J Clin Microbiol, 2004. **42**(11): p. 5394-96.
45. Shishido-Hara Y, *Progressive multifocal leukoencephalopathy and promyelocytic leukemia nuclear bodies: a review of clinical, neuropathological, and virological aspects of JC virus-induced demyelinating disease*. Acta Neuropathol, 2010. **120**(3): p. 403-17.
46. Allander T, Andreasson K, Gupta S, Bjerckner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, and Andersson B, *Identification of a third human polyomavirus*. J Virol, 2007. **81**(8): p. 4130-36.
47. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, Sloots TP, and Wang D, *Identification of a novel polyomavirus from patients with acute respiratory tract infections*. PLoS Pathog, 2007. **3**(5): p. e64.
48. Feng H, Shuda M, Chang Y, and Moore PS, *Clonal integration of a polyomavirus in human Merkel cell carcinoma*. Science, 2008. **319**(5866): p. 1096-100.
49. Dalianis T and Hirsch HH, *Human polyomaviruses in disease and cancer*. Virology, 2013. **437**(2): p. 63-72.
50. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, and Buck CB, *Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin*. Cell Host Microbe, 2010. **7**(6): p. 509-15.
51. van der Meijden E, Janssens RWA, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, and Feltkamp MCW, *Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient*. PLoS Pathog, 2010. **6**(7): p. e1001024.
52. Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kuhn J, Hengel H, and Ehlers B, *A novel human polyomavirus closely related to the*

- african green monkey-derived lymphotropic polyomavirus*. J Virol, 2011. **85**(9): p. 4586-90.
53. Buck CB, Phan GQ, Raiji MT, Murphy PM, McDermott DH, and McBride AA, *Complete genome sequence of a tenth human polyomavirus*. J Virol, 2012. **86**(19): p. 10887.
 54. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, Gordon JI, and Wang D, *Identification of MW polyomavirus, a novel polyomavirus in human stool*. J Virol, 2012. **86**(19): p. 10321-26.
 55. Yu G, Greninger AL, Isa P, Phan TG, Martínez MA, de la Luz Sanchez M, Contreras JF, Santos-Preciado JI, Parsonnet J, Miller S, DeRisi JL, Delwart E, Arias CF, and Chiu CY, *Discovery of a novel polyomavirus in acute diarrheal samples from children*. PLoS One, 2012. **7**(11): p. e49449.
 56. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I, Voigt S, Schmuck R, and Ehlers B, *Identification of a novel human polyomavirus in organs of the gastrointestinal tract*. PLoS One, 2013. **8**(3): p. e58021.
 57. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, Brennan DC, Mkakosya RS, Manary MJ, Gordon JI, and Wang D, *Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing*. Virology, 2013. **436**(2): p. 295-303.
 58. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese T, Faust PL, and Lipkin WI, *Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy*. J Infect Dis, 2014. **210**(10): p. 1595-59.
 59. Johne R, Buck CB, Allander T, Atwood WJ, Garcea RL, Imperiale MJ, Major EO, Ramqvist T, and Norkin LC, *Taxonomical developments in the family Polyomaviridae*. Arch Virol, 2011. **156**(9): p. 1627-34.
 60. Treisman R, Cowie A, Favaloro J, Jat P, and Kamen R, *The structures of the spliced mRNAs encoding polyoma virus early region proteins*. J Mol Appl Genet, 1981. **1**(2): p. 83-92.
 61. Jay G, Nomura S, Anderson CW, and Khoury G, *Identification of the SV40 agnogene product: a DNA binding protein*. Nature, 1981. **291**(5813): p. 346-49.
 62. Gerits N and Moens U, *Agnoprotein of mammalian polyomaviruses*. Virology, 2012. **432**(2): p. 316-26.
 63. Khalili K, White MK, Sawa H, Nagashima K, and Safak M, *The agnoprotein of polyomaviruses: a multifunctional auxiliary protein*. J Cell Physiol, 2005. **204**(1): p. 1-7.
 64. Stehle T, Yan Y, Benjamin TL, and Harrison SC, *Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment*. Nature, 1994. **369**(6476): p. 160-63.
 65. Barouch DH and Harrison SC, *Interactions among the major and minor coat proteins of polyomavirus*. J Virol, 1994. **68**(6): p. 3982-89.
 66. Chen XS, Stehle T, and Harrison SC, *Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry*. EMBO J, 1998. **17**(12): p. 3233-40.
 67. Streuli CH and Griffin BE, *Myristic acid is coupled to a structural protein of polyoma virus and SV40*. Nature, 1987. **326**(6113): p. 619-22.

68. Daniels R, Sadowicz D, and Hebert DN, *A very late viral protein triggers the lytic release of SV40*. PLoS Pathog, 2007. **3**(7): p. e98.
69. Raghava S, Giorda KM, Romano FB, Heuck AP, and Hebert DN, *The SV40 late protein VP4 is a viroporin that forms pores to disrupt membranes for viral release*. PLoS Pathog, 2011. **7**(6): p. e1002116.
70. Stehle T, Gamblin SJ, Yan Y, and Harrison SC, *The structure of simian virus 40 refined at 3.1 Å resolution*. Structure, 1996. **4**(2): p. 165-82.
71. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, and Rapoport TA, *Gangliosides are receptors for murine polyoma virus and SV40*. EMBO J, 2003. **22**(17): p. 4346-55.
72. Cahan LD, Singh R, and Paulson JC, *Sialyloligosaccharide receptors of binding variants of polyoma virus*. Virology, 1983. **130**(2): p. 281-89.
73. Fried H, Cahan LD, and Paulson JC, *Polyoma virus recognizes specific sialyloligosaccharide receptors on host cells*. Virology, 1981. **109**(1): p. 188-92.
74. Neu U, Allen SA, Blaum BS, Liu Y, Frank M, Palma AS, Ströh LJ, Feizi T, Peters T, Atwood WJ, and Stehle T, *A structure-guided mutation in the major capsid protein retargets BK polyomavirus*. PLoS Pathog, 2013. **9**(10): p. e1003688.
75. Buch MHC, Liaci AM, O'Hara SD, Garcea RL, Neu U, and Stehle T, *Structural and functional analysis of murine polyomavirus capsid proteins establish the determinants of ligand recognition and pathogenicity*. PLoS Pathog, 2015. **11**(10): p. e1005104.
76. Campanero-Rhodes MA, Smith A, Chai W, Sonnino S, Mauri L, Childs RA, Zhang Y, Ewers H, Helenius A, Imberty A, and Feizi T, *N-glycolyl GM1 ganglioside as a receptor for simian virus 40*. J Virol, 2007. **81**(23): p. 12846-58.
77. Neu U, Bauer J, and Stehle T, *Viruses and sialic acids: rules of engagement*. Curr Opin Struct Biol, 2011. **21**(5): p. 610-18.
78. Ströh LJ, Gee GV, Blaum BS, Dugan AS, Feltkamp MCW, Atwood WJ, and Stehle T, *Trichodysplasia spinulosa-associated polyomavirus uses a displaced binding site on VP1 to engage sialylated glycolipids*. PLoS Pathog, 2015. **11**(8): p. e1005112.
79. Neu U, Stehle T, and Atwood WJ, *The Polyomaviridae: contributions of virus structure to our understanding of virus receptors and infectious entry*. Virology, 2009. **384**(2): p. 389-99.
80. Neu U, Maginnis MS, Palma AS, Ströh LJ, Nelson CDS, Feizi T, Atwood WJ, and Stehle T, *Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif*. Cell Host Microbe, 2010. **8**(4): p. 309-19.
81. Dörries K, *Molecular biology and pathogenesis of human polyomavirus infections*. Dev Biol Stand, 1998. **94**: p. 71-79.
82. Silverman L and Rubinstein LJ, *Electron microscopic observations on a case of progressive multifocal leukoencephalopathy*. Acta Neuropathol, 1965. **5**(2): p. 215-24.
83. ZuRhein G and Chou SM, *Particles resembling papova viruses in human cerebral demyelinating disease*. Science, 1965. **148**(3676): p. 1477-79.
84. Maginnis MS, Ströh LJ, Gee GV, O'Hara BA, Derdowski A, Stehle T, and Atwood WJ, *Progressive multifocal leukoencephalopathy-associated*

- mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding.* MBio, 2013. **4**(3): p. e00247-13.
85. Haley SA and Atwood WJ, *An animal model for progressive multifocal leukoencephalopathy.* J Clin Invest, 2014. **124**(12): p. 5103-06.
 86. Ray U, Cinque P, Gerevini S, Longo V, Lazzarin A, Schippiling S, Martin R, Buck CB, and Pastrana DV, *JC polyomavirus mutants escape antibody-mediated neutralization.* Sci Transl Med, 2015. **7**(306): p. 306ra151.
 87. Jelcic I, Combaluzier B, Jelcic I, Faigle W, Senn L, Reinhart BJ, Ströh L, Nitsch RM, Stehle T, Sospedra M, Grimm J, and Martin R, *Broadly neutralizing human monoclonal JC polyomavirus VP1-specific antibodies as candidate therapeutics for progressive multifocal leukoencephalopathy.* Sci Transl Med, 2015. **7**(306): p. 306ra150.
 88. Dawe CJ, Freund R, Mandel G, Ballmer-Hofer K, Talmage DA, and Benjamin TL, *Variations in polyoma virus genotype in relation to tumor induction in mice. Characterization of wild type strains with widely differing tumor profiles.* Am J Pathol, 1987. **127**(2): p. 243-61.
 89. Feunteun J, Sompayrac L, Fluck M, and Benjamin TL, *Localization of gene functions in polyoma virus DNA.* Proc Natl Acad Sci USA, 1976. **73**(11): p. 4169-73.
 90. Dawe CJ, Law LW, and Dunn TB, *Studies of parotid-tumor agent in cultures of leukemic tissues of mice.* J Natl Cancer Inst, 1959. **23**: p. 717-97.
 91. Gross L, *"Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK-leukemic extracts, or AK-embryos.* Proc Soc Exp Biol Med, 1951. **76**(1): p. 27-32.
 92. Carroll J, Dey D, Kreisman L, Velupillai P, Dahl J, Telford S, Bronson R, and Benjamin TL, *Receptor-binding and oncogenic properties of polyoma viruses isolated from feral mice.* PLoS Pathog, 2007. **3**(12): p. e179.
 93. Main JH and Dawe CJ, *Tumor induction in transplanted tooth buds infected with polyoma virus.* J Natl Cancer Inst, 1966. **36**(6): p. 1121-36.
 94. Rowe WP, Hartley JW, Estes JD, and Huebner RJ, *Studies of mouse polyoma virus infection. I. Procedures for quantitation and detection of virus.* J Exp Med, 1959. **109**(4): p. 379-91.
 95. Bolen JB, Fisher SE, Chowdhury K, Shan TC, Williams JE, Dawe CJ, and Israel MA, *A determinant of polyomavirus virulence enhances virus growth in cells of renal origin.* J Virol, 1985. **53**(1): p. 335-39.
 96. Bauer PH, Bronson RT, Fung SC, Freund R, Stehle T, Harrison SC, and Benjamin TL, *Genetic and structural analysis of a virulence determinant in polyomavirus VP1.* J Virol, 1995. **69**(12): p. 7925-31.
 97. Bauer PH, Cui C, Liu WR, Stehle T, Harrison SC, DeCaprio JA, and Benjamin TL, *Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse.* J Virol, 1999. **73**(7): p. 5826-32.
 98. Freund R, Calderone A, Dawe CJ, and Benjamin TL, *Polyomavirus tumor induction in mice: effects of polymorphisms of VP1 and large T antigen.* J Virol, 1991. **65**(1): p. 335-41.
 99. Freund R, Garcea RL, Sahli R, and Benjamin TL, *A single-amino-acid substitution in polyomavirus VP1 correlates with plaque size and hemagglutination behavior.* Journal of virology, 1991. **65**(1): p. 350-55.

100. Gilbert JM and Benjamin TL, *Uptake pathway of polyomavirus via ganglioside GD1a*. J Virol, 2004. **78**(22): p. 12259-67.
101. Gilbert JM and Benjamin TL, *Early steps of polyomavirus entry into cells*. J Virol, 2000. **74**(18): p. 8582-88.
102. Damm EM, Pelkmans L, Kartenbeck J, Mezzacasa A, Kurzchalia T, and Helenius A, *Clathrin- and caveolin-1-independent endocytosis: entry of simian virus 40 into cells devoid of caveolae*. J Cell Biol, 2005. **168**(3): p. 477-88.
103. Richterová Z, Liebl D, Horák M, Palková Z, Stokrová J, Hozák P, Korb J, and Forstová J, *Caveolae are involved in the trafficking of mouse polyomavirus virions and artificial VP1 pseudocapsids toward cell nuclei*. J Virol, 2001. **75**(22): p. 10880-91.
104. Engel S, Heger T, Mancini R, Herzog F, Kartenbeck J, Hayer A, and Helenius A, *Role of endosomes in simian virus 40 entry and infection*. J Virol, 2011. **85**(9): p. 4198-211.
105. Pelkmans L, Kartenbeck J, and Helenius A, *Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the ER*. Nat Cell Biol, 2001. **3**(5): p. 473-83.
106. Eash S, Querbes W, and Atwood WJ, *Infection of vero cells by BK virus is dependent on caveolae*. J Virol, 2004. **78**(21): p. 11583-90.
107. Moriyama T, Marquez JP, Wakatsuki T, and Sorokin A, *Caveolar endocytosis is critical for BK virus infection of human renal proximal tubular epithelial cells*. J Virol, 2007. **81**(16): p. 8552-62.
108. Jiang M, Abend JR, Tsai B, and Imperiale MJ, *Early events during BK virus entry and disassembly*. J Virol, 2009. **83**(3): p. 1350-58.
109. Ashok A and Atwood WJ, *Contrasting roles of endosomal pH and the cytoskeleton in infection of human glial cells by JC virus and simian virus 40*. J Virol, 2003. **77**(2): p. 1347-56.
110. Elphick GF, Querbes W, Jordan JA, Gee GV, Eash S, Manley K, Dugan A, Stanifer M, Bhatnagar A, Kroeze WK, Roth BL, and Atwood WJ, *The human polyomavirus, JCV, uses serotonin receptors to infect cells*. Science, 2004. **306**(5700): p. 1380-83.
111. Pho MT, Ashok A, and Atwood WJ, *JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis*. J Virol, 2000. **74**(5): p. 2288-92.
112. Magnuson B, Rainey EK, Benjamin TL, Baryshev M, Mkrtchian S, and Tsai B, *ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding*. Mol Cell, 2005. **20**(2): p. 289-300.
113. Mkrtchian S, Baryshev M, Matvijenko O, Sharipo A, Sandalova T, Schneider G, and Ingelman-Sundberg M, *Oligomerization properties of ERp29, an endoplasmic reticulum stress protein*. FEBS Lett, 1998. **431**(3): p. 322-26.
114. Schelhaas M, Malmstrom J, Pelkmans L, Haugstetter J, Ellgaard L, Grunewald K, and Helenius A, *Simian virus 40 depends on ER protein folding and quality control factors for entry into host cells*. Cell, 2007. **131**(3): p. 516-29.
115. Bennett SM, Jiang M, and Imperiale MJ, *Role of cell-type-specific endoplasmic reticulum-associated degradation in polyomavirus trafficking*. J Virol, 2013. **87**(16): p. 8843-52.

116. Geiger R, Andritschke D, Friebe S, Herzog F, Luisoni S, Heger T, and Helenius A, *BAP31 and BiP are essential for dislocation of SV40 from the endoplasmic reticulum to the cytosol*. Nat Cell Biol, 2011. **13**(11): p. 1305-14.
117. Goodwin EC, Lipovsky A, Inoue T, Magaldi TG, Edwards APB, Van Goor KEY, Paton AW, Paton JC, Atwood WJ, Tsai B, and DiMaio D, *BiP and multiple DNAJ molecular chaperones in the endoplasmic reticulum are required for efficient simian virus 40 infection*. MBio, 2011. **2**(3): p. e00101-11.
118. Lilley BN, Gilbert JM, Ploegh HL, and Benjamin TL, *Murine polyomavirus requires the endoplasmic reticulum protein Derlin-2 to initiate infection*. J Virol, 2006. **80**(17): p. 8739-44.
119. Davison AJ, Eberle R, Ehlers B, Hayward GS, McGeoch DJ, Minson AC, Pellett PE, Roizman B, Studdert MJ, and Thiry E, *The order Herpesvirales*. Arch Virol, 2009. **154**(1): p. 171-77.
120. Whitley RJ and Roizman B, *Herpes simplex virus infections*. Lancet, 2001. **357**(9267): p. 1513-18.
121. Roizman B and Furlong D, *The replication of herpesviruses*, in *Reproduction*, H. Fraenkel-Conrat and R. Wagner, Editors. 1974, Springer US. p. 229-403.
122. Fields BN, Knipe DM, and Howley PM, *Fields virology 2013*, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
123. Antoine TE, Park PJ, and Shukla D, *Glycoprotein targeted therapeutics: a new era of anti-herpes simplex virus-1 therapeutics*. Rev Med Virol, 2013. **23**(3): p. 194-208.
124. Eisenberg RJ, Atanasiu D, Cairns TM, Gallagher JR, Krummenacher C, and Cohen GH, *Herpes virus fusion and entry: a story with many characters*. Viruses, 2012. **4**(5): p. 800-32.
125. Grünewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, and Steven AC, *Three-dimensional structure of herpes simplex virus from cryo-electron tomography*. Science, 2003. **302**(5649): p. 1396-98.
126. Nahmias AJ and Dowdle WR, *Antigenic and biologic differences in herpesvirus hominis*. Prog Med Virol, 1968. **10**: p. 110-59.
127. Daniels CA and LeGoff SG, *Shedding of infectious virus/antibody complexes from vesicular lesions of patients with recurrent herpes labialis*. Lancet, 1975. **2**(7934): p. 524-28.
128. Bloom DC, Giordani NV, and Kwiatkowski DL, *Epigenetic regulation of latent HSV-1 gene expression*. Biochim Biophys Acta, 2010. **1799**(3-4): p. 246-56.
129. Whitley RJ, *Herpes simplex encephalitis: adolescents and adults*. Antiviral Res, 2006. **71**(2-3): p. 141-48.
130. Smith MG, Lennette EH, and Reames HR, *Isolation of the virus of herpes simplex and the demonstration of intranuclear inclusions in a case of acute encephalitis*. Am J Pathol, 1941. **17**(1): p. 55-68.1.
131. Zarafonitis CJ and Smadel JE, *Fatal herpes simplex encephalitis in man*. Am J Pathol, 1944. **20**(3): p. 429-45.
132. Gilden DH, Mahalingam R, Cohrs RJ, and Tyler KL, *Herpesvirus infections of the nervous system*. Nat Clin Pract Neurol, 2007. **3**(2): p. 82-94.

133. Campadelli-Fiume G, Amasio M, Avitabile E, Cerretani A, Forghieri C, Gianni T, and Menotti L, *The multipartite system that mediates entry of herpes simplex virus into the cell*. Rev Med Virol, 2007. **17**(5): p. 313-26.
134. Heldwein EE and Krummenacher C, *Entry of herpesviruses into mammalian cells*. Cell Mol Life Sci, 2008. **65**(11): p. 1653-68.
135. Šedý JR, Spear PG, and Ware CF, *Cross-regulation between herpesviruses and the TNF superfamily members*. Nat Rev Immunol, 2008. **8**(11): p. 861-73.
136. Spear PG, *Herpes simplex virus: receptors and ligands for cell entry*. Cell Microbiol, 2004. **6**(5): p. 401-10.
137. WuDunn D and Spear PG, *Initial interaction of herpes simplex virus with cells is binding to heparan sulfate*. J Virol, 1989. **63**(1): p. 52-8.
138. MacLeod DT, Nakatsuji T, Yamasaki K, Kobzik L, and Gallo RL, *HSV-1 exploits the innate immune scavenger receptor MARCO to enhance epithelial adsorption and infection*. Nat Commun, 2013. **4**: p. 1963.
139. Herold BC, WuDunn D, Soltys N, and Spear PG, *Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity*. J Virol, 1991. **65**(3): p. 1090-98.
140. Lycke E, Johansson M, Svennerholm B, and Lindahl U, *Binding of herpes simplex virus to cellular heparan sulphate, an initial step in the adsorption process*. J Gen Virol, 1991. **72** (Pt 5): p. 1131-37.
141. Banfield BW, Leduc Y, Esford L, Visalli RJ, Brandt CR, and Tufaro F, *Evidence for an interaction of herpes simplex virus with chondroitin sulfate proteoglycans during infection*. Virology, 1995. **208**(2): p. 531-39.
142. Mårdberg K, Trybala E, Tufaro F, and Bergström T, *Herpes simplex virus type 1 glycoprotein C is necessary for efficient infection of chondroitin sulfate-expressing gro2C cells*. J Gen Virol, 2002. **83**(Pt 2): p. 291-300.
143. Friedman HM, Wang L, Fishman NO, Lambris JD, Eisenberg RJ, Cohen GH, and Lubinski J, *Immune evasion properties of herpes simplex virus type 1 glycoprotein gC*. J Virol, 1996. **70**(7): p. 4253-60.
144. Oh MJ, Akhtar J, Desai P, and Shukla D, *A role for heparan sulfate in viral surfing*. Biochem Biophys Res Commun, 2010. **391**(1): p. 176-81.
145. Ligas MW and Johnson DC, *A herpes simplex virus mutant in which glycoprotein D sequences are replaced by beta-galactosidase sequences binds to but is unable to penetrate into cells*. J Virol, 1988. **62**(5): p. 1486-94.
146. Montgomery RI, Warner MS, Lum BJ, and Spear PG, *Herpes simplex virus-1 entry into cells mediated by a novel member of the TNF/NGF receptor family*. Cell, 1996. **87**(3): p. 427-36.
147. Warner MS, Geraghty RJ, Martinez WM, Montgomery RI, Whitbeck JC, Xu R, Eisenberg RJ, Cohen GH, and Spear PG, *A cell surface protein with herpesvirus entry activity (HveB) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus*. Virology, 1998. **246**(1): p. 179-89.
148. Eberlé F, Dubreuil P, Mattei MG, Devilard E, and Lopez M, *The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is the true homolog of the murine MPH gene*. Gene, 1995. **159**(2): p. 267-72.
149. Lopez M, Eberlé F, Mattei MG, Gabert J, Birg F, Bardin F, Maroc C, and Dubreuil P, *Complementary DNA characterization and chromosomal*

- localization of a human gene related to the poliovirus receptor-encoding gene.* Gene, 1995. **155**(2): p. 261-65.
150. Shukla D, Liu J, Blaiklock P, Shworak NW, Bai X, Esko JD, Cohen GH, Eisenberg RJ, Rosenberg RD, and Spear PG, *A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry.* Cell, 1999. **99**(1): p. 13-22.
 151. Perez A, Li QX, Perez-Romero P, Delassus G, Lopez SR, Sutter S, McLaren N, and Fuller AO, *A new class of receptor for herpes simplex virus has heptad repeat motifs that are common to membrane fusion proteins.* J Virol, 2005. **79**(12): p. 7419-30.
 152. Suenaga T, Satoh T, Somboonthum P, Kawaguchi Y, Mori Y, and Arase H, *Myelin-associated glycoprotein mediates membrane fusion and entry of neurotropic herpesviruses.* Proc Natl Acad Sci USA, 2010. **107**(2): p. 866-71.
 153. Arii J, Goto H, Suenaga T, Oyama M, Kozuka-Hata H, Imai T, Minowa A, Akashi H, Arase H, Kawaoka Y, and Kawaguchi Y, *Non-muscle myosin IIA is a functional entry receptor for herpes simplex virus-1.* Nature, 2010. **467**(7317): p. 859-62.
 154. Arii J, Hirohata Y, Kato A, and Kawaguchi Y, *Nonmuscle myosin heavy chain IIb mediates herpes simplex virus 1 entry.* J Virol, 2015. **89**(3): p. 1879-88.
 155. Chowdhury S, Chouljenko VN, Naderi M, and Kousoulas KG, *The amino terminus of herpes simplex virus 1 glycoprotein K is required for virion entry via the paired immunoglobulin-like type-2 receptor alpha.* J Virol, 2013. **87**(6): p. 3305-13.
 156. Satoh T, Arii J, Suenaga T, Wang J, Kogure A, Uehori J, Arase N, Shiratori I, Tanaka S, Kawaguchi Y, Spear PG, Lanier LL, and Arase H, *PILR α is a herpes simplex virus-1 entry coreceptor that associates with glycoprotein B.* Cell, 2008. **132**(6): p. 935-44.
 157. Milne RS, Nicola AV, Whitbeck JC, Eisenberg RJ, and Cohen GH, *Glycoprotein D receptor-dependent, low-pH-independent endocytic entry of herpes simplex virus type 1.* J Virol, 2005. **79**(11): p. 6655-63.
 158. Heldwein EE, Lou H, Bender FC, Cohen GH, Eisenberg RJ, and Harrison SC, *Crystal structure of glycoprotein B from herpes simplex virus 1.* Science, 2006. **313**(5784): p. 217-20.
 159. Roche S, Bressanelli S, Rey FA, and Gaudin Y, *Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G.* Science, 2006. **313**(5784): p. 187-91.
 160. Svennerholm B, Jeansson S, Vahlne A, and Lycke E, *Involvement of glycoprotein C (gC) in adsorption of herpes simplex virus type 1 (HSV-1) to the cell.* Arch Virol, 1991. **120**(3-4): p. 273-79.
 161. Olofsson S, *Carbohydrates in herpesvirus infections.* APMIS Suppl, 1992. **27**: p. 84-95.
 162. Hansen JE, Lund O, Engelbrecht J, Bohr H, Nielsen JO, and Hansen JE, *Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase.* Biochem J, 1995. **308** (Pt 3): p. 801-13.
 163. Ekblad M, Adamiak B, Bergefall K, Nenonen H, Roth A, Bergström T, Ferro V, and Trybala E, *Molecular basis for resistance of herpes simplex virus type 1 mutants to the sulfated oligosaccharide inhibitor PI-88.* Virology, 2007. **367**(2): p. 244-52.

164. Mårdberg K, Nyström K, Tarp MA, Trybala E, Clausen H, Bergström T, and Olofsson S, *Basic amino acids as modulators of an O-linked glycosylation signal of the herpes simplex virus type 1 glycoprotein gC: functional roles in viral infectivity*. *Glycobiology*, 2004. **14**(7): p. 571-81.
165. Altgårde N, Eriksson C, Peerboom N, Phan-Xuan T, Moeller S, Schnabelrauch M, Svedhem S, Trybala E, Bergström T, and Bally M, *Mucin-like region of herpes simplex virus type 1 attachment protein glycoprotein C (gC) modulates the virus-glycosaminoglycan interaction*. *J Biol Chem*, 2015. **290**(35): p. 21473-85.
166. Marlin SD, Holland TC, Levine M, and Glorioso JC, *Epitopes of herpes simplex virus type 1 glycoprotein gC are clustered in two distinct antigenic sites*. *J Virol*, 1985. **53**(1): p. 128-36.
167. Wu CT, Levine M, Homa F, Highlander SL, and Glorioso JC, *Characterization of the antigenic structure of herpes simplex virus type 1 glycoprotein C through DNA sequence analysis of monoclonal antibody-resistant mutants*. *J Virol*, 1990. **64**(2): p. 856-63.
168. Mårdberg K, Trybala E, Glorioso JC, and Bergström T, *Mutational analysis of the major heparan sulfate-binding domain of herpes simplex virus type 1 glycoprotein C*. *J Gen Virol*, 2001. **82**(Pt 8): p. 1941-50.
169. Trybala E, Bergström T, Svennerholm B, Jeansson S, Glorioso JC, and Olofsson S, *Localization of a functional site on herpes simplex virus type 1 glycoprotein C involved in binding to cell surface heparan sulphate*. *J Gen Virol*, 1994. **75** (Pt 4): p. 743-52.
170. Rux AH, Moore WT, Lambris JD, Abrams WR, Peng C, Friedman HM, Cohen GH, and Eisenberg RJ, *Disulfide bond structure determination and biochemical analysis of glycoprotein C from herpes simplex virus*. *J Virol*, 1996. **70**(8): p. 5455-65.
171. Hung SL, Srinivasan S, Friedman HM, Eisenberg RJ, and Cohen GH, *Structural basis of C3b binding by glycoprotein C of herpes simplex virus*. *J Virol*, 1992. **66**(7): p. 4013-27.
172. Hung SL, Peng C, Kostavasili I, Friedman HM, Lambris JD, Eisenberg RJ, and Cohen GH, *The interaction of glycoprotein C of herpes simplex virus types 1 and 2 with the alternative complement pathway*. *Virology*, 1994. **203**(2): p. 299-312.
173. Liljeqvist JA, Svennerholm B, and Bergström T, *Typing of clinical herpes simplex virus type 1 and type 2 isolates with monoclonal antibodies*. *J Clin Microbiol*, 1999. **37**(8): p. 2717-18.
174. Adamiak B, Trybala E, Mårdberg K, Johansson M, Liljeqvist JA, Olofsson S, Grabowska A, Bienkowska-Szewczyk K, Szewczyk B, and Bergström T, *Human antibodies to herpes simplex virus type 1 glycoprotein C are neutralizing and target the heparan sulfate-binding domain*. *Virology*, 2010. **400**(2): p. 197-206.
175. Olofsson S, Bolmstedt A, Biller M, Mårdberg K, Leckner J, Malmström BG, Trybala E, and Bergström T, *The role of a single N-linked glycosylation site for a functional epitope of herpes simplex virus type 1 envelope glycoprotein gC*. *Glycobiology*, 1999. **9**(1): p. 73-81.
176. Varki A and Lowe JB, *Biological roles of glycans*, in *Essentials of glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley,

- C.R. Bertozzi, G.W. Hart, and M.E. Etzler, Editors. 2009: Cold Spring Harbor (NY).
177. Barry JD and McCulloch R, *Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite*. *Adv Parasitol*, 2001. **49**: p. 1-70.
 178. de Virgilio M, Kitzmuller C, Schwaiger E, Klein M, Kreibich G, and Ivessa NE, *Degradation of a short-lived glycoprotein from the lumen of the endoplasmic reticulum: the role of N-linked glycans and the unfolded protein response*. *Mol Biol Cell*, 1999. **10**(12): p. 4059-73.
 179. Shental-Bechor D and Levy Y, *Folding of glycoproteins: toward understanding the biophysics of the glycosylation code*. *Curr Opin Struct Biol*, 2009. **19**(5): p. 524-33.
 180. Feizi T, *Oligosaccharides that mediate mammalian cell-cell adhesion*. *Curr Opin Struct Biol*, 1993. **3**(5): p. 701-710.
 181. Rosen SD and Bertozzi CR, *The selectins and their ligands*. *Curr Opin Cell Biol*, 1994. **6**(5): p. 663-73.
 182. Connor RJ, Kawaoka Y, Webster RG, and Paulson JC, *Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates*. *Virology*, 1994. **205**(1): p. 17-23.
 183. Rogers GN and D'Souza BL, *Receptor binding properties of human and animal H1 influenza virus isolates*. *Virology*, 1989. **173**(1): p. 317-22.
 184. Rogers GN and Paulson JC, *Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin*. *Virology*, 1983. **127**(2): p. 361-73.
 185. Yu RK, Nakatani Y, and Yanagisawa M, *The role of glycosphingolipid metabolism in the developing brain*. *J Lipid Res*, 2009. **50 Suppl**: p. S440-45.
 186. Yu RK, Tsai YT, Ariga T, and Yanagisawa M, *Structures, biosynthesis, and functions of gangliosides-an overview*. *J Oleo Sci*, 2011. **60**(10): p. 537-44.
 187. Schnaar RL, Suzuki A, and Stanley P, *Glycosphingolipids*, in *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, Editors. 2009: Cold Spring Harbor (NY).
 188. Hakomori S, Handa K, Iwabuchi K, Yamamura S, and Prinetti A, *New insights in glycosphingolipid function: "glycosignaling domain," a cell surface assembly of glycosphingolipids with signal transducer molecules, involved in cell adhesion coupled with signaling*. *Glycobiology*, 1998. **8**(10): p. xi-xix.
 189. Anderson RGW, *The caveolae membrane system*. *Annu Rev Biochem*, 1998. **67**: p. 199-225.
 190. Simons K and Toomre D, *Lipid rafts and signal transduction*. *Nat Rev Mol Cell Biol*, 2000. **1**(1): p. 31-39.
 191. Ledeen RW and Wu G, *Nuclear sphingolipids: metabolism and signaling*. *J Lipid Res*, 2008. **49**(6): p. 1176-86.
 192. Ströh LJ and Stehle T, *Glycan engagement by viruses: receptor switches and specificity*. *Annu Rev of Virol*, 2014. **1**(1): p. 285-306.

193. Varki A and Schauer R, *Sialic acids*, in *Essentials of glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, Editors. 2009: Cold Spring Harbor (NY).
194. Angata T and Varki A, *Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective*. Chem Rev, 2002. **102**(2): p. 439-69.
195. Nie H, Li Y, and Sun XL, *Recent advances in sialic acid-focused glycomics*. J Proteomics, 2012. **75**(11): p. 3098-112.
196. Miyata S, Sato C, Kitamura S, Toriyama M, and Kitajima K, *A major flagellum sialoglycoprotein in sea urchin sperm contains a novel polysialic acid, an α 2,9-linked poly-N-acetylneuraminic acid chain, capped by an 8-O-sulfated sialic acid residue*. Glycobiology, 2004. **14**(9): p. 827-40.
197. Fukuda M, *Possible roles of tumor-associated carbohydrate antigens*. Cancer Res, 1996. **56**(10): p. 2237-44.
198. Hildebrandt H, Becker C, Glüer S, Rösner H, Gerardy-Schahn R, and Rahmann H, *Polysialic acid on the neural cell adhesion molecule correlates with expression of polysialyltransferases and promotes neuroblastoma cell growth*. Cancer Res, 1998. **58**(4): p. 779-84.
199. Helander A, Silvey KJ, Mantis NJ, Hutchings AB, Chandran K, Lucas WT, Nibert ML, and Neutra MR, *The viral σ 1 protein and glycoconjugates containing α 2-3-linked sialic acid are involved in type 1 reovirus adherence to M cell apical surfaces*. J Virol, 2003. **77**(14): p. 7964-77.
200. Schultze B and Herrler G, *Bovine coronavirus uses N-acetyl-9-O-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells*. J Gen Virol, 1992. **73** (Pt 4): p. 901-06.
201. Han L, Tan M, Xia M, Kitova EN, Jiang X, and Klassen JS, *Gangliosides are ligands for human noroviruses*. J Am Chem Soc, 2014. **136**(36): p. 12631-37.
202. Suzuki Y, Suzuki T, Matsunaga M, and Matsumoto M, *Gangliosides as paramyxovirus receptor. Structural requirement of sialo-oligosaccharides in receptors for hemagglutinating virus of Japan (Sendai virus) and Newcastle disease virus*. J Biochem, 1985. **97**(4): p. 1189-99.
203. Matrosovich M, Herrler G, and Klenk HD, *Sialic Acid Receptors of Viruses*. Top Curr Chem, 2013.
204. Arnberg N, Kidd AH, Edlund K, Nilsson J, Pring-Akerblom P, and Wadell G, *Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction*. Virology, 2002. **302**(1): p. 33-43.
205. Arnberg N, Pring-Akerblom P, and Wadell G, *Adenovirus type 37 uses sialic acid as a cellular receptor on Chang C cells*. J Virol, 2002. **76**(17): p. 8834-41.
206. Imberty A, Lortat-Jacob H, and Pérez S, *Structural view of glycosaminoglycan-protein interactions*. Carbohydr Res, 2007. **342**(3-4): p. 430-39.
207. Esko JD, Kimata K, and Lindahl U, *Proteoglycans and sulfated glycosaminoglycans*, in *Essentials of glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, Editors. 2009: Cold Spring Harbor (NY).
208. Rabenstein DL, *Heparin and heparan sulfate: structure and function*. Nat Prod Rep, 2002. **19**(3): p. 312-31.

209. Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, and Carrell RW, *The anticoagulant activation of antithrombin by heparin*. Proc Natl Acad Sci USA, 1997. **94**(26): p. 14683-88.
210. Funderburgh JL, *Keratan sulfate: structure, biosynthesis, and function*. Glycobiology, 2000. **10**(10): p. 951-58.
211. Trowbridge JM and Gallo RL, *Dermatan sulfate: new functions from an old glycosaminoglycan*. Glycobiology, 2002. **12**(9): p. 117R-25R.
212. Sugahara K, Mikami T, Uyama T, Mizuguchi S, Nomura K, and Kitagawa H, *Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate*. Curr Opin Struct Biol, 2003. **13**(5): p. 612-20.
213. Laurent TC, Laurent UBG, and Fraser JRE, *The structure and function of hyaluronan: an overview*. Immunol Cell Biol, 1996. **74**(2): p. A1-7.
214. Fry EE, Lea SM, Jackson T, Newman JWI, Ellard FM, Blakemore WE, Abu-Ghazaleh R, Samuel A, King AMQ, and Stuart DI, *The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex*. EMBO J, 1999. **18**(3): p. 543-54.
215. Dehecchi MC, Tamanini A, Bonizzato A, and Cabrini G, *Heparan sulfate glycosaminoglycans are involved in adenovirus type 5 and 2-host cell interactions*. Virology, 2000. **268**(2): p. 382-90.
216. Tuve S, Wang H, Jacobs JD, Yumul RC, Smith DF, and Lieber A, *Role of cellular heparan sulfate proteoglycans in infection of human adenovirus serotype 3 and 35*. PLoS Pathog, 2008. **4**(10): p. e1000189.
217. Giroglou T, Florin L, Schäfer F, Streeck RE, and Sapp M, *Human papillomavirus infection requires cell surface heparan sulfate*. J Virol, 2001. **75**(3): p. 1565-70.
218. Schowalter RM, Pastrana DV, and Buck CB, *Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry*. PLoS Pathog, 2011. **7**(7): p. e1002161.
219. Salvador B, Sexton NR, Carrion R, Jr., Nunneley J, Patterson JL, Steffen I, Lu K, Muench MO, Lembo D, and Simmons G, *Filoviruses utilize glycosaminoglycans for their attachment to target cells*. J Virol, 2013. **87**(6): p. 3295-304.
220. Nokhbeh MR, Hazra S, Alexander DA, Khan A, McAllister M, Suuronen EJ, Griffith M, and Dimock K, *Enterovirus 70 binds to different glycoconjugates containing α 2,3-linked sialic acid on different cell lines*. J Virol, 2005. **79**(11): p. 7087-94.
221. Nilsson EC, Jamshidi F, Johansson SMC, Oberste MS, and Arnberg N, *Sialic acid is a cellular receptor for coxsackievirus A24 variant, an emerging virus with pandemic potential*. J Virol, 2008. **82**(6): p. 3061-68.
222. Zocher G, Mistry N, Frank M, Hähnlein-Schick I, Ekström JO, Arnberg N, and Stehle T, *A sialic acid binding site in a human picornavirus*. PLoS Pathog, 2014. **10**(10): p. e1004401.
223. You J, O'Hara SD, Velupillai P, Castle S, Levery S, Garcea RL, and Benjamin TL, *Ganglioside and non-ganglioside mediated host responses to the mouse polyomavirus*. PLoS Pathog, 2015. **11**(10): p. e1005175.
224. Stehle T and Harrison SC, *High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding*. EMBO J, 1997. **16**(16): p. 5139-48.

225. Caruso M, Belloni L, Sthandier O, Amati P, and Garcia MI, *$\alpha 4\beta 1$ integrin acts as a cell receptor for murine polyomavirus at the postattachment level.* J Virol, 2003. **77**(7): p. 3913-21.
226. Burmeister WP, Guilligay D, Cusack S, Wadell G, and Arnberg N, *Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites.* J Virol, 2004. **78**(14): p. 7727-36.
227. Ströh LJ, Maginnis MS, Blaum BS, Nelson CDS, Neu U, Gee GV, O'Hara BA, Motamedi N, DiMaio D, Atwood WJ, and Stehle T, *The greater affinity of JC polyomavirus capsid for $\alpha 2,6$ -linked lactoseries tetrasaccharide c than for other sialylated glycans is a major determinant of infectivity.* J Virol, 2015. **89**(12): p. 6364-75.
228. Koromyslova AD, Leuthold MM, Bowler MW, and Hansman GS, *The sweet quartet: binding of fucose to the norovirus capsid.* Virology, 2015. **483**: p. 203-08.
229. Neu U, Wang J, Macejak D, Garcea RL, and Stehle T, *Structures of the major capsid proteins of the human Karolinska Institutet and Washington University polyomaviruses.* J Virol, 2011. **85**(14): p. 7384-92.
230. Bialasiewicz S, Whiley DM, Lambert SB, Jacob K, Bletchly C, Wang D, Nissen MD, and Sloots TP, *Presence of the newly discovered human polyomaviruses KI and WU in Australian patients with acute respiratory tract infection.* J Clin Virol, 2008. **41**(2): p. 63-68.
231. Debiaggi M, Canducci F, Brerra R, Sampaolo M, Marinozzi MC, Parea M, Arghittu M, Alessandrino EP, Nava S, Nucleo E, Romero E, and Clementi M, *Molecular epidemiology of KI and WU polyomaviruses in infants with acute respiratory disease and in adult hematopoietic stem cell transplant recipients.* J Med Virol, 2010. **82**(1): p. 153-56.
232. Rennspiess D, Pujari S, Keijzers M, Abdul-Hamid MA, Hochstenbag M, Dingemans AM, Kurz AK, Speel EJ, Haugg A, Pastrana DV, Buck CB, De Baets MH, and Zur Hausen A, *Detection of human polyomavirus 7 in human thymic epithelial tumors.* J Thorac Oncol, 2015. **10**(2): p. 360-6.
233. Neu U, Hengel H, Blaum BS, Schowalter RM, Macejak D, Gilbert M, Wakarchuk WW, Imamura A, Ando H, Kiso M, Arnberg N, Garcea RL, Peters T, Buck CB, and Stehle T, *Structures of Merkel cell polyomavirus VP1 complexes define a sialic acid binding site required for infection.* PLoS Pathog, 2012. **8**(7): p. e1002738.
234. Air GM and Laver WG, *The neuraminidase of influenza virus.* Proteins, 1989. **6**(4): p. 341-56.
235. Hadigal SR, Agelidis AM, Karasneh GA, Antoine TE, Yakoub AM, Ramani VC, Djalilian AR, Sanderson RD, and Shukla D, *Heparanase is a host enzyme required for herpes simplex virus-1 release from cells.* Nat Commun, 2015. **6**: p. 6985.
236. Müller S, *Structural and functional characterization of the protein-protein interaction between the HCMV immunoevasin UL16 and the NKG2D ligand MICB,* in Faculty of Chemistry and Pharmacy 2010, University of Tübingen.
237. Buch MHC, *Structural and functional analysis of the major capsid protein of murine polyomavirus, VP1,* in Faculty of Science 2011, University of Tübingen.
238. Huang XQ and Miller W, *A time-efficient, linear-space local similarity algorithm.* Adv Appl Math, 1991. **12**(3): p. 337-357.

239. Vincze T, Posfai J, and Roberts RJ, *NEBcutter: A program to cleave DNA with restriction enzymes*. *Nucleic Acids Res*, 2003. **31**(13): p. 3688-91.
240. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, and Bairoch A, *Protein identification and analysis tools on the ExPASy Server*, in *The proteomics protocols handbook*, J.M. Walker, Editor 2005, Humana Press. p. 571-607.
241. Stothard P, *The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences*. *Biotechniques*, 2000. **28**(6): p. 1102, 1104.
242. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, and Erlich H, *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction*. *Cold Spring Harb Symp Quant Biol*, 1986. **51 Pt 1**: p. 263-73.
243. Rychlik W, Spencer WJ, and Rhoads RE, *Optimization of the annealing temperature for DNA amplification in vitro*. *Nucleic Acids Res*, 1990. **18**(21): p. 6409-12.
244. Breslauer KJ, Frank R, Blöcker H, and Marky LA, *Predicting DNA duplex stability from the base sequence*. *Proc Natl Acad Sci USA*, 1986. **83**(11): p. 3746-50.
245. Borer PN, Dengler B, Tinoco I, Jr., and Uhlenbeck OC, *Stability of ribonucleic acid double-stranded helices*. *J Mol Biol*, 1974. **86**(4): p. 843-53.
246. Dalby B, Cates S, Harris A, Ohki EC, Tilkins ML, Price PJ, and Ciccarone VC, *Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications*. *Methods*, 2004. **33**(2): p. 95-103.
247. Sjöquist J, Meloun B, and Hjelm H, *Protein A isolated from Staphylococcus aureus after digestion with lysostaphin*. *Eur J Biochem*, 1972. **29**(3): p. 572-78.
248. Laemmli UK, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. *Nature*, 1970. **227**(5259): p. 680-85.
249. Burnette WN, *"Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A*. *Anal Biochem*, 1981. **112**(2): p. 195-203.
250. Carrington JC and Dougherty WG, *A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing*. *Proc Natl Acad Sci USA*, 1988. **85**(10): p. 3391-95.
251. Stanley P, *Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity*. *Mol Cell Biol*, 1989. **9**(2): p. 377-83.
252. Casasnovas JM, Larvie M, and Stehle T, *Crystal structure of two CD46 domains reveals an extended measles virus-binding surface*. *EMBO J*, 1999. **18**(11): p. 2911-22.
253. Müller S, Zocher G, Steinle A, and Stehle T, *Structure of the HCMV UL16-MICB complex elucidates select binding of a viral immunoevasin to diverse NKG2D ligands*. *PLoS Pathog*, 2010. **6**(1): p. e1000723.
254. Neu U, *Structural and functional investigations on VP1, the major capsid protein of simian virus 40*, in *Faculty of Chemistry and Pharmacy 2006*, University of Tübingen.

255. Åkerström B and Björck L, *A physicochemical study of protein G, a molecule with unique immunoglobulin G-binding properties.* J Biol Chem, 1986. **261**(22): p. 10240-47.
256. Klostermann F, *Purification of herpes simplex type 1 glycoprotein C and characterisation of an unknown protein from Staphylococcus aureus,* in *Faculty of Science 2015,* University of Tübingen.

7. PUBLICATIONS AND CONTRIBUTIONS

Buch MHC, Liaci AM, O'Hara SD, Garcea RL, Neu U, Stehle T (2015) Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity. *PLoS Pathogens* 11: e1005104

© 2015, Buch *et al.*

MHCB purified RA VP1 and solved the structures presented in the publication. Also, he developed the density integration procedure together with AML and UN, and he carried it out. Furthermore, he wrote the paper together with AML, UN, SDO, RLG, and TS.

Ströh LJ, Neu U, Blaum BS, **Buch MHC**, Garcea RL, Stehle T (2014) Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site. *J Virol.* 88(18): 10,831-10,839

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

MHCB cloned and provided the expression constructs for HPyV6 And HPyV7 VP1.

Magaldi TG, **Buch MHC**, Murata H, Erickson KD, Neu U, Garcea RL, Peden K, Stehle T, DiMaio D (2012) Mutations in the GM1 Binding site of Simian Virus 40 Alter Receptor Usage and Cell Tropism. *J Virol.* 86(13): 7,028-7,042

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

MHCB modeled and verified the A70L, A70L/N138Y, and A70L/F75L/H129Q mutations and generated Figures 2 and 10.

RESEARCH ARTICLE

Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity

Michael H. C. Buch¹*, A. Manuel Liaci¹*, Samantha D. O'Hara², Robert L. Garcea², Ursula Neu^{1,3*}, Thilo Stehle^{1,3*}

1 Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany, **2** Department of Molecular, Cellular, and Developmental Biology, and the BioFrontiers Institute, University of Colorado, Boulder, Colorado, United States of America, **3** Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America

* These authors contributed equally to this work.

‡ Current Address: MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London, UK

* uneu@nimr.mrc.ac.uk (UN); thilo.stehle@uni-tuebingen.de (TS)



 OPEN ACCESS

Citation: Buch MHC, Liaci AM, O'Hara SD, Garcea RL, Neu U, Stehle T (2015) Structural and Functional Analysis of Murine Polyomavirus Capsid Proteins Establish the Determinants of Ligand Recognition and Pathogenicity. *PLoS Pathog* 11(10): e1005104. doi:10.1371/journal.ppat.1005104

Editor: Craig Meyers, Penn State University School of Medicine, UNITED STATES

Received: March 18, 2015

Accepted: July 22, 2015

Published: October 16, 2015

Copyright: © 2015 Buch et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files except for the protein sequences of MuPyV RA VP1 and PTA VP1, which are available from UniProt under the accession numbers P49302 and Q76TX8, respectively, as well as the PDB entries of the constructs used in this publication (PDB IDs 5CPU, 5CPW, 5CPX, 5CPY, 5CPZ, and 5CQ0). Please also refer to [Table 2](#) in the manuscript.

Funding: MHCB, AML, UN, and TS are grateful for financial support from the Deutsche Forschungsgemeinschaft SFB 685 (www.sfb685.de).

Abstract

Murine polyomavirus (MuPyV) causes tumors of various origins in newborn mice and hamsters. Infection is initiated by attachment of the virus to ganglioside receptors at the cell surface. Single amino acid exchanges in the receptor-binding pocket of the major capsid protein VP1 are known to drastically alter tumorigenicity and spread in closely related MuPyV strains. The virus represents a rare example of differential receptor recognition directly influencing viral pathogenicity, although the factors underlying these differences remain unclear. We performed structural and functional analyses of three MuPyV strains with strikingly different pathogenicities: the low-tumorigenicity strain RA, the high-pathogenicity strain PTA, and the rapidly growing, lethal laboratory isolate strain LID. Using ganglioside deficient mouse embryo fibroblasts, we show that addition of specific gangliosides restores infectability for all strains, and we uncover a complex relationship between virus attachment and infection. We identify a new infectious ganglioside receptor that carries an additional linear [α-2,8]-linked sialic acid. Crystal structures of all three strains complexed with representative oligosaccharides from the three main pathways of ganglioside biosynthesis provide the molecular basis of receptor recognition. All strains bind to a range of sialylated glycans featuring the central [α-2,3]-linked sialic acid present in the established receptors GD1a and GT1b, but the presence of additional sialic acids modulates binding. An extra [α-2,8]-linked sialic acid engages a protein pocket that is conserved among the three strains, while another, [α-2,6]-linked branching sialic acid lies near the strain-defining amino acids but can be accommodated by all strains. By comparing electron density of the oligosaccharides within the binding pockets at various concentrations, we show that the [α-2,8]-linked sialic acid increases the strength of binding. Moreover, the amino acid

SDO and RLG acknowledge the National Institutes of Health R01-CA37667 (www.nih.gov). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

exchanges have subtle effects on their affinity for the validated receptor GD1a. Our results indicate that both receptor specificity and affinity influence MuPyV pathogenesis.

Author Summary

Viruses are obligate intracellular pathogens, and all of them share one crucial step in their life cycle—the attachment to their host cell via cellular receptors, which are usually proteins or carbohydrates. This step is decisive for the selection of target cells and virus entry. In this study, we investigated murine polyomavirus (MuPyV), which attaches to host gangliosides with its major capsid protein, VP1. We have solved the crystal structures of VP1 in complex with previously known interaction partners as well as with the ganglioside GT1a, which we have identified as a novel functional receptor for MuPyV. Earlier studies have shown that different strains with singular amino acid exchanges in the receptor binding pocket of VP1 display altered pathogenicity and viral spread. Our investigations show that, while these exchanges do not abolish binding or significantly alter interaction modes to our investigated carbohydrates, they have subtle effects on glycan affinity. The combination of receptor specificity, abundance, and affinity reveals a much more intricate regulation of pathogenicity than previously believed. Our results exemplify how delicate changes to the receptor binding pocket of MuPyV VP1 are able to drastically alter virus behavior. This system provides a unique example to study how the first step in the life cycle of a virus can dictate its biological properties.

Introduction

The engagement of one or several host cell receptors is the first step in the infectious cycle of a virus. A large number of viruses, including many human pathogens, depend on carbohydrate recognition for initial attachment to the cell surface. Viral tropism and the internalization pathway are usually determined by the specificity and affinity of the receptor interaction as well as the glycan distribution on different cell surfaces (reviewed in [1]). Many viruses use glycoproteins, glycolipids, or both as receptors for cell entry [2]. Gangliosides are ubiquitous glycolipids on the outer leaflet of mammalian cell membranes that serve as receptors for a number of viruses. They are composed of a membrane-embedded ceramide moiety linked to a complex carbohydrate structure that projects away from the cell. Gangliosides almost always contain α -5-*N*-acetyl-neuraminic acid (sialic acid, Neu5Ac) that can be attached to the core of the molecule with $[\alpha$ -2,3], $[\alpha$ -2,6], or $[\alpha$ -2,8] linkages (Fig 1). Gangliosides exist on cell surfaces in complex and poorly understood patterns that are cell type-, age-, and tissue-dependent ([3,4], reviewed in [5]).

Murine Polyomavirus (MuPyV) is a double-stranded DNA virus that can induce tumors in newborn animals. It was long known to engage glycan receptors that contain a minimal motif of sialic acid $[\alpha$ -2,3]-linked to galactose [6,7], and more recently gangliosides GD1a and GT1b were identified as MuPyV receptors [8]. Viral attachment is mediated by the major capsid protein, VP1, which forms pentameric capsomers that assemble into the $T = 7d$ icosahedral capsid of the virus [9,10]. Sialylated oligosaccharide receptors are engaged in a shallow groove on top of VP1 formed by loop structures on the protein surface [11–13], similar to other polyomaviruses [1].

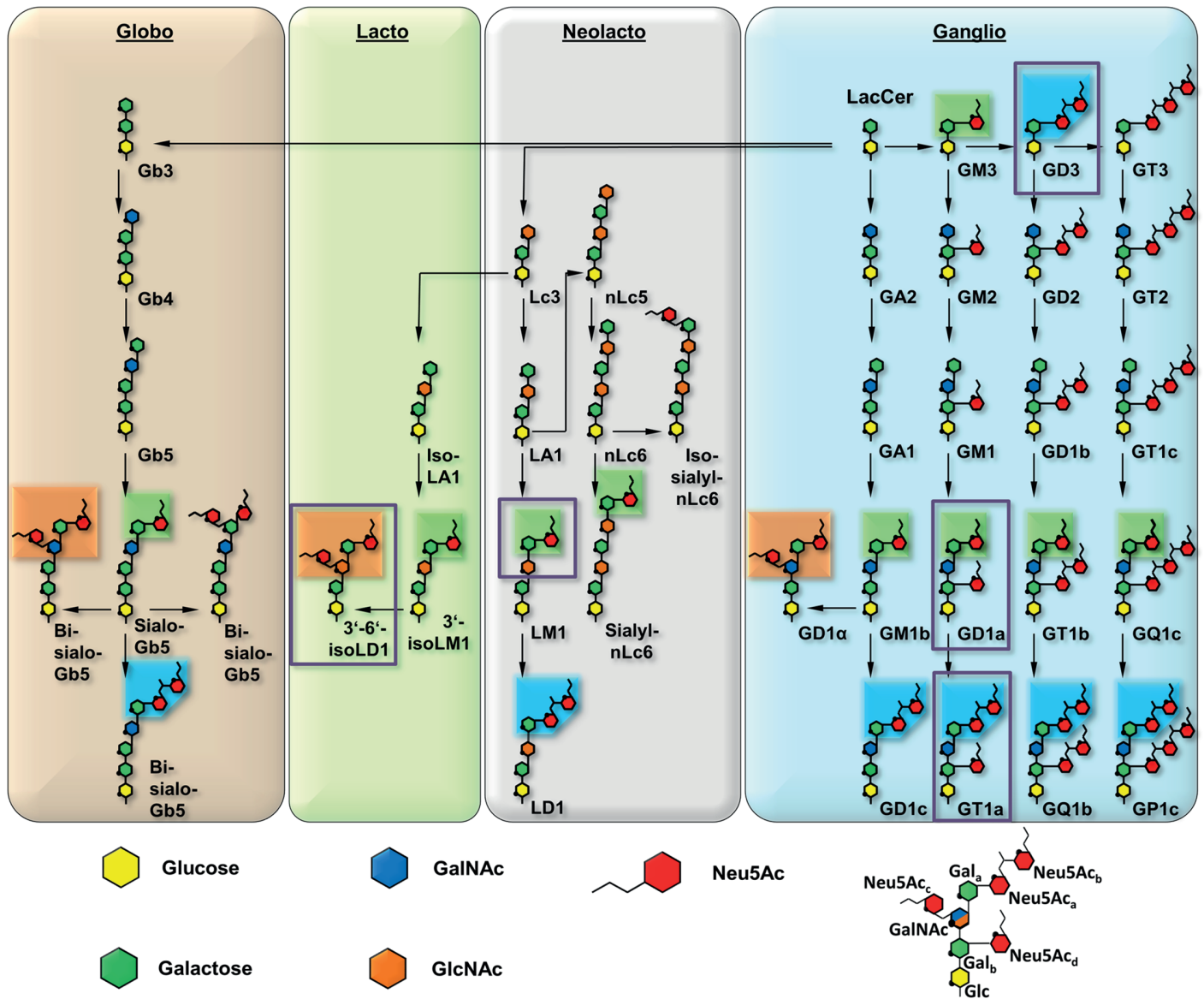


Fig 1. Overview and biosynthetic pathway of the four most prominent ganglioside series. The glycan parts of important members are shown for each series. The downstream biosynthetic steps are identical for all members of a row, although they may vary in linkage orientation. The six-membered pyranose rings are numbered counterclockwise, starting from the bottom (C1, except for C2 in Neu5Ac), and the ring oxygen is symbolized with a black dot. Neu5Ac moieties are rearranged for clarity, and all linkages are mediated by O2 or O8. Most of the gangliosides (e.g. LM1) can be further modified, e.g. by fucosylation. Linkages involving Neu5Ac are present in the α conformation, all other linkages are in the β conformation. Boxes represent three distinguishable sialoglycotopes that contain linkages found in GT1a (blue, representative for $[\alpha-2,8]$), GD1a (green, $[\alpha-2,3]$), and 3'-6'-iso-LD1 (also referred to as DSLNT, orange, $[\alpha-2,6]$). The naming is according to the corresponding gangliosides; if possible, the Svennerholm shorthand is used [64–66]. All biosynthetic routes were verified using the KEGG metabolic pathway database [67]. A prototype glycan that exemplifies the different positions of Gal and Neu5Ac moieties is depicted on the lower right. The glycan portions investigated in this study are highlighted by purple boxes.

doi:10.1371/journal.ppat.1005104.g001

MuPyV displays striking differences in pathogenicity and spread among three closely related prototype strains upon infection of newborn virus-free mice. The laboratory-derived RA strain [14] shows limited spread and induces few tumors of strictly mesenchymal origin after a long latency period, while the naturally occurring PTA strain [15,16] has disseminated infection and causes multiple tumors of epithelial and mesenchymal origin within a short time.

Table 1. Description of the investigated MuPyV strains.

	RA Strain	PTA Strain	LID Strain
Distinctive amino acids	G91, V296	E91, V296	E91, A296
Pathogenicity	No or only singular tumors, mesenchymal origin.	High tumor density of epithelial and mesenchymal origin.	Virulent. Damage of host tissues, early death due to brain hemorrhages and kidney failure.
Latency	Long	Short	Very short

doi:10.1371/journal.ppat.1005104.t001

LID [17,18], another laboratory isolate MuPyV strain, spreads most rapidly, causing early death by damaging host tissues, leading to brain hemorrhages and kidney failure [19]. The differences among the three strains have been mapped to amino acid variations at two positions, 91 and 296, within the receptor-binding region of VP1 [20–24]. While RA bears a glycine residue at position 91, this residue is replaced with a glutamate in both PTA and LID. An additional valine-to-alanine exchange at position 296 is present in LID (Table 1). The pathogenicity profile of one strain can be introduced into the other strains by mutating these two residues, confirming that these substitutions are necessary and sufficient to generate a specific phenotype [25]. The same substitutions have also been observed for other strains of MuPyV [21,22]. MuPyV found in feral mice has the VP1 sequence of PTA [26], but the virus is controlled by an intact immune system. As studies of viral spread can be conducted *in vivo* and virus infectivity can be tested *in vitro* using ganglioside deficient mouse cells, MuPyV represents an attractive and rare model system to define the relationships between receptor binding and viral spread and tropism.

Crystal structures of the low pathogenicity strain RA have shown how this virus engages 3'-sialyllactose, a short, linear trisaccharide terminating in [α -2,3]-linked sialic acid, as well as an oligosaccharide that additionally contains a second, branching [α -2,6]-linked sialic acid [11,12]. These structures also identified the location of residues 91 and 296 in the carbohydrate-binding region, suggesting that they might modulate interactions of VP1 with its receptors in the higher pathogenicity strains PTA and LID. Modelling suggested that a glutamate side chain at position 91 would lead to electrostatic repulsion of the [α -2,6]-branched sialic acid, thereby preventing binding of such branched structures by either LID or PTA. Branched sugars carrying an [α -2,6]-linked sialic acid could thus act as pseudoreceptors that will not facilitate productive infection but hamper the spread of RA within the host, in contrast to PTA and LID [8,12]. In line with this hypothesis, gangliosides GD1a and GT1b, which do not contain an [α -2,6]-branched sialic acid, have been identified as entry receptors for the PTA [8,16] and RA strain [27] of MuPyV. However, the molecular determinants of GD1a or GT1b receptor interactions with PTA and LID are not understood, because all structural information is limited to date to RA MuPyV.

To define the interactions of the three MuPyV strains with receptors on the cell surface, we have solved high-resolution structures of RA, PTA, and LID VP1 pentamers in complex with three ganglioside glycans that represent common motifs found in members of the four most prominent ganglioside biosynthesis series and that feature [α -2,3]-, [α -2,6]-, and [α -2,8]-linked sialic acids (for carbohydrate structures, nomenclature, and annotations see Fig 1). We have also conducted crystallographic soaking experiments at different ligand concentrations to compare the relative affinities of each of the three strains for their interaction partners. We find that expanding the well-characterized Neu5Ac- $[\alpha$ -2,3]-Gal epitope with a linear [α -2,8]-linked sialic acid (as found for example in GT1a vs. GD1a) leads to additional interactions between carbohydrate and VP1 in all three strains. Consequently, we identify ganglioside GT1a as an infectious receptor for all three strains. Moreover, the branching [α -2,6]-linked sialic acid is

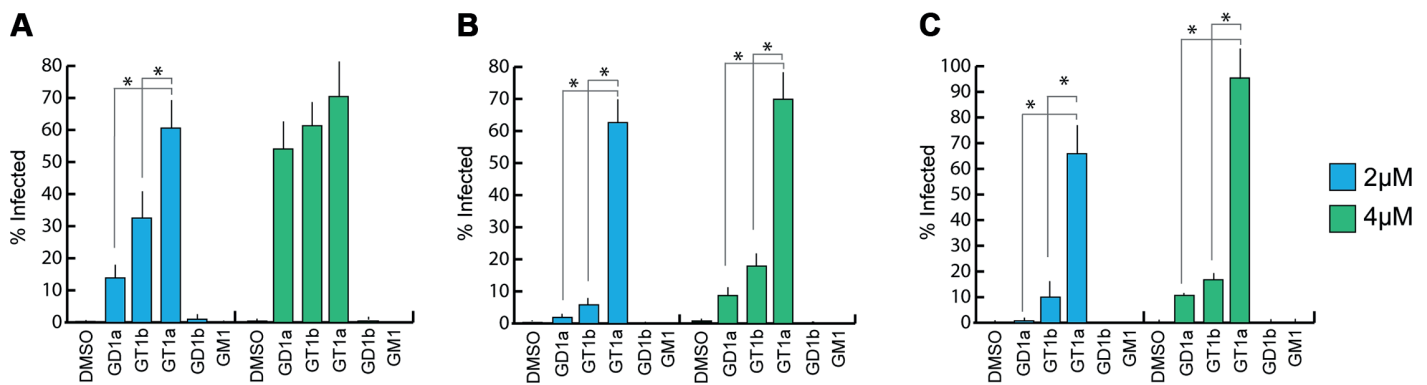


Fig 2. GT1a, GT1b, and GD1a supplementation rescues MuPyV infection of Gang^{-/-} MEFs. Ganglioside knock-out (Gang^{-/-}) MEFs were completely resistant to infection of all strains of MuPyV as shown by the absence of T-antigen positive nuclei at 24 hours post infection (DMSO control). GD1a, GT1b, and GT1a ganglioside supplementation of Gang^{-/-} MEFs restored RA (A), PTA (B), and LID MuPyV (C) infection, while GD1b and GM1 supplementation resulted in little to no infection by any virus strain. Infection levels were quantified at both 2 μM and 4 μM ganglioside supplementation (blue and green bars, respectively). Infection levels are normalized to MuPyV infection of WT MEFs, and error bars correspond to standard error.

doi:10.1371/journal.ppat.1005104.g002

close to the strain-defining amino acids, but can be accommodated by all strains, in contrast to the earlier model. However, the amino acid exchanges defining each strain have subtle effects on their affinity for the validated receptor GD1a. Our results exemplify the effect of minimal changes in a binding pocket on the receptor binding properties of a virus.

Results

GT1a, GD1a, and GT1b gangliosides are infectious receptors for MuPyV

Previous efforts to identify receptors for MuPyV used immortalized cell lines, such as Vero or C6 glioma cells that were supplemented with candidate gangliosides before infection [8,28]. We utilized a mouse embryo knock-out fibroblast cell line (Gang^{-/-} MEFs) specifically deficient in ganglioside synthesis and completely resistant to MuPyV infection (S1A Fig and [29]) to test the ability of ganglioside receptors to rescue infection by different strains of MuPyV. Gang^{-/-} MEFs were supplemented with individual gangliosides followed by infection with RA, PTA, and LID MuPyV (Fig 2). Importantly, it should be noted that the three MuPyV strains we used do not have the same particle to PFU ratio. The viruses have been normalized to similar MOIs, but they cannot be quantitatively compared to one another. However, each strain has been normalized to its own infection rate of WT MEFs; therefore, infection rates upon supplementation of gangliosides can be compared within a strain. The previously identified ganglioside receptors GD1a and GT1b [8] rescued RA, PTA, and LID infection of Gang^{-/-} MEFs in a dose-responsive manner. We also analyzed the GT1a ganglioside that had not been previously investigated as a candidate infectious receptor for MuPyV. We found that GT1a, a member of the ganglio-series synthesized from GD1a (Fig 1), also rescued RA, PTA, and LID infection in a dose responsive manner (Fig 2). Moreover, GT1a supplementation of Gang^{-/-} MEFs conferred higher levels of RA, PTA, and LID MuPyV infection than the previously identified receptors GD1a and GT1b. Finally, we tested the ability of the gangliosides GD1b and GM1 to rescue MuPyV infection of Gang^{-/-} MEFs. GD1b and GT1b supplementation has previously been shown to restore BK polyomavirus infection of ganglioside deficient cells [30]; however, GD1b restored little to no MuPyV infection of Gang^{-/-} MEFs. GM1 supplementation has previously been shown to restore infection by SV40 [8]; however, GM1 did not rescue MuPyV infection of Gang^{-/-} MEFs. These data confirm that GT1a is an infectious receptor for all strains of MuPyV.

We also investigated whether MuPyV cell surface binding to infectious or non-infectious ganglioside receptors correlated with infection. To this end, we measured the levels of free (unbound) virus in each ganglioside supplemented sample at 4 hours post infection. We did not detect significant differences in MuPyV cell surface binding to different ganglioside receptors or WT MEFs, indicating that cell surface binding alone does not determine infection (S1B Fig). Instead, a considerable amount of virus binds to Gang^{-/-} MEFs even in the absence of ganglioside supplementation (S1A Fig). MuPyV is also endocytosed in Gang^{-/-} MEFs, which however does not lead to infection [29]. Taken together, these data confirm that gangliosides are not required for cell surface binding. They are, however, required for infection, and GT1a appears to be more efficient than GD1a and GT1b.

Structure of MuPyV VP1 bound to GT1a

In order to define the mode of recognition of GT1a, particularly to the naturally occurring PTA strain of MuPyV, we have soaked VP1 crystals with the glycan portion of GT1a and solved the structure of the complex (Table 2). While the receptor interaction pocket of RA VP1 has been described [11–13], no structural information for the pathogenicity-defining amino acids at positions 91 and 296 in the pockets of PTA and LID has been available. PTA and LID both carry a glutamate at position 91, and this side chain is being held in a characteristic position with the carboxyl group facing away from the glycan receptor due to a salt bridge formed with K186 (Fig 3), as previously predicted [12]. The GT1a glycan is a branched structure with a long disialylated arm, which has the sequence Neu5Ac_b-[α-2,8]-Neu5Ac_a-[α-2,3]-Gal_a-[β-1,3]-GalNAc, and a second short arm, which consists of a single Neu5Ac_a [α-2,3]-linked to Gal_b (for carbohydrate structures, nomenclature, and moiety indexing see Fig 1). The disialylated arm of GT1a is clearly visible in the crystal structure of PTA VP1; it is well defined by electron density and makes extensive contacts with the protein (Fig 4B–4D). Overall, the GT1a glycan adopts a twisted horseshoe-like shape, with Neu5Ac_a and Neu5Ac_b wrapping around the side chains of Y72 and R77 of VP1. Its longer, disialylated arm contains a Neu5Ac_a-[α-2,3]-Gal_a sequence that is also present in GD1a and simpler compounds such as 3'-sialyllactose (3SL), and the interactions of this motif with VP1 are essentially identical to those seen in previous structures [11–13]. However, our structure visualizes an additional network of contacts made by the terminal [α-2,8]-linked Neu5Ac_b (Fig 4C and 4D). Its carboxyl group engages Y72 and forms water-mediated hydrogen bonds with Q71, Y72, as well as D85 of the neighboring monomer (D85^{*}). In addition, the *N*-acetyl nitrogen of Neu5Ac_b forms a hydrogen bond with the backbone carbonyl of T67, and O8 and O9 in the glycerol chain of the sugar are hydrogen-bonded with the R77 side chain. The carboxyl groups of Neu5Ac_a and Neu5Ac_b are about 4 Å apart, and the positively charged side chain of R77 counteracts their negative charges (Fig 4C and 4D). Neu5Ac_a and Neu5Ac_b contribute binding interfaces of approximately 160 Å² and 190 Å², respectively (calculated using the PISA server [31]). The remaining Gal_a-GalNAc-Gal_b stem of GT1a forms fewer contacts with the protein, which include a hydrogen bond between G78 and the Gal_a O4 hydroxyl group (Fig 4) as well as several van der Waals interactions. Notably, the C_β and C_γ atoms of E91 are within van-der-Waals range of O6 and C6 of Gal_a, and the E91 carboxylate group is close to C6 of GalNAc. The total contact surface for this portion of the glycan is 142 Å².

Because the differences in tumorigenicity and host spread among strains have been mapped to the glycan binding pocket of VP1, and because GT1a appears to be particularly efficient in facilitating productive infection, we set out to determine how the three strains engage GT1a. By solving the crystal structures of RA and LID VP1 complexed with GT1a using the identical strategy used for the PTA-GT1a complex, we found that the overall binding mode of GT1a is

Table 2. Data collection and refinement statistics.

	PTA VP1 Native	PTA VP1 + GT1a	PTA VP1 + GD1a	PTA VP1 + DSLNT	RA VP1 + GT1a	RA VP1 + GD1a
Data Collection						
Beamline	SLS, X06DA	SLS, X06DA	SLS, X06DA	ESRF, ID 14–1	SLS, X06DA	SLS, X06DA
Space Group	P3 ₁ 21	P3 ₁ 21	P3 ₁ 21	P3 ₁ 21	P3 ₁ 21	P3 ₁ 21
Cell Dimensions						
a = b, c [Å]	219.61, 99.82	219.60, 99.74	220.45, 99.71	219.73, 100.00	219.55, 99.60	219.06, 99.40
α = β, γ [°]	90, 120	90, 120	90, 120	90, 120	90, 120	90, 120
Resolution [Å]	50–1.64 (1.68–1.64)	50–1.75 (1.79–1.75)	50–1.93 (1.98–1.93)	50–1.87 (1.92–1.87)	50–1.71 (1.75–1.71)	50–1.90 (1.95–1.90)
R _{meas} [%]	10.5 (68.7)	13.3 (68.3)	11.3 (68.8)	15.2 (69.1)	7.4 (74.1)	11.1 (71.7)
I/σ(I)	10.5 (2.3)	7.08 (1.51)	12.8 (3.0)	7.6 (3.0)	15.71 (2.17)	11.3 (1.9)
Completeness [%]	99.9 (99.9)	97.6 (96.6)	96.0 (98.1)	99.8 (99.8)	99.8 (99.9)	99.0 (99.1)
Redundancy	5.0 (4.8)	3.4 (3.2)	5.4 (5.5)	3.7 (3.7)	4.3 (3.9)	2.9 (2.8)
Wilson B-Factor [Å ²]	23.1	25.3	25.7	23.1	26.0	23.3
Refinement						
Resolution [Å]	48.2–1.64	47.6–1.75	47.8–1.93	48.4–1.83	50–1.71	48.1–1.90
No. of Reflections	324,802	261,253	192,327	220,105	285,887	205,733
R _{work} / R _{free} [%]	15.85 / 17.30	16.0 / 18.13	15.27 / 17.42	15.38 / 17.56	15.27 / 16.98	15.46 / 17.84
No. of Atoms						
Protein	11,117	11,088	11,150	10,996	11,323	11,225
Solvent	1,840	1,827	1,884	1,632	2,059	1,971
Carbohydrate	-	425	285	202	385	285
B-Factors [Å ²]						
Protein	19.3	20.7	21.4	20.1	21.1	19.9
Solvent	29.6	30.9	31.5	30.5	32.6	31.1
Carbohydrate	-	30.3	35.1	34.2	36.0	38.2
R. m. s. Deviations						
Bond Lengths [Å]	0.007	0.007	0.006	0.008	0.007	0.008
Bond Angles [°]	1.16	1.17	1.06	1.20	1.13	1.19
Ramachandran Plot						
Favored	1,340 (97.2%)	1,336 (97.0%)	1,335 (96.9%)	1,334 (96.9%)	1,342 (97.0%)	1,340 (97.0%)
Allowed	38 (2.8%)	41 (3.0%)	42 (3.2%)	42 (3.1%)	42 (3.0%)	42 (3.0%)
Disallowed	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
PDB ID	5CPU	5CPW	5CPY	5CPX	5CPZ	5CQO

doi:10.1371/journal.ppat.1005104.t002

very similar across the three strains (Fig 5A), with a conserved binding mode of the [α-2,8]-linked Neu5Ac_b. Although the replacement of glutamate with glycine at position 91 leads to a contact area decrease of 33 Å² in RA, the orientation of GT1a in this strain is not altered (compare Fig 5B and 5C). Likewise, the substitution of valine with alanine at position 296 in LID removes a hydrophobic contact but does not affect the conformation of GT1a (Fig 5E; S2 Fig).

The Neu5Ac_a-Gal_a-GalNAc linkages in the long arm of GT1a adopt conformations that have been reported in numerous structures (for example [32–34]). While the [α-2,3] linkage between Neu5Ac_a and Gal_a adopts the conformation that has been reported for DSLNT and 3SL, the branching Neu5Ac_d-[α-2,3]-Gal_b linkage adopts a different conformation, which has been reported for structures containing O-4-substituted galactoses (as in [35,36]). While a higher variability is observed for Neu5Ac-[α-2,8]-Neu5Ac linkages (S2E Fig), this linkage adopts torsion angles that are in agreement with other, related structures such as in the

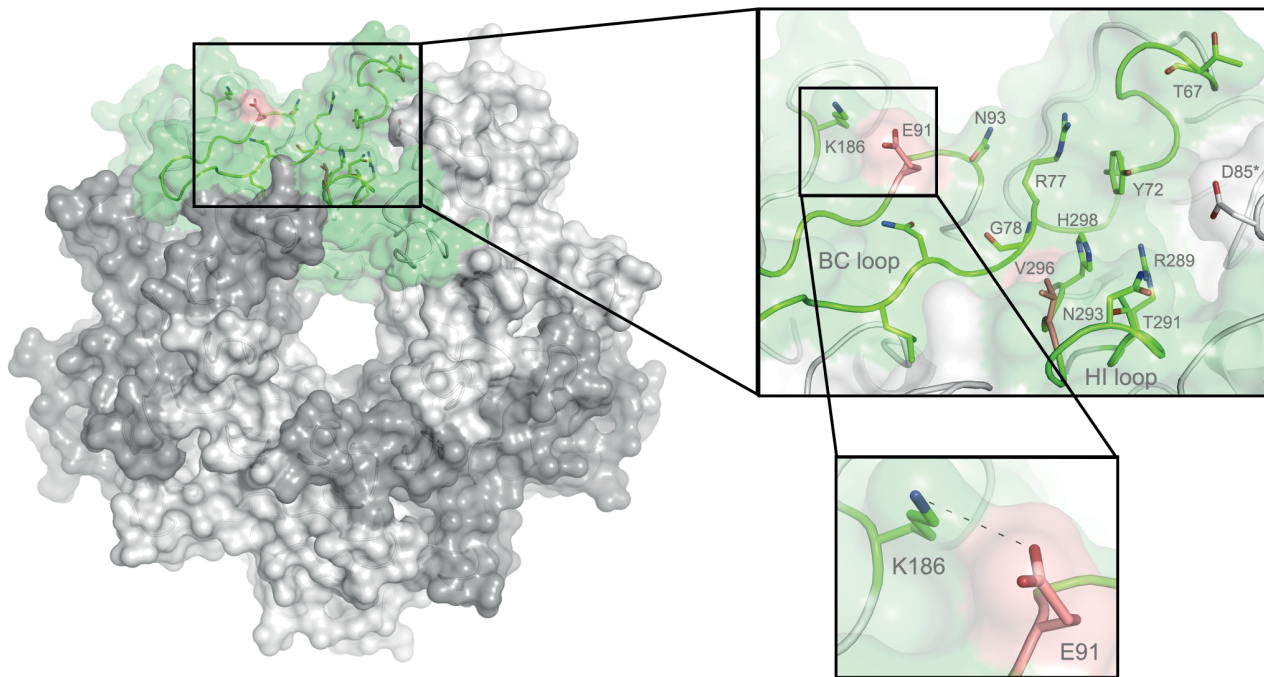


Fig 3. The MuPyV binding pocket. Top view on the receptor-binding region of PTA, which is shown with E91 and V296 highlighted in salmon. Residues that are known to participate in receptor binding are contributed by the BC and HI loops and are highlighted as stick models. One monomer is shaded in green and the other monomers are alternatingly shaded light and dark grey for better distinction.

doi:10.1371/journal.ppat.1005104.g003

structure of human liver fructose-1,6-bisphosphatase in complex with an allosteric inhibitor [37] or in the complex of tetanus toxin with a GT1b analog [38]. The overall structure is in good agreement with a molecular dynamics simulation using an AMBER force field in an aqueous environment [39]. A well-defined set of water molecules mediates bridged hydrogen bonds between the pyranose moieties, especially between Neu5Ac_b and Neu5Ac_d (S3 Fig). Due to these steric constraints, the GT1a complexes feature well-defined electron density not only for the binding epitope, but also for the non-binding, branching NeuNAc_d in its preferred solution conformation [40], which brings this moiety to about 5 Å near the end of the long arm and gives the glycan the characteristic, horseshoe-like topology that is observable in all complex structures.

Structures of MuPyV VP1 strains bound to other sialylated glycans

As RA, PTA, and LID VP1 all bind GT1a in a highly similar conformation, we hypothesized that the differences in pathogenicity and spread among the three strains might be due to the recognition of additional carbohydrates by only a subset of MuPyV strains. As shown in Fig 1, the many different gangliosides share a relatively small set of common sialoglycotopes. We therefore investigated the ability of all three VP1 proteins to bind other glycan structures that are representative for these epitopes. We solved structures of VP1 bound to the glycan portions of two of these gangliosides: The GD1a glycan is an established infectious receptor and essentially a truncated version of GT1a lacking the [α-2,8]-linked Neu5Ac_b in the long arm. The human milk hexasaccharide DSLNT is the glycan portion of the lacto-series ganglioside 3'-6'-isoLD1 (Fig 1) [41], which is overexpressed in the central nervous system. In contrast to GT1a and GD1a, DSLNT does not contain an [α-2,3]-linked Neu5Ac_d as a short arm but instead a

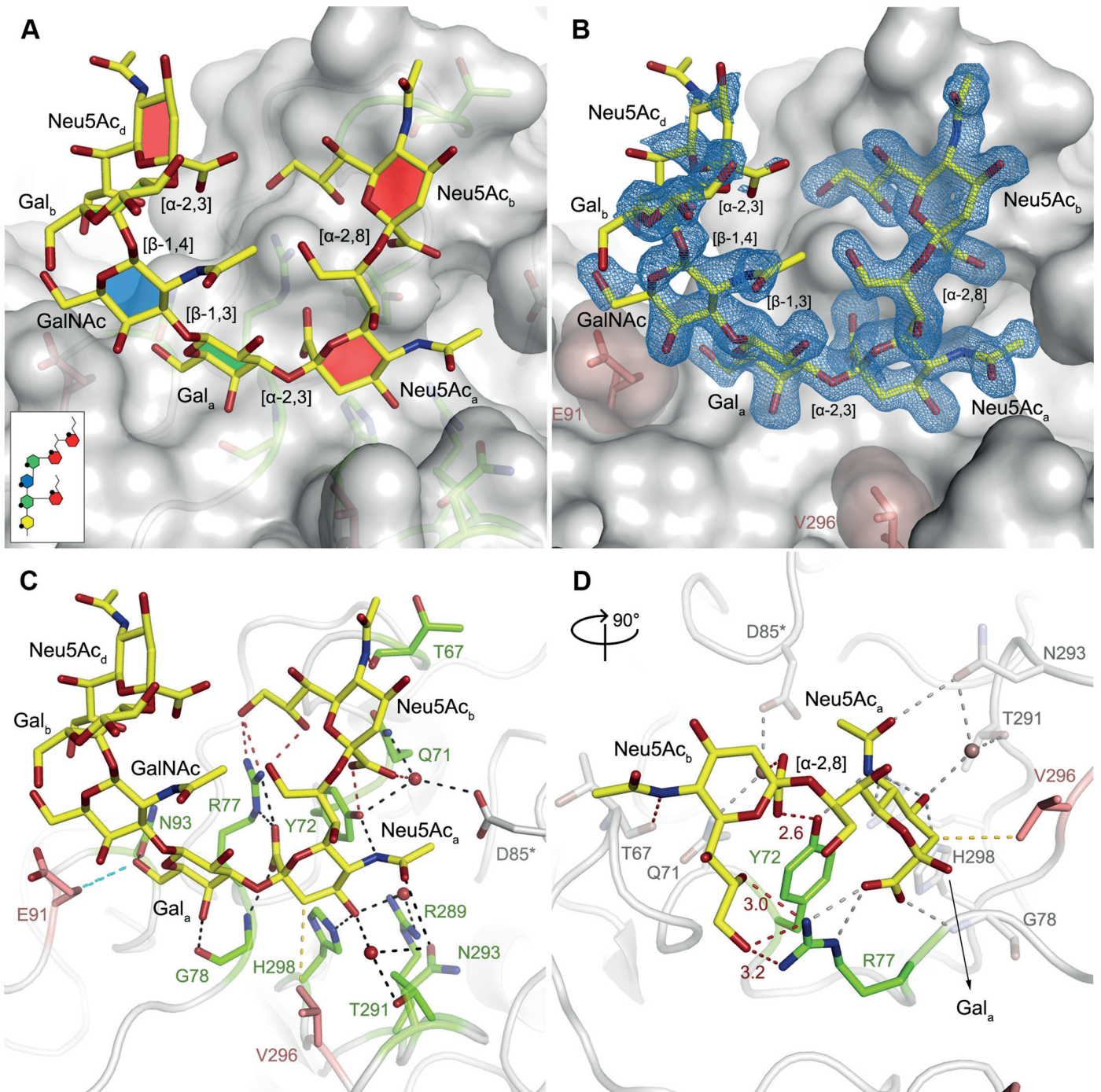


Fig 4. Binding of GT1a to PTA. **A** The PTA binding pocket and the GT1a conformation upon binding are shown from an angle parallel to the fivefold axis. A scheme of the glycan is shown in the inset, and the sugar rings are filled according to the coloring scheme from Fig 1. **B** Simulated annealing $F_{obs}-F_{calc}$ omit map (resolution 1.71 Å, calculated at 3.5 σ , carved 2.3 Å around the glycan). **C** Possible binding interactions of GT1a and PTA. E91 and V296 are highlighted in salmon. Hydrogen bonds are shown in black, the hydrophobic contact mediated by V296 in gold, and the van-der-Waals contacts of E91 are shown in cyan. Waters that mediate key hydrogen bonds are shown as red spheres. Unique interactions mediated by the novel GT1a-like binding motif are shown in red. **D** Zoomed view of the binding to the two terminal Neu5Ac moieties. The rest of the glycan is omitted for clarity. Residues except Y72 and R77 as well as waters involved in contacts with these two glycan moieties are pale grey and salmon, respectively.

doi:10.1371/journal.ppat.1005104.g004

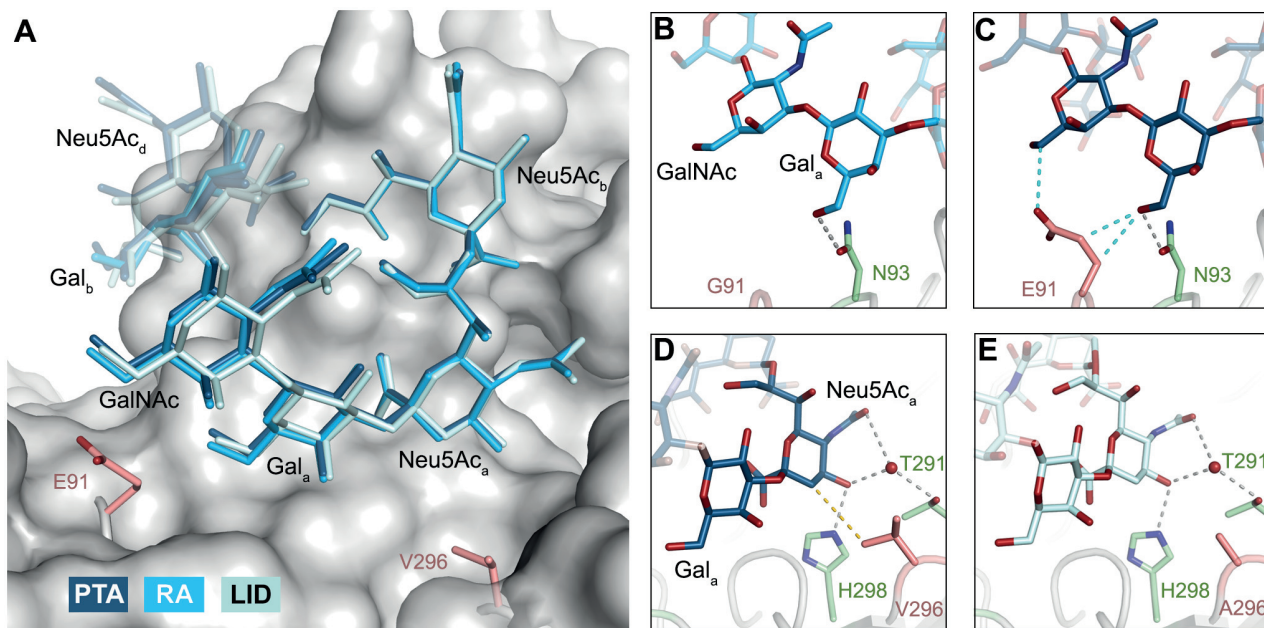


Fig 5. Binding modes of GT1a to the different MuPyV strains. **A** Superposition of the GT1a-binding mode of RA (GT1a in sky blue), PTA (dark blue), and LID (pale blue). The Neu5Ac_b-[α -2,8]-Neu5Ac_a-[α -2,3]-Gal_a motif is shown in solid sticks, together with the adjacent GalNAc moiety. All superpositionings were carried out in PyMOL [61] using 'align' for the protein chains only. Surface, E91 and V296 are from PTA/GT1a. All 'align' rmsd values are below 0.16 Å. **B & C** Close view of the van-der-Waals contacts introduced by the E91 side chain present in PTA and LID (C), but not in RA (B). Hydrogen bonds are shown in grey, van-der-Waals contacts in cyan. **D & E** Close view of the hydrophobic contact mediated by V296 in RA and PTA (D), but not by A296 in LID (E). The 4.0 Å hydrophobic contact is not present in the LID strain, whose pocket is opened to the right. Hydrogen bonds are shown in grey, hydrophobic contacts are shown in gold.

doi:10.1371/journal.ppat.1005104.g005

branching [α -2,6]-linked Neu5Ac_c. This structure is similar to a very common epitope on O-linked glycoproteins [42–44]. DSLNT was used in previous studies of MuPyV as a model “pseudoreceptor” [12] and was investigated here to help rationalize these earlier data, to facilitate a comparison among strains, and to establish a binding profile for glycans containing an [α -2,6]-linked sialic acid.

GD1a. The previously identified MuPyV receptor GD1a is similar to a truncated GT1a structure, containing only a Neu5Ac_a-[α -2,3]-Gal_a motif instead of Neu5Ac_b-[α -2,8]-Neu5Ac_a-[α -2,3]-Gal. The disaccharide engages all three strains in a very similar manner (Fig 6A). Neither the longer E91 side chain (in PTA and LID) nor the shorter A296 side chain (in LID) result in an altered conformation of the ligand.

DSLNT. The DSLNT glycan terminates in a Neu5Ac_a-[α -2,3]-Gal_a motif, which is the part of the molecule best defined by electron density in all complexes. DSLNT also contains an additional [α -2,6]-linked, branched Neu5Ac_c residue, which is not present in either GT1a or GD1a. There is weak electron density for Neu5Ac_c in one of the five binding pockets of the RA strain, but it only engages in few interactions [12]. While PTA and LID do bind DSLNT, the complex structures do not show any electron density for the Neu5Ac_c, indicating that this sugar is conformationally flexible and does not contribute contacts. When bound to the PTA strain, the stem of DSLNT is moderately rearranged (Fig 6B). In comparison to GalNAc in GT1a, the GlcNAc moiety is slightly tilted away from E91 due to a \sim 20° rotation of the psi angle in the Gal_a-[β -1,3]-GlcNAc linkage (Fig 6B and 6C, assessed using CARP) that propagates throughout the sugar. In addition, there is no visible electron density for the GlcNAc O6 that is engaged in the [α -2,6]-branching as well as an increased B-Factor variance within the

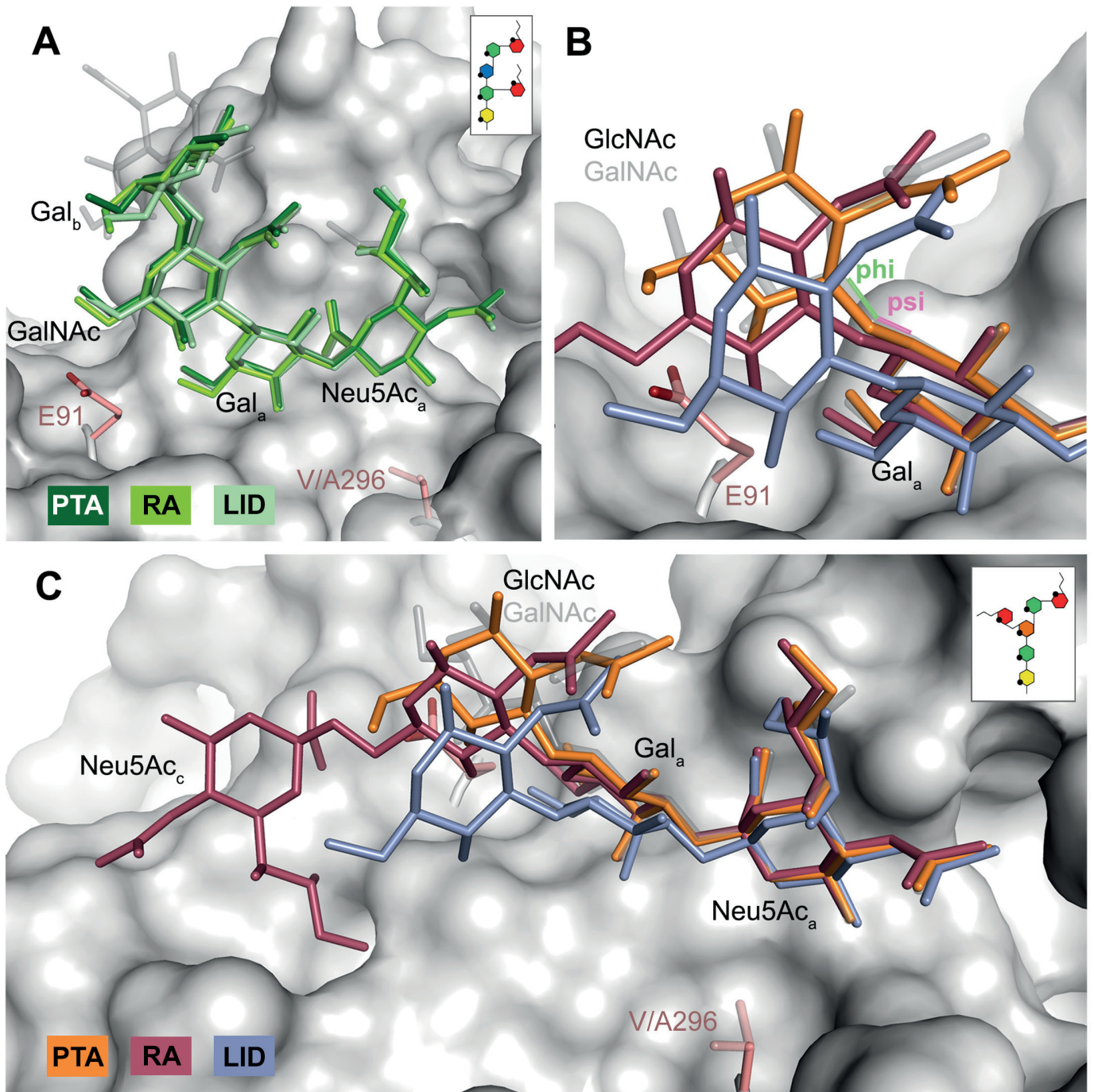


Fig 6. Differences in receptor binding patterns across glycans. **A** Superposition of the binding modes of GD1a to RA (light green), PTA (dark green), and LID (pale green). The sequence of GD1a is shown in the inset. The Neu5Ac_a-[α -2,3]-Gal_a motif is shown in solid sticks, together with the adjacent GalNAc moiety. In all figures, GT1a bound to PTA is overlaid as a grey ghost for comparison, with Neu5Ac_b omitted for clarity. Deviations exceeding the atomic error of the structure and alignment rmsd values are only found in the stem region of the sugar, starting at Gal_b. All superpositionings were carried out in PyMOL [61] using 'align' for the protein chains only. Surface, E91 and V296 are from PTA/GT1a. All 'align' rmsd values are below 0.16 Å. **B & C** Comparison of the DSLNT binding modes to RA (red), PTA (orange), and LID (violet). In PTA-DSLNT, [α -2,6]-branching causes a 15–20° psi angle shift of the GlcNAc moiety compared to GD1a and GT1a, resulting in a 1 Å sideward twist movement of the stem. In RA-DSLNT, combination of this shift with a 15° shift in the phi angle results in a downward movement of GlcNAc and its branching Neu5Ac_c compared to PTA-DSLNT. In LID, the shift is already observable for Gal_a and results in the loss of ordered density for GlcNAc. All angles were analyzed with CARP. The sequence of DSLNT is shown in the inset of panel C.

doi:10.1371/journal.ppat.1005104.g006

glycan (S4 Fig). The reason for the sideward twist and the missing electron density for Neu5Ac_c observed in PTA is probably the electrostatic repulsion between the carboxyl groups of Neu5Ac_c and E91. While the charge of E91 is compensated by K186 (Fig 3), as was hypothesized before [12], the two carboxylate groups would come within about 2 Å of one another if DSLNT bound to PTA in the same way as observed in RA. This hypothesis is confirmed by a PTA E91Q mutant that rescued binding of Neu5Ac_c (S5 Fig). In turn, when bound to RA, DSLNT exhibits a stronger conformational rearrangement (Fig 6B and 6C). Due to the missing side chain at position 91, the psi angle rotation between Gal_a and GlcNAc is accompanied by an additional 15° rotation of the phi angle, bringing the GlcNAc moiety and the attached Neu5Ac_c closer to the protein surface [12]. In LID, the valine to alanine mutation at position 296 reduces its van-der-Waals radius. This change results in a broader binding pocket compared to the other strains and the loss of a hydrophobic interaction between position 296 in VP1 and C3 of Neu5Ac_a for all glycans. This gives room for a stark alteration in the binding mode of DSLNT that starts with a slight tilt of Neu5Ac_a and propagates through the sugar (Fig 6C), ultimately resulting in a prominent sideward shift of the whole glycan stem. The resulting increased conformational freedom of DSLNT is reflected by a lack of electron density in its stem region as well as by an elevated temperature factor variance (S4 Fig). This alteration of the binding mode in LID is likely to be observed for other glycans that are not conformationally restrained by the [α-2,3]-linked Neu5Ac_d.

Relative affinities of MuPyV strains for sialylated glycans

Since all three MuPyV strains are able to engage the three different glycan structures in a largely identical manner, we reasoned that the differences in pathogenicity and spread might be attributable to subtle differences in affinity, rather than specificity, among the strains. The affinities of RA VP1 for 3'-sialyllactose and DSLNT were estimated to be in the low mM range [11]. Coupled with the high costs of glycans and the high amount required due to their low binding affinity, weak binding poses technical challenges for classical affinity measurements. We therefore utilized a crystallographic approach to quantitatively compare ligand binding. We crystallized all three VP1 pentamers in the same condition, and soaked each with the oligosaccharide portions of GT1a, GD1a, and DSLNT at different concentrations in parallel. X-ray data of all crystals were collected in the same manner, and the data sets were processed using the same protocol and integrated as described previously [45]. All data sets were processed in the same unit cell, scaled, and the bias-reduced difference electron density around the central Neu5Ac_a-[α-2,3]-Gal_a motif was quantified for each data set (see the Methods section for details). Our crystallization condition contains a high amount of ammonium sulfate, which competes with the carboxyl group of Neu5Ac_a and has to be displaced by the carbohydrates. Therefore, our observed binding is weaker than in a physiological setting. However, while not yielding dissociation constants in the traditional sense, this method enables us to compare relative levels of binding across our three different strains and three different glycans.

The GT1a glycan exhibits the strongest binding in all three VP1 variants compared with DSLNT or GD1a (Fig 7A–7C), with no detectable difference between the strains (Fig 7D). This finding is in accord with our ganglioside add-back experiments in cell culture (Fig 2), which consistently showed higher levels of infection mediated by GT1a compared to GD1a. The stronger overall binding of GT1a can be attributed to the additional [α-2,8]-linked sialic acid present in GT1a (Neu5Ac_b), which contributes several interactions and an increased buried surface area. These contacts seem to outweigh the differences in van der Waals contacts with the side chains of E91 or V296, at least to the extent discernable in our assay.

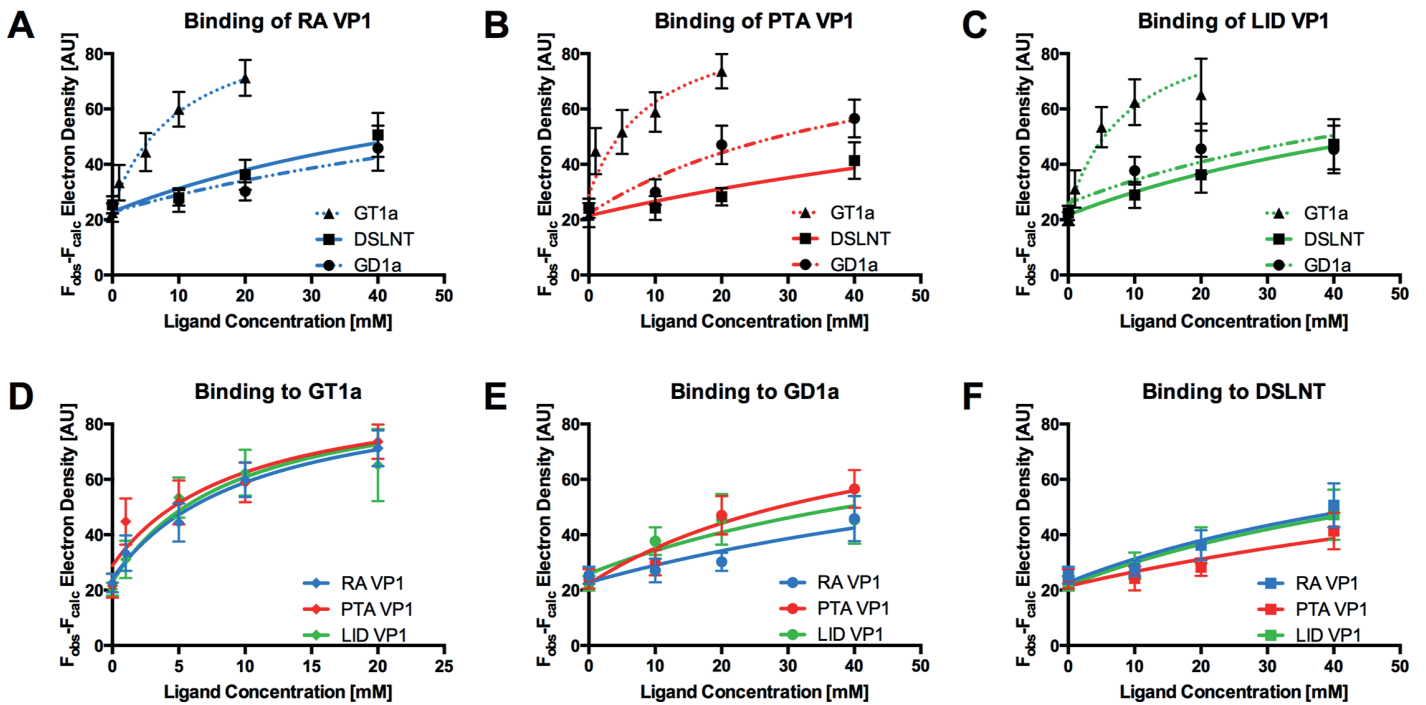


Fig 7. Binding of ligands to MuPyV VP1. The average simulated annealing $F_{obs}-F_{calc}$ electron density for the Neu5Ac₆-[α -2,3]-Gal₆ in GT1a, GD1a, and DSLNT is plotted against ligand concentration. RA VP1 is colored blue, PTA VP1, is colored red, and LID VP1 is colored green. GT1a is displayed in a dotted line with triangles, GD1a in a dashed line with circles, and DSLNT in a solid line with squares. The error bars correspond to the standard deviation of the mean electron density observed in the five chains of VP1. **A** Electron density of GT1a, GD1a, and DSLNT in RA VP1. **B** Same as in **A**, but for PTA VP1. **C** Same as in **A**, but for LID VP1. **D** Comparison of GT1a-derived electron density in RA VP1, PTA VP1, and LID VP1. **E** Same as in **D**, but for GD1a. **F** Same as in **D**, but for DSLNT.

doi:10.1371/journal.ppat.1005104.g007

GD1a binds less well to all strains compared to GT1a. In addition, there are differences in binding strength among the three strains. PTA and LID VP1 appear to bind GD1a at the same level and better compared with RA (Fig 7A–7C and 7E) because these two strains gain additional interaction surface and van-der-Waals contacts from their E91 side chain. This effect is more pronounced than in GT1a, because in GD1a it cannot be masked by the additional contacts of the [α -2,8]-linked Neu5Ac₆.

DSLNT displays the lowest overall affinity to all strains, with levels comparable to GD1a in RA for all three strains (Fig 7A–7C) despite the DSLNT conformation being slightly different in each VP1 complex (Fig 7F). Neither the blocking of Neu5Ac₆ binding by E91, nor the increased conformational freedom in LID appears to alter binding affinity. It is possible that Neu5Ac₆ in RA adopts a conformation that might not be favorable and therefore not heavily contribute to affinity, in spite of the added contact surface. Combined with the fact that electron density for Neu5Ac₆ could only be observed in one binding pocket of RA VP1 [12], we believe that this conformation is possible but not probable in solution. Instead, an increased number of conformational options might make up for a loss of binding contacts.

Discussion

Many viruses engage cell-surface glycans to mount an infection, and subtle differences in the recognition of such receptors can be linked with altered tropism and pathogenicity. Examples include the canine parvovirus and feline panleukopenia virus [46,47], the human BK polyomavirus [48], B-lymphotropic polyomavirus [49,50] as well as avian and human influenzaviruses

[51,52]. However, MuPyV is a rare example of a virus in which drastic differences in pathogenicity directly correlate with single amino acid substitutions in the viral capsid.

In order to provide a structural basis for understanding the profoundly different pathogenicities of the three MuPyV strains RA, PTA and LID, we have solved structures of their VP1 proteins and characterized their receptor-binding properties. We show that the ganglioside GT1a serves as a MuPyV receptor and promotes infection with higher potency than the previously identified receptors GD1a and GT1b. Structurally, the increased potency of GT1a can be directly explained by a set of additional contacts involving the $[\alpha\text{-}2,8]$ -linked Neu5Ac_c, that is only present in this glycan and that gives it a characteristic horseshoe-like shape. It had previously been suggested that the G91E mutation present in PTA and LID abolishes binding to branched glycans containing $[\alpha\text{-}2,6]$ -linked Neu5Ac and thus allows the virus to spread more efficiently in the host [8,11]. However, our analyses show that the presence of a glutamate at position 91 still allows binding of the branched oligosaccharides GT1a, GD1a, or DSLNT to all three strains, albeit with subtle differences in binding affinity. While all three strains bind GT1a with comparable affinity, PTA and LID bind GD1a better than RA. The DSLNT glycan binds similarly to all three strains, with the lowest overall affinity. This is again in line with the structures, which show that the branched Neu5Ac_c of DSLNT does not engage in any specific contacts. The limited contacts between Neu5Ac_c and RA observed in an earlier structure [12] have to be considered a crystallization artifact as they were only observed in one out of five binding sites, and this visible Neu5Ac_c moiety was located near a crystal contact.

The ligand binding promiscuity of MuPyV is surprisingly high. Binding mostly requires a ubiquitous minimal Neu5Ac- $[\alpha\text{-}2,3]$ -Gal motif, in agreement with earlier findings [6,7]. It therefore seems plausible that the virus also recognizes other glycans bearing this motif, resulting in differences in pathogenicity and spread. Preliminary studies show that glycans with an N-acetylactose core (Neu5Ac- $[\alpha\text{-}2,3]$ -Gal- $[\beta\text{-}1,4]$ -GlcNAc), as found in neolacto gangliosides such as the predominant ganglioside of peripheral nerve cells, LM1 [53,54], can also be bound in a manner similar to DSLNT and with higher flexibility than GT1a or GD1a (S6 Fig).

Based on our structures, certain requirements that contribute to receptor specificity can be established. For example, branches at Gal-O4 within the minimal motif produce clashes and cannot be tolerated. Therefore, although the GD1a glycan possesses two Neu5Ac- $[\alpha\text{-}2,3]$ -Gal motifs, it prefers the one on its longer arm for complex formation. For the same reason, glycans such as GM1 or GM2 that only possess such a branched Neu5Ac- $[\alpha\text{-}2,3]$ -Gal epitope cannot engage MuPyV productively. In support of this, the GM1 ganglioside is not able to rescue MuPyV infection of Gang^{-/-} MEFs (Fig 2, [29]), although low-level and probably non-specific interactions with cells can be detected (S1B Fig). GT1b possesses a disialylated arm at Gal_b and is monosialylated at Gal_a. We predict that GT1b engages VP1 with its monosialylated arm. The second, disialylated arm is likely to be accommodated in such a binding mode, and the $[\alpha\text{-}2,8]$ -linked sialic acid might contribute additional contacts. Binding via the monosialylated arm is in line with our findings that supplementation of Gang^{-/-} cells with GT1b rescues infection at a level between GD1a and GT1a. Some gangliosides whose glycan epitopes are capable of engaging VP1 *in vitro* might not be infectious receptors *in vivo*, mainly because of steric complications in the context of the cell membrane. For example, while the crystal structure of PTA with the glycan portion of GD3 shows an identical binding mechanism to GT1a (S7 Fig), supplementation of Gang^{-/-} MEFs with GD3 does not restore infectivity [29]. We reason that the glycan stem of GD3 (and of gangliosides with a similar length such as GM3) is too short to allow efficient attachment of the MuPyV capsid to the cell membrane.

The discrepancy in pathogenicity in MuPyV strains that differ only at one single position is stark. In sharp contrast, the differences among receptor binding between the three strains investigated here are subtle, and a correlation of the structural data with the observed

pathogenicity profiles remains challenging. One reason for this is that avidity effects in the virus capsid, which can engage many ligands simultaneously, multiply subtle changes in the affinity of capsomers for single glycans. It was shown for influenza viruses that small changes between millimolar binding affinities of single binding sites can result in dramatically altered viral binding properties [52]. As discussed above, we found the main difference between RA and PTA/LID to be a differing affinity for GD1a, which appears to bind better to the latter strains due to the larger E91 side chain. This might facilitate attachment and productive infection by these strains to cells that display GD1a, and may thus give them an advantage over RA. While we could not show differences between the PTA and LID strain in terms of glycan affinity to isolated VP1 pentamers, it is unclear how this translates to avidity effects. As such, it is possible that capsid avidities differ enough to explain the more limited spread of PTA. Although direct correlations cannot be made, it becomes increasingly clear that the virus needs to uphold a delicate equilibrium between efficient infection and release from infected and lysed cells as well as selective affinity for productive receptors. The absence of the RA and LID strains outside the laboratory [26] emphasizes that this equilibrium is affected by minute changes in the receptor binding properties.

The MuPyV receptor pocket can clearly accommodate several related but distinct glycan structures (Figs 1 and 4–6). These structures also decorate glycoproteins on many cell surfaces. It therefore seems likely that MuPyV can also engage glycans that are not attached to gangliosides. For instance, the glycan stem of GD1 α , which is very similar to DSLNT and prominent on glycoproteins [42–44], is a likely receptor candidate. The different cell-surface distribution patterns of glycoproteins and gangliosides may likewise influence MuPyV spread [8]. Glycoprotein receptors with unknown identity have in fact been shown to promote non-productive internalization of MuPyV, which in turn elicits innate immune responses by the host [29]. Along these lines, our results suggest that virus particles adhere to and enter ganglioside deficient MEFs to levels that are not significantly lower than for wild-type and ganglioside supplemented Gang^{-/-} cells, although without detectable infection. Although not representative for other cell types, these results suggest that the amount of non-productive “pseudoreceptors” on the MEF cell surface is much higher than anticipated.

Our data demonstrate that varying affinities for different gangliosides are the key determinants of a successful MuPyV infection, in line with earlier reports [6–8]. Perhaps unexpectedly, we also find that (even non-specific) attachment of the virus to a host cell can lead to successful internalization, but that this does not necessarily lead to an infection. Thus, we propose that the ratio between productive (ganglioside bound) and non-productive (ganglioside and glycoprotein bound) glycotopes on the host cell itself or in its microenvironment helps to determine the productivity of infection through diverging entry routes, and that differential affinities to these receptors dictate this equilibrium. The nature of these diverging routes, their underlying driving forces, and potential biological consequences other than immune stimulation [29] remain unknown—as does the point at which they diverge. We cannot exclude the possibility that the distribution and binding properties of (pseudo-)receptors are of importance mostly for the post-entry stage rather than for events taking place at the cell surface. A better understanding of the distribution patterns and densities of glycans on specific cells is clearly needed to fully appreciate the many aspects of pathogenesis and tropism of MuPyV as well as many other glycan-binding viruses.

Methods

Ganglioside supplementation and quantification of MuPyV infection

WT and Gang^{-/-} MEFs were seeded onto 96-well Costar 3906 imaging plates in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (FBS). WT (B4+/+St8

+ / +) and Gang^{-/-} MEFs (B4^{-/-}-St8^{-/-}) were provided by Thomas Benjamin at Harvard Medical School. Gangliosides were purchased from Matreya LLC and resuspended in DMSO upon arrival, aliquoted, and stored at -20°C until use. Cells were incubated overnight in serum free media prior to infection. For ganglioside supplemented Gang^{-/-} MEFs, cells were starved in serum free media containing the indicated concentration of ganglioside. Gangliosides were then removed, and cells were washed with serum free media to remove any free ganglioside. Cells were then infected with NG59RA, PTA, and LID MuPyV (MOI ~10–30). At 24 hours post infection cells were washed in phosphate buffered saline and fixed with 4% paraformaldehyde at room temperature for 10 minutes. Cells were then permeabilized with 0.1% Triton X-100, blocked in 10% FBS in PBS, and then stained for the viral protein, T-antigen (E1). Samples were then incubated with Alexa Fluor labeled secondary antibodies (546). Plates were imaged with the Molecular Devices ImageXpress Micro XL High-Content Screener. The percent infected was calculated for each well (5 images were taken per well). Three wells were quantified per sample and the average percent infected, standard error, and standard deviation were calculated for each sample. To quantify infection, T-antigen staining was measured per each DAPI labeled nucleus. For image analysis, the DAPI channel on each image was thresholded, and nuclei were counted using ImageJ (Analyze Particles). These particles were marked as “Regions of Interest” (ROI), and then the average pixel intensity of T-antigen staining was measured for each nucleus (ROI). These were then binned into T-antigen positive or T-antigen negative nuclei to create % infected.

VP1 immunofluorescence staining

WT and Gang^{-/-} MEFs were seeded onto glass coverslips in Dulbecco's Modified Eagle's Medium supplemented with 10% (FBS). Cells were incubated overnight in serum free media prior to infection. For ganglioside supplementation, Gang^{-/-} MEFs were starved in serum free media containing the indicated concentration of ganglioside. Gangliosides were then removed and cells were washed with serum free media to remove any free ganglioside. Cells were then infected with NG59RA. At indicated times post infection the cells were fixed with 4% paraformaldehyde at room temperature. Cells were blocked in 10% FBS in PBS and then stained for GD1a using mAb MAB5606 (Millipore). Cells were then permeabilized with 0.1% Triton X-100 and stained for the viral proteins, VP1 (I58 antibody) and T-antigen (E1 antibody). Samples were washed and then incubated with Alexa Fluor labeled secondary antibodies (488, 546, 647). Slides were then mounted using DAPI prolong gold mounting media. Slides were imaged with a Nikon A1R confocal microscope. All images were taken as a 9 to 13 step (.25µm) z-stacks on a laser scanning confocal microscope. Each z-stack was aligned and compressed into a max intensity Z projection image.

Virus binding to ganglioside supplemented Gang^{-/-} MEFs

WT and Gang^{-/-} MEFs were seeded onto a 24 well dish in Dulbecco's Modified Eagle's Medium supplemented with 10% (FBS). Cells were incubated overnight in serum free media prior to infection. For ganglioside supplemented Gang^{-/-} MEFs, cells were starved in serum free media containing the indicated concentration of ganglioside. Gangliosides were then removed and cells were washed with serum free media to remove any free ganglioside. Cells were then infected with either NG59RA, PTA, or LID at an MOI ~10–30 (250 µL/well). At 4 hours post infection 150 µL of virus supernatant was removed and placed into a microcentrifuge tube. This virus supernatant was then used to infect WT MEFs seeded onto a 96-well plate (50 µL/well). The amount of free virus was then quantified as percent of infection of the 96-well reinfection plate. At 24 hours post virus addition the plate was washed in PBS and

fixed with 4% PFA at RT for 10 minutes. Cells were then permeabilized with 0.1% Triton X-100, blocked in 10% FBS in PBS, and then stained for the viral protein, T-antigen (E1). Samples were then incubated with Alexa Fluor labeled secondary antibodies (546). Plates were imaged with the Molecular Devices ImageXpress Micro XL High-Content Screener. The percent infected was calculated for each well (5 images were taken per well) as indicated by T-antigen positive nuclei. Three wells were quantified per sample and the average percent infected, standard error, and standard deviation were calculated for each sample. For image analysis, the DAPI channel on each image was thresholded and nuclei were counted using ImageJ (Analyze Particles). These particles were marked as “Regions of Interest” (ROI) and then the average pixel intensity of T-antigen staining was measured for each nucleus (ROI). These were then binned into T-antigen positive or T-antigen negative nuclei to create % infected.

Expression and purification of VP1 pentamers

DNA encoding residues 33–316 of RA (GenBank # M34958.1) or PTA VP1 (GenBank # PSU27812) was cloned into the expression vector pET15b (Novagen) in frame with an N-terminal hexahistidine tag (His-tag) and a thrombin cleavage site. DNA for LID VP1 (GenBank # PSU27813) was generated by site-directed mutagenesis of PTA VP1 residue 296. VP1 pentamers were overexpressed in *E. coli* (BL21) after IPTG induction, and purified by nickel affinity chromatography. The His-tag was removed by thrombin cleavage on column for 72 hours (leaving the non-native residues GSHM at the N-terminus), followed by size exclusion chromatography on a Superdex-200 column.

Crystallization and crystal soaking

Pure VP1 pentamers were supplemented with 20 mM DTT, concentrated to 7.5–8 mg/mL (RA VP1) or 8.5–9 mg/mL (PTA and LID VP1), and crystallized by sitting-drop vapor diffusion. RA VP1 was crystallized at 20°C against reservoir solutions containing a range of 1.25–1.8 M ammonium sulfate and 1–10% (*v/v*) isopropanol. PTA and LID were crystallized at 4°C against reservoir solutions containing 0.1 M HEPES pH 7–8.5 and 1.6–1.8 M K-Na phosphate. For complex formation, the crystals were soaked in the reservoir solution supplemented with the glycan. The detailed crystallization and soaking procedures are listed in [S1 Table](#). The GT1a and GD1a glycans were purchased from Elicityl SA (France), and the DSLNT glycan was purchased from Carbosynth (United Kingdom).

For concentration-dependent soaking VP1 pentamers of all three strains were crystallized at 20°C against a mother liquor containing 1.5 M ammonium sulfate and 6% (*v/v*) isopropanol. These crystals were soaked in drops of mother liquor containing the appropriate concentration of glycan for 30 minutes.

All crystals were cryoprotected by incubation in mother liquor supplemented with the appropriate concentration of glycan and 25% (*v/v*) glycerol. They were then flash-frozen in liquid nitrogen.

Structure determination and electron density quantification

Data reduction was carried out in XDS [55], and the structure of native RA VP1 was solved in Molrep [56] using a model generated from the previously solved structure of P16 VP1 (PDB code 1VPN [12]). Other structures were solved by molecular replacement using the RA VP1 structure in Phenix [57]. All structures were completed by alternating rounds of manual model building in Coot [58], followed by restrained coordinate and isomorphous B-factor refinement including TLS refinement and five-fold non-crystallographic symmetry restraints in Refmac5 [59]. TLS parameters were obtained from the TLSMD server [60]. All models agree well with

the experimental data and have good geometry (Table 2). The PDB accession codes for the structures are listed in Table 2. Structural figures were prepared in PyMOL [61].

Data collected for concentration-dependent soaking experiments was processed as described above. The unit cell parameters for all datasets were treated as equal for all datasets and isomorphous to the dataset “RA Nat” (S2 Table). They were scaled against “RA Nat” in Scaleit [62] and then subjected to B-factor refinement and simulated annealing in Phenix against models of RA, PTA, or LID VP1, which lacked atoms of all solvent molecules in the receptor binding pocket as well as those of tryptophan residues 98 and 227 as controls. The resulting bias-reduced $F_{\text{obs}} - F_{\text{calc}}$ electron density for Neu5Ac_a-[α-2,3]-Gal_a and the two marker tryptophans was calculated as a summation of values of the grid points in a mask generated 1 Å around these groups using the program Mapman [63]. The overall binding of a sugar at different concentrations influences the electron density of the Neu5Ac_a-[α-2,3]-Gal_a portion, which is included in GT1a, GD1a, and DSLNT. In contrast, it has no effect on the electron density of the marker tryptophan residues, which do not differ significantly for all data points. For each data point, the average density of the five chains was plotted against ligand concentration and submitted to a non-linear least squares fit using the equation

$$Y = \frac{X}{(K_D + X)} \cdot (B_{\text{max}} - B_0) + B_0 \quad (1)$$

where B_{max} was the highest observed electron density value overall (constrained to 95.03 AU) and B_0 the electron density in the binding pocket at 0 mM ligand concentration. Plotting and fitting was done using the program Prism 6 (GraphPad Software, Inc., La Jolla, California, USA).

Supporting Information

S1 Fig. Gangliosides are required for MuPyV infection, but not for cell surface binding. (A) WT, Gang^{-/-} MEFs, and Gang^{-/-} MEFs supplemented with GD1a were infected with NG59RA MuPyV. The MuPyV ganglioside receptor GD1a can be detected on the WT MEFs and GD1a-supplemented Gang^{-/-} MEFs (green), but is absent in Gang^{-/-} MEFs. Virus binds WT, Gang^{-/-}, and GD1a-supplemented Gang^{-/-} MEFs as shown by VP1 staining (red) on the cell surface at 1 hour post infection. At 24 hours post infection, WT and GD1a-supplemented Gang^{-/-} MEFs show robust infection as indicated by nuclear T-antigen staining (magenta). Despite high levels of virus binding, Gang^{-/-} MEFs are completely resistant to infection as shown by lack of T-antigen staining at 24 hours post infection. (B) Gang^{-/-} MEFs were supplemented with 2 μM GD1a, GT1b, GT1a, GD1b, and GM1 followed by infection with RA, PTA, and LID MuPyV. At 4 hours post infection, virus supernatant was removed and the amount of free virus was quantified for each sample by re-infection of WT MEFs. Virus bound to all cells at similar levels, and there were no significant differences in virus binding to infectious *versus* non-infectious ganglioside receptors. Error bars are standard error, and virus binding to WT MEFs is normalized to one. (TIF)

S2 Fig. CARP Plots of GT1a bound to the PTA VP1 pentamer. The observed phi and psi torsion angles for the linkages occurring in the PTA-GT1a complex have been plotted and compared to other linkages found in the PDB using CARP with the crystallographic definition of torsion angles. The observed linkages are: Neu5Ac-[α-2,3]-Gal (A,B), Gal-[β-1,3]-GalNAc (C), GalNAc-[β-1,4]-Gal (D), and Neu5Ac-[α-2,8]-Neu5Ac (E). The inset on the lower right shows the schematic and observed structure of GT1a. The linkages are named according to the panels;

the coloring of the glycan rings was adopted from Fig 1.
(TIF)

S3 Fig. Ordered water molecules between the two branches of GT1a. Possible hydrogen bonds between the glycan and ordered water molecules are depicted in grey. In GD1a, the glycerol-like tail of Neu5Ac_a could principally also stabilize the glycan, but preferentially adopts a conformation that does not participate in intramolecular contacts.
(TIF)

S4 Fig. B-factor variance across different strains and glycans. The glycans are colored by B factor on an absolute scale between 0 (dark blue) and 80 (deep red). GT1a/PTA is shown as a grey ghost for comparison. For GD1a and DSLNT, the intramolecular B-factor variance when bound to LID is considerably higher, while values for GT1a are comparable to the other strains.
(TIF)

S5 Fig. Mutation to a glutamine at position 91 of PTA VP1 restores the DSLNT binding mode of RA VP1. Shown are the superimposed DSLNT complex structures of RA (PDB-ID 1VPS [12], transparent red) and PTA E91Q (yellow, r.m.s.d. value for the superposition in PyMOL: 0.159 Å). An F_{obs}-F_{calc} omit map (2σ, carved 1.6 Å around the ligand) is shown for the PTA E91Q complex. On the lower right, DSLNT is represented schematically. As for RA VP1, visible electron density for Neu5Ac_c in PTA E91Q can be seen in one of the five chains.
(TIF)

S6 Fig. The reduced van-der-Waals radius at position 296 in LID allows for a more versatile glycan binding. The van-der-Waals radius of 3.5 Å is indicated as dotted sphere for V296 (PTA, yellow) and A296 (LID, blue). The mutation opens the pocket to one side and allows for a more flexible binding mode of glycans without internal stabilization (DSLNT and 3'-N-Acetyl-sialyllactosamine (3'-SLN), the prototype glycan of the LM1 ganglioside). O4 of Gal_a is pointing directly towards the surface in all complexes. In this binding mode, no branching at this point (as is the case e.g. for GD1b) can be tolerated. Glycans that adopt a binding mode similar to the rigid GT1a are colored in green tones, glycans that exhibit shifts of their moieties are colored in shades of red. GT1a bound to PTA is shown as a grey ghost for comparison.
(TIF)

S7 Fig. GD3 binding to VP1. The complex structures of PTA VP1 with GT1a (dark blue) and GD3 (light pink) are superimposed in PyMOL. On the upper right, the structure of GD3 is represented schematically.
(TIF)

S1 Table. Crystallization and soaking conditions.
(DOCX)

S2 Table. Data set statistics for concentration-dependent soaking experiments.
(DOCX)

Acknowledgments

SDO and RLG would like to thank the BioFrontiers Advanced Light Microscopy Core for microscope facility use and imaging support. We thank Luisa Ströh, Georg Zocher, and Bärbel Blaum for helpful discussions. We are also grateful to the staff at the Swiss Light Source (SLS, Villigen, Switzerland), the European Synchrotron Radiation Facility (ESRF, Grenoble, France),

and the Deutsches Elektronen-Synchrotron (DESY, Hamburg, Germany) for beam time and beam line assistance.

Author Contributions

Conceived and designed the experiments: MHCBA AML UN TS SDO RLG. Performed the experiments: MHCBA AML SDO. Analyzed the data: MHCBA AML UN TS SDO RLG. Wrote the paper: MHCBA AML UN TS SDO RLG.

References

1. Ströh LJ, Stehle T (2014) Glycan Engagement by Viruses: Receptor Switches and Specificity. *Annu Rev Virol* 1: 285–306.
2. Matrosovich M, Herrler G, Klenk HD (2015) Sialic Acid Receptors of Viruses. *Top Curr Chem.* 367:1–28 doi: [10.1007/128_2013_466](https://doi.org/10.1007/128_2013_466) PMID: [23873408](https://pubmed.ncbi.nlm.nih.gov/23873408/)
3. Ngamukote S, Yanagisawa M, Ariga T, Ando S, Yu RK (2007) Developmental changes of glycosphingolipids and expression of glycogenes in mouse brains. *J Neurochem* 103: 2327–2341. PMID: [17883393](https://pubmed.ncbi.nlm.nih.gov/17883393/)
4. Yu RK, Macala LJ, Taki T, Weinfield HM, Yu FS (1988) Developmental changes in ganglioside composition and synthesis in embryonic rat brain. *J Neurochem* 50: 1825–1829. PMID: [3131485](https://pubmed.ncbi.nlm.nih.gov/3131485/)
5. Yu RK, Tsai YT, Ariga T, Yanagisawa M (2011) Structures, biosynthesis, and functions of gangliosides—an overview. *J Oleo Sci* 60: 537–544. PMID: [21937853](https://pubmed.ncbi.nlm.nih.gov/21937853/)
6. Cahan LD, Singh R, Paulson JC (1983) Sialyloligosaccharide receptors of binding variants of polyoma virus. *Virology* 130: 281–289. PMID: [6316632](https://pubmed.ncbi.nlm.nih.gov/6316632/)
7. Fried H, Cahan LD, Paulson JC (1981) Polyoma virus recognizes specific sialyloligosaccharide receptors on host cells. *Virology* 109: 188–192. PMID: [6258307](https://pubmed.ncbi.nlm.nih.gov/6258307/)
8. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, et al. (2003) Gangliosides are receptors for murine polyoma virus and SV40. *EMBO J* 22: 4346–4355. PMID: [12941687](https://pubmed.ncbi.nlm.nih.gov/12941687/)
9. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, et al. (1991) Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354: 278–284. PMID: [1659663](https://pubmed.ncbi.nlm.nih.gov/1659663/)
10. Stehle T, Gamblin SJ, Yan Y, Harrison SC (1996) The structure of simian virus 40 refined at 3.1 Å resolution. *Structure* 4: 165–182. PMID: [8805523](https://pubmed.ncbi.nlm.nih.gov/8805523/)
11. Stehle T, Harrison SC (1996) Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. *Structure* 4: 183–194. PMID: [8805524](https://pubmed.ncbi.nlm.nih.gov/8805524/)
12. Stehle T, Harrison SC (1997) High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J* 16: 5139–5148. PMID: [9305654](https://pubmed.ncbi.nlm.nih.gov/9305654/)
13. Stehle T, Yan Y, Benjamin TL, Harrison SC (1994) Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* 369: 160–163. PMID: [8177322](https://pubmed.ncbi.nlm.nih.gov/8177322/)
14. Feunteun J, Sompayrac L, Fluck M, Benjamin T (1976) Localization of gene functions in polyoma virus DNA. *Proc Natl Acad Sci U S A* 73: 4169–4173. PMID: [186787](https://pubmed.ncbi.nlm.nih.gov/186787/)
15. Dawe CJ, Law LW, Dunn TB (1959) Studies of parotid-tumor agent in cultures of leukemic tissues of mice. *J Natl Cancer Inst* 23: 717–797. PMID: [13814555](https://pubmed.ncbi.nlm.nih.gov/13814555/)
16. Gross L (1951) "Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK-leukemic extracts, or AK-embryos. *Proc Soc Exp Biol Med* 76: 27–32. PMID: [14816382](https://pubmed.ncbi.nlm.nih.gov/14816382/)
17. Main JH, Dawe CJ (1966) Tumor induction in transplanted tooth buds infected with polyoma virus. *J Natl Cancer Inst* 36: 1121–1136. PMID: [4287623](https://pubmed.ncbi.nlm.nih.gov/4287623/)
18. Rowe WP, Hartley JW, Estes JD, Huebner RJ (1959) Studies of mouse polyoma virus infection. 1. Procedures for quantitation and detection of virus. *J Exp Med* 109: 379–391. PMID: [13641563](https://pubmed.ncbi.nlm.nih.gov/13641563/)
19. Bolen JB, Fisher SE, Chowdhury K, Shan TC, Williams JE, et al. (1985) A determinant of polyomavirus virulence enhances virus growth in cells of renal origin. *J Virol* 53: 335–339. PMID: [2981359](https://pubmed.ncbi.nlm.nih.gov/2981359/)
20. Bauer PH, Bronson RT, Fung SC, Freund R, Stehle T, et al. (1995) Genetic and structural analysis of a virulence determinant in polyomavirus VP1. *J Virol* 69: 7925–7931. PMID: [7494305](https://pubmed.ncbi.nlm.nih.gov/7494305/)
21. Dawe CJ, Freund R, Mandel G, Ballmer-Hofer K, Talmage DA, et al. (1987) Variations in polyoma virus genotype in relation to tumor induction in mice. Characterization of wild type strains with widely differing tumor profiles. *The American journal of pathology* 127: 243–261. PMID: [2437801](https://pubmed.ncbi.nlm.nih.gov/2437801/)
22. Freund R, Calderone A, Dawe CJ, Benjamin TL (1991) Polyomavirus tumor induction in mice: effects of polymorphisms of VP1 and large T antigen. *J Virol* 65: 335–341. PMID: [1845894](https://pubmed.ncbi.nlm.nih.gov/1845894/)

23. Freund R, Dawe CJ, Benjamin TL (1988) The middle T proteins of high and low tumor strains of polyomavirus function equivalently in tumor induction. *Virology* 167: 657–659. PMID: [2849243](#)
24. Freund R, Garcea RL, Sahli R, Benjamin TL (1991) A single-amino-acid substitution in polyomavirus VP1 correlates with plaque size and hemagglutination behavior. *Journal of virology* 65: 350–355. PMID: [1845896](#)
25. Bauer PH, Cui C, Liu WR, Stehle T, Harrison SC, et al. (1999) Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse. *J Virol* 73: 5826–5832. PMID: [10364334](#)
26. Carroll J, Dey D, Kreisman L, Velupillai P, Dahl J, et al. (2007) Receptor-binding and oncogenic properties of polyoma viruses isolated from feral mice. *PLoS Path* 3: e179.
27. Gilbert J, Benjamin T (2004) Uptake pathway of polyomavirus via ganglioside GD1a. *J Virol* 78: 12259–12267. PMID: [15507613](#)
28. O'Hara SD, Stehle T, Garcea R (2014) Glycan receptors of the Polyomaviridae: structure, function, and pathogenesis. *Curr Opin Virol* 7: 73–78. doi: [10.1016/j.coviro.2014.05.004](#) PMID: [24983512](#)
29. You J, O'Hara SD, Velupillai P, Castle S, Lavery S, et al. (2015) Ganglioside and non-Ganglioside Receptors for the Mouse Polyomavirus. *PLoS Path* 11: e1005175.
30. Low JA, Magnuson B, Tsai B, Imperiale MJ (2006) Identification of gangliosides GD1b and GT1b as receptors for BK virus. *J Virol* 80: 1361–1366. PMID: [16415013](#)
31. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797. PMID: [17681537](#)
32. Blaum BS, Hannan JP, Herbert AP, Kavanagh D, Uhrin D, et al. (2015) Structural basis for sialic acid-mediated self-recognition by complement factor H. *Nat Chem Biol* 11: 77–82. doi: [10.1038/nchembio.1696](#) PMID: [25402769](#)
33. Chen Y, Tan M, Xia M, Hao N, Zhang XC, et al. (2011) Crystallography of a Lewis-binding norovirus, elucidation of strain-specificity to the polymorphic human histo-blood group antigens. *PLoS Path* 7: e1002152.
34. Neu U, Hengel H, Blaum BS, Schowalter RM, Macejak D, et al. (2012) Structures of Merkel cell polyomavirus VP1 complexes define a sialic acid binding site required for infection. *PLoS Path* 8: e1002738.
35. Merritt EA, Kuhn P, Sarfaty S, Erbe JL, Holmes RK, et al. (1998) The 1.25 Å resolution refinement of the cholera toxin B-pentamer: evidence of peptide backbone strain at the receptor-binding site. *J Mol Biol* 282: 1043–1059. PMID: [9753553](#)
36. Neu U, Woellner K, Gauglitz G, Stehle T (2008) Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc Natl Acad Sci U S A* 105: 5219–5224. doi: [10.1073/pnas.0710301105](#) PMID: [18353982](#)
37. Cubrilovic D, Haap W, Barylyuk K, Ruf A, Badertscher M, et al. (2014) Determination of protein-ligand binding constants of a cooperatively regulated tetrameric enzyme using electrospray mass spectrometry. *ACS Chem Biol* 9: 218–226. doi: [10.1021/cb4007002](#) PMID: [24128068](#)
38. Fotinou C, Emsley P, Black I, Ando H, Ishida H, et al. (2001) The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin. *J Biol Chem* 276: 32274–32281. PMID: [11418600](#)
39. Sharmila DJ, Veluraja K (2006) Conformations of higher gangliosides and their binding with cholera toxin—investigation by molecular modeling, molecular mechanics, and molecular dynamics. *J Biomol Struct Dyn* 23: 641–656. PMID: [16615810](#)
40. Zhang Y, Yamamoto S, Yamaguchi T, Kato K (2012) Application of paramagnetic NMR-validated molecular dynamics simulation to the analysis of a conformational ensemble of a branched oligosaccharide. *Molecules* 17: 6658–6671. doi: [10.3390/molecules17066658](#) PMID: [22728360](#)
41. Kuan CT, Chang J, Mansson JE, Li J, Pegram C, et al. (2010) Multiple phenotypic changes in mice after knockout of the B3gnt5 gene, encoding Lc3 synthase—a key enzyme in lacto-neolacto ganglioside synthesis. *BMC Dev Biol* 10: 114. doi: [10.1186/1471-213X-10-114](#) PMID: [21087515](#)
42. Parkkinen J, Rogers GN, Korhonen T, Dahr W, Finne J (1986) Identification of the O-linked sialyloligosaccharides of glycophorin A as the erythrocyte receptors for S-fimbriated *Escherichia coli*. *Infect Immun* 54: 37–42. PMID: [2875951](#)
43. Schauer R (2000) Achievements and challenges of sialic acid research. *Glycoconj J* 17: 485–499. PMID: [11421344](#)
44. Thomsson KA, Holmen-Larsson JM, Angstrom J, Johansson ME, Xia L, et al. (2012) Detailed O-glycomics of the Muc2 mucin from colon of wild-type, core 1- and core 3-transferase-deficient mice highlights differences compared with human MUC2. *Glycobiology* 22: 1128–1139. doi: [10.1093/glycob/cws083](#) PMID: [22581805](#)

45. Burmeister WP, Guilligay D, Cusack S, Wadell G, Arnberg N (2004) Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J Virol* 78: 7727–7736. PMID: [15220447](#)
46. Parrish CR (1990) Emergence, natural history, and variation of canine, mink, and feline parvoviruses. *Adv Virus Res* 38: 403–450. PMID: [2171302](#)
47. Parrish CR (1991) Mapping specific functions in the capsid structure of canine parvovirus and feline panleukopenia virus using infectious plasmid clones. *Virology* 183: 195–205. PMID: [1647068](#)
48. Neu U, Allen SA, Blaum BS, Liu Y, Frank M, et al. (2013) A structure-guided mutation in the major capsid protein retargets BK polyomavirus. *PLoS Path* 9: e1003688.
49. Neu U, Khan ZM, Schuch B, Palma AS, Liu Y, et al. (2013) Structures of B-lymphotropic polyomavirus VP1 in complex with oligosaccharide ligands. *PLoS Path* 9: e1003714.
50. Kanda T, Furuno A, Yoshiike K (1986) Mutation in the VP-1 gene is responsible for the extended host range of a monkey B-lymphotropic papovavirus mutant capable of growing in T-lymphoblastoid cells. *J Virol* 59: 531–534. PMID: [3488416](#)
51. Xiong X, Martin SR, Haire LF, Wharton SA, Daniels RS, et al. (2013) Receptor binding by an H7N9 influenza virus from humans. *Nature* 499: 496–499. doi: [10.1038/nature12372](#) PMID: [23787694](#)
52. Xiong X, Coombs PJ, Martin SR, Liu J, Xiao H, et al. (2013) Receptor binding by a ferret-transmissible H5 avian influenza virus. *Nature* 497: 392–396. doi: [10.1038/nature12144](#) PMID: [23615615](#)
53. Ogawa-Goto K, Funamoto N, Ohta Y, Abe T, Nagashima K (1992) Myelin gangliosides of human peripheral nervous system: an enrichment of GM1 in the motor nerve myelin isolated from cauda equina. *J Neurochem* 59: 1844–1849. PMID: [1402926](#)
54. Ishikawa Y, Gasa S, Minami R, Makita A (1987) Characterization of neutral glycosphingolipids from fetal human brain: evidence for stage-specific expression of the globo, ganglio, and neolacto series in the central nervous system. *J Biochem* 101: 1369–1375. PMID: [3667553](#)
55. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66: 133–144. doi: [10.1107/S0907444909047374](#) PMID: [20124693](#)
56. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66: 22–25. doi: [10.1107/S0907444909042589](#) PMID: [20057045](#)
57. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213–221. doi: [10.1107/S0907444909052925](#) PMID: [20124702](#)
58. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132. PMID: [15572765](#)
59. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240–255. PMID: [15299926](#)
60. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62: 439–450. PMID: [16552146](#)
61. The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC.
62. Howell PL, Smith GD (1992) Identification of heavy-atom derivatives by normal probability methods. *J Appl Cryst* 25: 81–86.
63. Kleywegt GJ, Jones TA (1996) xDIMPAN and xDATAMAN—Programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallogr D Biol Crystallogr* 52: 826–828. PMID: [15299647](#)
64. Kato Y, Kuan CT, Chang J, Kaneko MK, Ayriss J, et al. (2010) GMab-1, a high-affinity anti-3'-isoLM1/3',6'-isoLD1 IgG monoclonal antibody, raised in lacto-series ganglioside-defective knockout mice. *Biochem Biophys Res Commun* 391: 750–755. doi: [10.1016/j.bbrc.2009.11.132](#) PMID: [19944071](#)
65. Svennerholm L (1969). *Comprehensive Biochemistry*.
66. Taube S, Jiang M, Wobus CE (2010) Glycosphingolipids as receptors for non-enveloped viruses. *Viruses* 2: 1011–1049. doi: [10.3390/v2041011](#) PMID: [21994669](#)
67. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27–30. PMID: [10592173](#)

Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site

Luisa J. Ströh,^a Ursula Neu,^{a*} Bärbel S. Blaum,^a Michael H. C. Buch,^a Robert L. Garcea,^b Thilo Stehle^{a,c}

Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany^a; Department of Molecular, Cellular and Developmental Biology and the BioFrontiers Institute, University of Colorado, Boulder, Colorado, USA^b; Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA^c

ABSTRACT

Human polyomavirus 6 (HPyV6) and HPyV7 are commonly found on human skin. We have determined the X-ray structures of their major capsid protein, VP1, at resolutions of 1.8 and 1.7 Å, respectively. In polyomaviruses, VP1 commonly determines antigenicity as well as cell-surface receptor specificity, and the protein is therefore linked to attachment, tropism, and ultimately, viral pathogenicity. The structures of HPyV6 and HPyV7 VP1 reveal uniquely elongated loops that cover the bulk of the outer virion surfaces, obstructing a groove that binds sialylated glycan receptors in many other polyomaviruses. In support of this structural observation, interactions of VP1 with α 2,3- and α 2,6-linked sialic acids could not be detected in solution by nuclear magnetic resonance spectroscopy. Single-cell binding studies indicate that sialylated glycans are likely not required for initial attachment to cultured human cells. Our findings establish distinct antigenic properties of HPyV6 and HPyV7 capsids and indicate that these two viruses engage nonsialylated receptors.

IMPORTANCE

Eleven new human polyomaviruses, including the skin viruses HPyV6 and HPyV7, have been identified during the last decade. In contrast to better-studied polyomaviruses, the routes of infection, cell tropism, and entry pathways of many of these new viruses remain largely mysterious. Our high-resolution X-ray structures of major capsid proteins VP1 from HPyV6 and from HPyV7 reveal critical differences in surface morphology from those of all other known polyomavirus structures. A groove that engages specific sialic acid-containing glycan receptors in related polyomaviruses is obstructed, and VP1 of HPyV6 and HPyV7 does not interact with sialylated compounds in solution or on cultured human cells. A comprehensive comparison with other structurally characterized polyomavirus VP1 proteins enhances our understanding of molecular determinants that underlie receptor specificity, antigenicity, and, ultimately, pathogenicity within the polyomavirus family and highlight the need for structure-based analysis to better define phylogenetic relationships within the growing polyomavirus family and perhaps also for other viruses.

Polyomaviruses are a group of nonenveloped double-stranded DNA (dsDNA) viruses that were initially identified in mice (1) but have been found since then in birds and in several species of mammals, including humans (reviewed in references 2, 3, and 4). Due to recently developed techniques such as high-throughput sequencing and rolling-circle amplification, a number of new polyomaviruses, including 11 new human polyomaviruses, have been identified during the last decade (5–14). This expansion of the *Polyomaviridae* family led to its division into three genera, the ortho-, wuki-, and avipolyomaviruses (15). With high sequence homology and conserved overall architecture across genera, the family forms an attractive platform for analyzing determinants of cell entry, cell tropism, and host range as well as other factors that contribute to pathogenesis (4). Asymptomatic and latent infections with polyomaviruses are common in the healthy human population (16–18). Individuals with impaired immune responses due to organ transplantation, monoclonal antibody therapy, hematological diseases, or human immunodeficiency virus (HIV) infection are found to be especially susceptible to reactivation of polyomaviruses, which can lead to severe or fatal diseases (reviewed in references 4 and 19).

Cutaneous human polyomavirus 6 (HPyV6) and HPyV7 (8) are commonly shed from the skin together with the oncogenic Merkel cell polyomavirus (MCPyV) (8, 20). Although no human

disease has been linked to HPyV6 and HPyV7 so far, initial studies indicate that persistent infections with both viruses are very common, resulting in seropositivity rates of 35% to 90% by adulthood (8, 17). Thus, an involvement of HPyV6 and/or HPyV7 in cutaneous tumors has to be considered (21–24). While the two viruses have the same tropism as MCPyV, they are more closely related in sequence to the Washington University and Karolinska Institute polyomaviruses (WUPyV and KIPyV, respectively), which were isolated from respiratory tract samples (5, 6). Hence, WUPyV, KIPyV, HPyV6, and HPyV7 have been classified together as wuki-polyomaviruses (15).

The polyomavirus major capsid protein VP1 determines antigenicity and mediates attachment to host-cell receptors. It is well

Received 17 April 2014 Accepted 1 July 2014

Published ahead of print 9 July 2014

Editor: M. J. Imperiale

Address correspondence to Thilo Stehle, thilo.stehle@uni-tuebingen.de.

* Present address: Ursula Neu, National Institute of Medical Research, London, United Kingdom.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.01084-14

established that VP1 possesses a jelly-roll topology and assembles into 72 pentamers, which in turn form a $T = 7d$ icosahedral capsid (25, 26). The VP1 pentamers associate with minor capsid proteins VP2 and VP3, which are located inside the capsid. The presence and roles of VP2 and VP3 seem to differ among polyomavirus species (27). All known structures of polyomavirus VP1 show extended and structurally variable surface loops that emanate from a conserved β -sheet core structure formed by strands B, I, D, and G and strands C, H, E, and F (25, 26, 28–35). These surface loops, named BC-, DE-, HI-, and EF-loops after the β -strands connected by them, are chiefly responsible for viral antigenicity. For each virus, they form a unique virus-host interaction platform that determines host range, cell tropism, viral spread, and pathogenicity. Engagement of sialylated glycan motifs during cell attachment and entry is a common feature of the better-studied orthopolyomaviruses (28–30, 32–36). In contrast, the routes of infection, transmission, cell tropism, receptor specificity, and entry pathways of wikipolyomaviruses remain largely mysterious. To provide an initial framework for investigating the molecular determinants of receptor specificity and tropism of HPyV6 and HPyV7, we determined high-resolution crystal structures of their recombinantly expressed VP1 proteins. While the core structures are highly conserved, the surface loops of HPyV6 and HPyV7 VP1 differ profoundly from those of related polyomavirus VP1 structures. Specific cell surface receptors remain to be unveiled for both viruses, but, interestingly, our findings indicate that sialylated glycans are likely not engaged during early infection steps. In support of the crystallographic analyses, interactions of VP1 with either α 2,3- or α 2,6-linked sialic acids could not be detected by flow cytometry cell binding studies and saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectroscopy.

MATERIALS AND METHODS

DNA cloning and protein expression. Following a strategy established for the expression of soluble, assembly-incompetent VP1 pentamers (36), DNAs coding for residues 20 to 291 of HPyV6 VP1 and residues 20 to 288 of HPyV7 VP1 (GenBank accession codes ADE45449 and ADE45474) were cloned into pET15b vectors (Novagen). Soluble pentamers were expressed and purified as described earlier (30, 33). Four nonnative residues (GSHM) are present at the N termini of both proteins after purification, and a nonnative glutamine forms the C terminus of HPyV7 VP1. For cell binding experiments, the JC polyomavirus (JCPyV) VP1 wild type and L54F mutant (residues 22 to 289) and murine polyomavirus VP1 (RA strain; residues 33 to 316) were expressed and purified accordingly (30, 37).

Crystallization, data collection, and structure determination. HPyV6 VP1 was concentrated to 6 mg/ml in 20 mM HEPES (pH 7.5)–150 mM NaCl–20 mM dithiothreitol (DTT), and crystals were obtained at 20°C using the sitting-drop vapor diffusion technique and a reservoir containing 100 mM Tris (pH 8.5)–200 mM NaSCN (sodium thiocyanate)–13.3% (wt/vol) polyethylene glycol (PEG) 3350. HPyV7 VP1 was crystallized using 7 mg/ml VP1–20 mM HEPES (pH 7.5)–150 mM NaCl and a reservoir containing 100 mM Na malonate (pH 4.5)–15% (wt/vol) PEG 3350 (hanging-drop vapor diffusion technique). Drops were set up with 1 μ l protein solution and 1 μ l reservoir solution in each case. Crystals were harvested into the respective reservoir solutions supplemented with 30% (vol/vol) glycerol prior to flash-freezing them in liquid nitrogen. Diffraction data were collected at beamlines ID14-4 at ESRF (Grenoble, France) (HPyV6) and X06DA at SLS (Villigen, Switzerland) (HPyV7). Data sets were processed with XDS (38), and structures were solved by molecular replacement with Phaser in CCP4 (39, 40). The WUPyV VP1 core structure (PDB accession no. 3S7X) (31) served as a search model to

solve the HPyV7 VP1 structure, and the refined HPyV7 VP1 coordinates were then used to determine the structure of HPyV6 VP1. Rigid-body and simulated annealing refinement was in both cases carried out with Phenix (41), followed by alternating rounds of model building in Coot (42) and restrained refinement, including the translation-libration-screw method (43) and 5-fold noncrystallographic symmetry restraints with Refmac5 (44). Structural superpositions were done using secondary-structure matching (45) and the program Superpose in CCP4 (40). PyMOL (The PyMOL Molecular Graphics System, Version 1.3; Schrödinger, LLC) and the APBS tool plugin (46) were used to create structure figures.

Cell culture. HeLa S3 and 293TT (47) cells were maintained in a humidified 37°C CO₂ chamber in Dulbecco's modified Eagle medium (DMEM) supplemented with 1% penicillin-streptomycin, 6 mM L-glutamine, 1 mM sodium pyruvate, and 10% heat-inactivated fetal bovine serum (FBS). HeLa S3 cells were kindly provided by Katharina Rehn and Dirk Schwarzer (IFIB, Tübingen, Germany) and 293TT cells by Christopher C. Buck (National Cancer Institute, NIH).

Flow cytometry experiments. VP1 pentamers were labeled with Alexa Fluor 488 C5 maleimide (Invitrogen). Proteins (1 mg/ml) were incubated in 20 mM HEPES (pH 7.5)–150 mM NaCl for 16 to 18 h with 80 mM DTT at 4°C. Excess DTT was removed using two 5-ml HiTrap desalting columns (GE Healthcare), and the dye (10 mM in 20 mM HEPES pH 7.0–150 mM NaCl) was added dropwise by gently mixing it into the protein solution (0.3 to 0.4 mg/ml in 20 mM HEPES pH 7.0–150 mM NaCl) to give a 8-fold molar excess of the dye. The reaction mixture was incubated for 18 h at 4°C. DTT (10 mM) was added, and the excess dye and DTT were removed by desalting (two 5-ml HiTrap desalting columns). The labeling efficiency was determined by UV light/visible light (UV/vis) absorption according to the manufacturer's protocol.

Cells (80% to 90% confluent) were detached nonenzymatically from flasks by incubation with Gibco enzyme-free cell dissociation buffer (Life Technologies) for 30 min at 37°C and were washed twice with phosphate-buffered saline (PBS). A total of 5×10^6 cells were suspended in 100 μ l PBS. Cells were mock treated or pretreated with 0.2 U/ml neuraminidase (*Clostridium perfringens* neuraminidase type V; Sigma-Aldrich) at 37°C for 30 min and then washed 3 times with 500 μ l PBS and pelleted after each wash at $200 \times g$ (4 min; 4°C). Cells were incubated in 100 μ l of labeled VP1 pentamer solution (50 μ g/ml in PBS) on ice for 2 h with agitation every 15 to 20 min. HeLa S3 cells were washed twice in 500 μ l PBS and fixed in 500 μ l PBS containing 1% formaldehyde for 30 min. 293TT cells were washed twice in PBS and then suspended in 500 μ l PBS for the measurement. DAPI (4',6'-diamidino-2-phenylindole) was added to 293TT cells to gate for live cells. Analysis was done using a BD FAC-SCanto (Becton, Dickinson and Company) flow cytometer equipped with a 488-nm excitation line. A total of 10,000 gated events were measured for each sample. Data were analyzed using FlowJo software (Tree Star, Inc.).

Saturation transfer difference NMR spectroscopy. STD NMR spectra were recorded at 283 K using 3-mm-inner-diameter Match tubes (200- μ l sample volume) and a Bruker AVIII-600 spectrometer equipped with a room temperature probe head and processed with TopSpin 3.0 (Bruker). Samples contained 1 mM α 2,3-sialyllactose and α 2,6-sialyllactose (Carbosynth) (each) and a 50 μ M concentration of either HPyV6 or HPyV7 VP1. Proteins were buffer exchanged prior to NMR experiments in centrifugal concentrators to 20 mM potassium phosphate (pH 7.4)–150 mM NaCl in D₂O. Off- and on-resonance irradiation frequencies were set to –80 ppm and 7.0 ppm, respectively. The irradiation power of the selective pulses was 57 Hz, the saturation time was 2 s, and the total relaxation delay was 3 s. A 50-ms continuous-wave spin-lock pulse with a strength of 3.2 kHz was employed to suppress residual protein signals. A total of 512 scans were recorded. A total of 10,000 points were collected, and spectra were multiplied with an exponential window function (line broadening, 1 Hz) prior to Fourier transformation. Spectra were referenced to 298 K using the α -D-Glc anomeric proton as an internal standard (48).

TABLE 1 Data collection and structural refinement statistics^a

Parameter	HPyV6 VP1	HPyV7 VP1
Space group	C2	C2
Unit cell dimensions		
a, b, c (Å)	183.9, 89.4, 125.3	209.7, 86.4, 84.2
α, β, γ (°)	90.0, 131.3, 90.0	90.0, 92.1, 90.0
Data collection		
Resolution (Å)	30.0–1.8 (1.85–1.80)	30.0–1.70 (1.75–1.70)
No. of unique reflections	140,450 (10,046)	161,009 (9,962)
Redundancy	4.1 (4.1)	3.9 (3.9)
<i>R</i> _{meas} (%)	7.0 (56.8)	6.7 (55.9)
<i>I</i> /σ	14.8 (2.8)	15.7 (2.4)
CC _{1/2} (%)	99.8 (84.7)	99.9 (75.4)
Completeness (%)	99.2 (96.4)	97.8 (82.1)
Wilson B-factor (Å ²)	28.9	26.0
Refinement		
<i>R</i> _{work} / <i>R</i> _{free} (%)	14.8/17.4	16.9/19.6
No. of atoms		
Protein	10,171	9,856
Water	1,070	990
Average B-factor (Å ²)		
Protein	28.7	24.8
Water	34.2	29.0
RMSDs		
Bond lengths (Å)	0.009	0.010
Bond angles (°)	1.381	1.391
Ramachandran plot (calculated using MolProbity Server)		
Favored (%)	95.9	97.5
Allowed (%)	4.1	2.4
Outliers (%)	0.0	0.1

^a Values for the highest-resolution bin are given in parentheses. CC_{1/2}, correlation between intensities from random half-data sets.

$$R_{\text{meas}} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl,j} \rangle|}{\sum_{hkl} \sum_j^n I_{hkl,j}}$$

where *n* is the number of observations of the reflection and $\langle I_{hkl,j} \rangle$ the intensity of symmetry (or Friedel)-related observations.

$$R_{\text{work}} = \frac{\sum_{hkl} |F_{\text{obs}}(hkl) - F_{\text{calc}}(hkl)|}{\sum_{hkl} F_{\text{obs}}(hkl)}$$

where *F*_{obs} and *F*_{calc} are the observed and calculated structure factors, respectively. A total of 5% of the reflections were not used during structure refinement to calculate *R*_{free}.

Protein structure accession numbers. Coordinates and structure factor amplitudes have been deposited under accession numbers 4PCG (HPyV6) and 4PCH (HPyV7) with the RCSB Protein Data Bank (www.rcsb.org).

RESULTS

Overall structures of HPyV6 and HPyV7 VP1. The HPyV6 and HPyV7 VP1 pentamer structures were solved at resolutions of 1.8 and 1.7 Å, respectively, and the refined structures have excellent statistics (Table 1). The final coordinates include residues 22 to 87 and 94 to 290 in the case of HPyV6 and residues 23 to 47, 55 to 85, and 96 to 286 of HPyV7 for all five VP1 chains in the asymmetric units. Each VP1 monomer within the pentamer adopts the iconic jelly-roll fold consisting of two apposed β-sheets (Fig. 1). The EF-loops fold into short three-stranded β-sheets (E', E'', and E''')

and decorate the side of the pentamer. The I-strand is split into two parts named I and I'. As is typical for polyomavirus VP1 structures, rather poor electron density was observed for the CD-loops at the base of the HPyV6 and HPyV7 pentamers (25, 26, 28–32). This loop is flexible and assumes different conformations even in the context of the intact virion (25, 26, 28). The long BC-loop is divided for clarity into BC1- and BC2-loops that face in different directions (Fig. 1C). The BC1-loops of the HPyV7 VP1 pentamer have elevated mobility, and they have continuous electron density only when contacting the neighboring protomer within the crystal lattice. Thus, the final coordinates contain only three ordered BC1-loops per pentamer. In contrast, the BC1-loops of HPyV6 VP1 have good electron density and share similar conformations that are independent of the presence of crystal contacts. They could therefore be built for all five chains in the asymmetric unit.

The Cα atoms of monomeric and pentameric HPyV6 and HPyV7 VP1 structures superpose with very low root mean square deviation (RMSD) values of 0.6 Å and 0.7 Å, respectively, reflecting their high sequence identity of 68% (49). RMSD values for individual residues exceed 1.5 Å only within the EF-loop, where three additional residues elongate the HPyV6 VP1 EF-loop somewhat so that it projects further away from the 5-fold axis (Fig. 1C).

In order to quantify the level of structural diversity of VP1 structures within the members of the *Polyomaviridae* family, the HPyV6 and HPyV7 VP1 coordinates were superposed with the most closely related WUPyV (PDB accession no. 3S7X) and KIPyV (PDB accession no. 3S7V) VP1 structures as well as the evolutionarily more distant MCPyV (PDB accession no. 4FMG) and simian virus 40 (SV40) (PDB accession no. 3BWQ) VP1 structures (15). Root mean square deviation (RMSD) values for superpositions of VP1 monomers from HPyV6 or HPyV7 onto their KIPyV and WUPyV counterparts are low (~1.2 Å), in line with the classification of these four viruses into the wukipolyomavirus family. Superpositions of HPyV6 and HPyV7 VP1 onto VP1 monomers from the orthopolyomaviruses MCPyV and SV40 yield slightly higher RMSD values (1.4 to 1.7 Å). However, when entire VP1 pentamers are superposed, the RMSD differences all lie in a range from 1.3 Å (HPyV6-KIPyV) to 1.7 Å (HPyV6-SV40), demonstrating that the pentameric VP1 arrangements are similar across the wuki- and orthopolyomaviruses.

Organization of surface loops. The top surface of the VP1 pentamer, which corresponds to the accessible surface of the virus, is almost entirely defined by the BC-, DE-, and HI-loops, and these loops endow each polyomavirus with a unique platform for specific interactions with individual receptors (Fig. 1C). Typically, this platform binds glycan receptors that terminate in sialic acid (Neu5Ac), but sequences and linkages of the recognized oligosaccharides differ among polyomaviruses, leading to specific interactions with a small subset of sialylated glycans in each case (28–30, 32–36). However, despite these differences, the location of the sialic acid binding site is conserved among the polyomaviruses for which VP1 structures have been available to date. The sialic acid binding site is typically located in a recessed area at the junction of the BC1-, BC2-, DE-, and HI-loops of a VP1 monomer, and additional contacts are contributed by the BC2- and DE-loops of counterclockwise (ccw) and clockwise (cw) neighboring monomers, respectively (28–30, 32–35). In order to assess the ability of the HPyV6 and HPyV7 surface loops to form such a sialic acid binding site, we compared the conformations and lengths of their

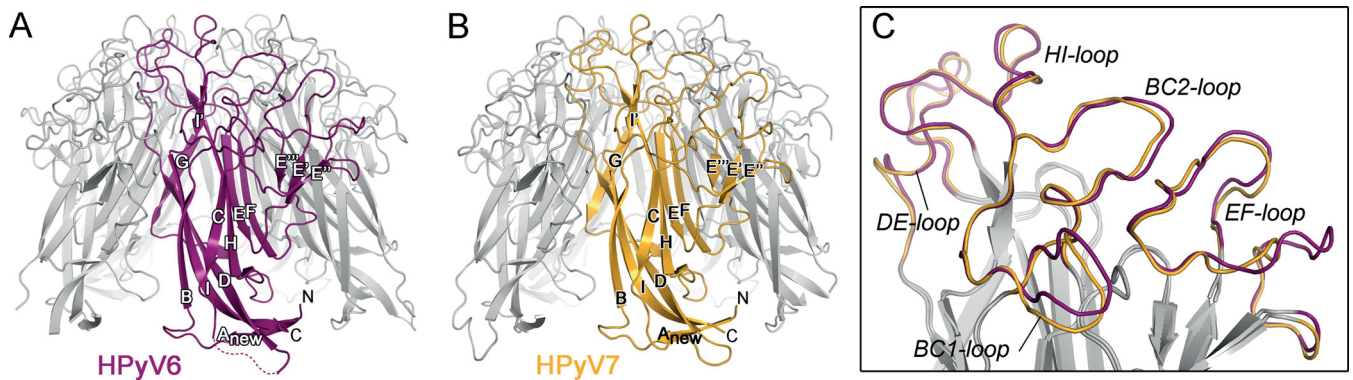


FIG 1 Architecture of HPyV6 and HPyV7 VP1 pentamers. (A and B) Overall folds of HPyV6 (A) and HPyV7 (B) VP1 pentamers shown in a ribbon representation, with VP1 monomers highlighted in magenta and gold, respectively. A dashed line represents the missing HPyV6 CD-loop residues, which are defined only by rather poor electron density due to structural flexibility. (C) Closeup view of the surface loop architectures of HPyV6 and HPyV7 VP1 monomers. The loops are colored as described for panels A and B, respectively. VP1 monomer structures were superposed using the secondary-structure-matching (SSM) superposition tool (45) in the program Coot (42).

surface-exposed loops with the equivalent loops in WUPyV and KIPyV, as well as in MCPyV and SV40. For clarity, only comparisons with the closely related KIPyV and the sialic acid-engaging SV40 polyomavirus are shown in Fig. 2. As the HPyV6 and HPyV7 loop structures are very similar (Fig. 1C), only HPyV6 VP1 is discussed and shown. The HI-loop of polyomavirus VP1 is only a short hairpin in all structures crystallized to date, and it typically forms a wall that closes the glycan-receptor binding site at one end. Strikingly, this loop is extended by 14 residues in HPyV6 compared to KIPyV and WUPyV and by 9 residues compared to SV40 and MCPyV (Fig. 2). Rather than projecting outward, the HPyV6 HI-loop folds on top of the pentamer and forms extensive

contacts within itself and with the DE-loop (not shown in detail). To accommodate the extended HI-loop, the DE-loop of the ccw monomer is displaced toward the 5-fold symmetry axis. The BC2-loop of HPyV6 is severely truncated in comparison to other VP1 structures and lies flat on the VP1 surface (Fig. 2B). Taking these data together, the elongated HI-loops and the truncated BC2-loops of HPyV6 (and also HPyV7) lead to a profoundly altered surface loop network.

The HI-loop participates in the recognition of sialylated glycan receptors in all sialic acid binding polyomaviruses whose structures have been determined to date (28–30, 32–35), by contributing parts of the shallow receptor binding groove on the protein

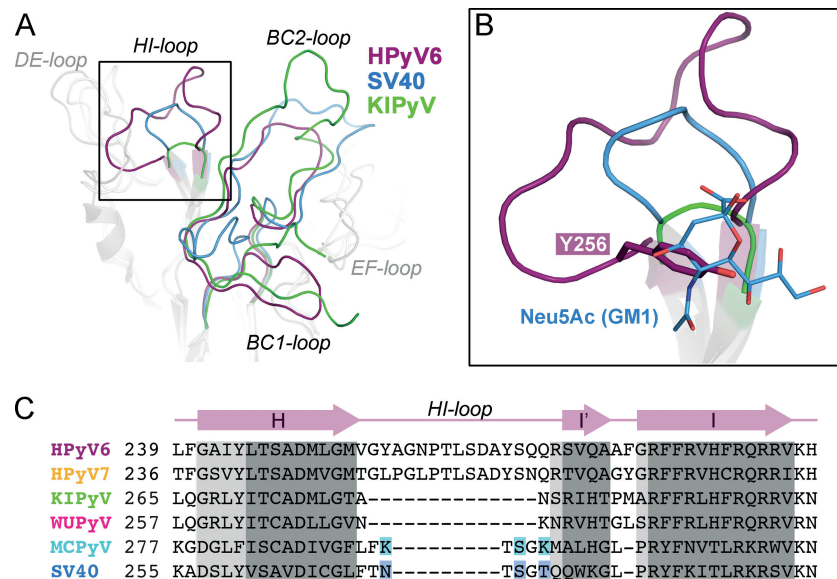


FIG 2 Architecture of VP1 surface loops. Superposition and comparison of VP1 surface loop structures are shown for wikipolyomaviruses HPyV6 and KIPyV (PDB accession no. 3S7V) and orthopolyomavirus SV40 (PDB accession no. 3BWR). VP1 monomers were superposed using the secondary-structure-matching (SSM) superposition tool (45). (A) HI-, BC1-, and BC2-loops are highlighted in color in the overview. (B) Closeup view of the HI-loop receptor binding pocket. The sialic acid moiety (Neu5Ac) of the GM1 glycan in the SV40 VP1-GM1 pentasaccharide complex structure (PDB accession no. 3BWR) and the Y256 side chain of HPyV6 VP1 are shown in stick representation. The same view is taken for panels A and B. (C) Sequence alignment of the HI-loop region. Key residues interacting with the sialic acid moiety in the SV40-GM1 glycan and MCPyV VP1- α 2,3-sialyllactosamine complex structures (PDB accession no. 4FMI) are highlighted in blue and cyan, respectively. Regions in which all VP1 structures align with root mean square deviation (RMSD) values of <1.0 Å (dark gray) and >1.5 Å (light gray) between C α atoms are shaded.

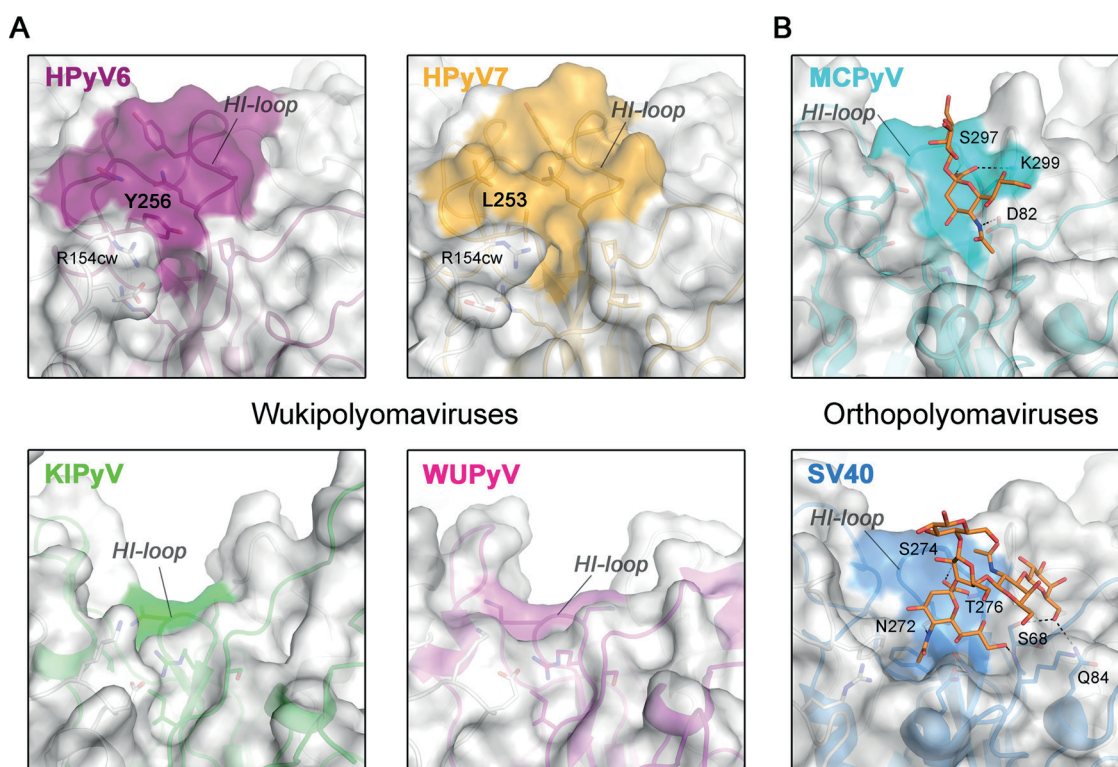


FIG 3 Surface structures of VP1 pentamers. Closeup views of VP1 pentamer top-surface regions that are involved in sialic acid engagement in the case of MCPyV and SV40 are shown. Equivalent surface sections are shown in surface and cartoon representations for VP1 pentamers from HPyV6, HPyV7, and KIPyV (PDB accession no. 3S7V) and WUPyV (PDB accession no. 3S7X) (A) and from SV40 (PDB accession no. 3BWR) and MCPyV (PDB accession no. 4FMI) (B). HI-loop residues are highlighted on the surface representations according to the colors assigned to the respective viruses. Carbohydrates (Neu5Ac and Gal of α 2,3-sialyllactosamine and GM1 pentasaccharide) in panel B are shown in stick representations (colored by atom type; carbons in orange, oxygen in red, and nitrogen in blue), and glycan-protein contacts (hydrogen bonding and salt bridges) are shown as dashed lines for MCPyV and SV40 VP1.

surface and providing direct or water-mediated contacts with glycan components. The extended HI-loop in HPyV6 and HPyV7 produces a prominent, elevated ridge on the virion surface that partially covers the glycan binding site groove present in orthopolyomaviruses (Fig. 3). Thus, the conformation of the HI-loop appears to hinder the binding of sialic acids to the HPyV6 and HPyV7 capsids. In particular, HI-loop residues Y256 (HPyV6) and L253 (HPyV7) (Fig. 2B and 3A) project into the groove and would collide with a potential sialic acid ligand. The R154 side chains in the cw HPyV6 and HPyV7 EF-loops also help to close the binding site (Fig. 3A). The uniquely elongated HI-loops of HPyV6 and HPyV7 are not conserved in either KIPyV or WUPyV (Fig. 3A). In fact, these two viruses have unusually short HI-loops (see also Fig. 2C), and surface analysis shows that they possess a deep groove leading toward the central pore (Fig. 3A). Thus, HPyV6 and HPyV7 feature a remodeled surface structure compared to all other known VP1 structures, with likely consequences for receptor binding. The HPyV6 and HPyV7 genomes sequenced so far have revealed naturally occurring VP1 amino acid variations (8), which are mostly buried in the assembled virus (not shown) and thus are likely not critical for glycan receptor engagement. The only exceptions are HPyV7 residues 63 (threonine or proline), 153 (asparagine or aspartate), and 167 (serine or threonine), which are distributed across the BC2- and EF-loops and, therefore, could theoretically account for modulated receptor interactions.

HPyV6 and HPyV7 VP1 do not bind sialic acids. Unas-

sembled recombinant VP1 pentamers are useful tools to study cell attachment and entry and cellular trafficking of polyomaviruses (30, 50). To investigate whether HPyV6 and HPyV7 engage sialic acids on cell surfaces during early steps of infection, we analyzed binding of their VP1 pentamers to two cultured human cell lines, HeLa S3 and 293TT, by flow cytometry (Fig. 4). Prior to single-cell binding experiments with Alexa Fluor 488-conjugated VP1 pentamers, cells were mock treated or incubated with *Clostridium perfringens* neuraminidase type V to remove terminal α 2,3-, α 2,6-, and α 2,8-linked sialic acids from the cell surface. JCPyV and murine polyomavirus (RA strain) VP1 pentamers bind to both cell lines in a neuraminidase-sensitive manner (30, 51) (Fig. 4B and E). In contrast, HPyV6 and HPyV7 VP1 attachment to both cell lines is not affected by enzymatic removal of sialic acids (Fig. 4A and D). The measured fluorescence signals for HPyV6 and HPyV7 VP1 pentamers are also similar to those seen with the control representing neuraminidase-insensitive binding, the JCPyV VP1 L54F mutant (Fig. 4C and F). This mutant has a disrupted VP1 sialic acid binding site and no longer engages sialylated receptors (37).

To probe for interactions in solution and to identify ligand atoms in contact with VP1, we analyzed binding of the HPyV6 and HPyV7 VP1 pentamers to the sialylated model compounds α 2,3- and α 2,6-linked sialyllactose by STD NMR spectroscopy (52). STD NMR spectroscopy has been successfully used to define, for

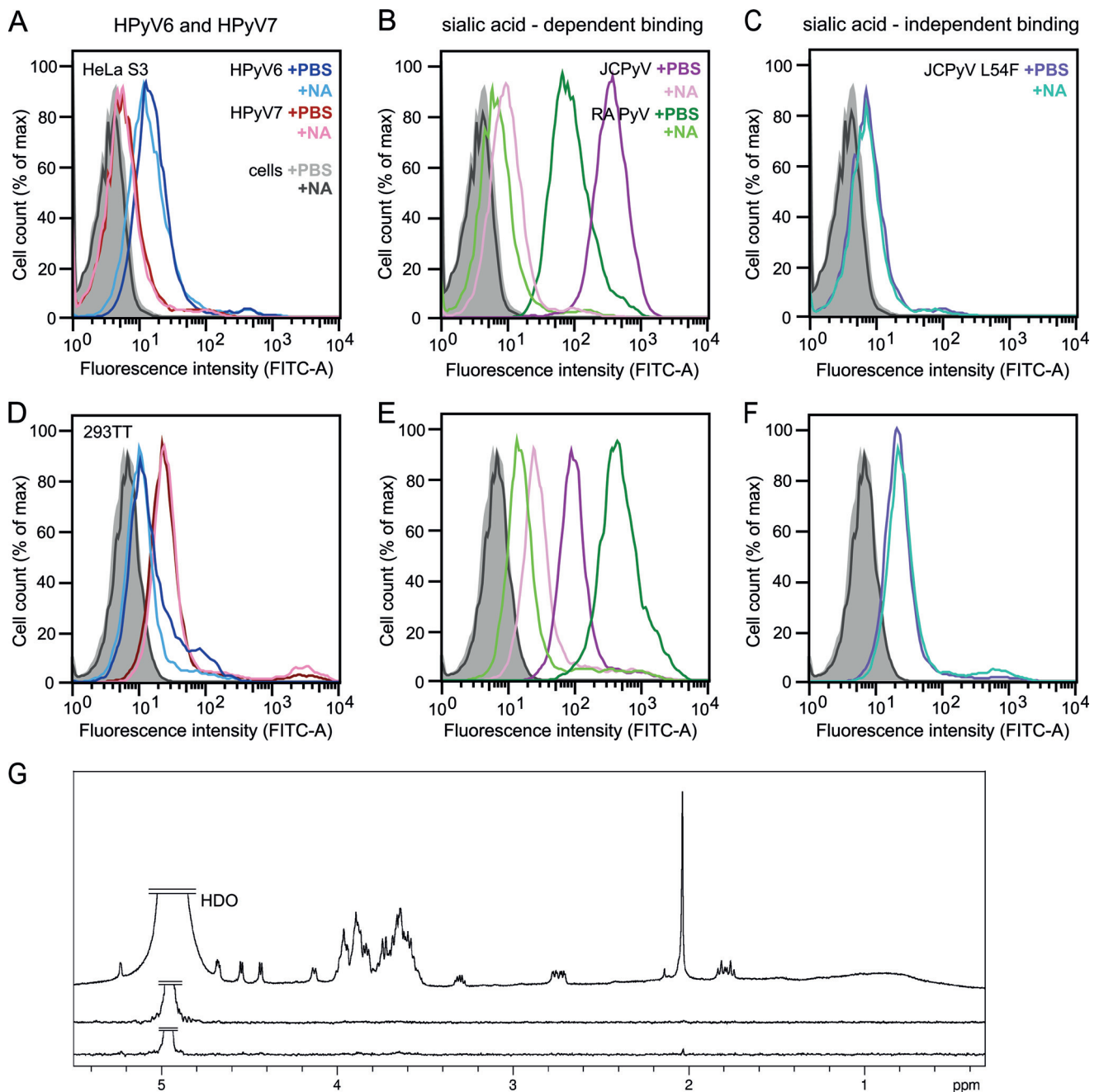


FIG 4 HPyV6 and HPyV7 VP1 do not engage sialic acids. (A to F) Cell binding analysis. (G) Saturation transfer difference (STD) NMR spectroscopy of HPyV6 and HPyV7 VP1 pentamers with α 2,3- and α 2,8-sialyllactose. (A to F) HeLa S3 (A to C) and 293TT (D to F) cells were subjected to mock treatment (PBS) or were pretreated with 0.2 U/ml *Clostridium perfringens* neuraminidase (NA), washed, and then incubated with Alexa Fluor 488-conjugated VP1 pentamers. VP1 pentamer binding was then analyzed by flow cytometry. Histograms represent the fluorescence intensity of Alexa Fluor 488 for 10,000 gated events in each case. Data for cells alone are colored gray and black for mock- and NA-treated cells, respectively. Three independent experiments were performed, and results of a typical experiment are presented. (B and E) JCPyV and murine polyomavirus (RA strain) VP1 pentamers are included as positive controls for neuraminidase-sensitive attachment (30, 51). (C and F) JCPyV L54F is a VP1 mutant with an abolished sialic acid binding site (37) and was used to test for sialic acid-independent cell binding. FITC, fluorescein isothiocyanate; max, maximum. (G) From top to bottom: ^1H reference spectrum of 50 μM HPyV7 VP1 with 1 mM α 2,3- and α 2,6-sialyllactose each; STD NMR difference spectrum recorded with the same sample; STD NMR difference spectrum of 50 μM HPyV6 VP1 with 1 mM (each) α 2,3- and α 2,8-sialyllactose. No significant saturation transfer to either capsid protein was observed. HDO peaks were truncated for clarity.

example, the interactions of MCPyV and BKPv VP1 with specific glycan receptor motifs (32, 33). No significant magnetization transfer was observed with either type of sialyllactose from HPyV6 or HPyV7 VP1, suggesting that neither protein interacts with sim-

ple α 2,3- or α 2,6-linked sialylated oligosaccharides (Fig. 4G). These findings are consistent with the crystal structure analysis, and they are also in agreement with experiments employing crystal soaking and glycan array screening, neither of which

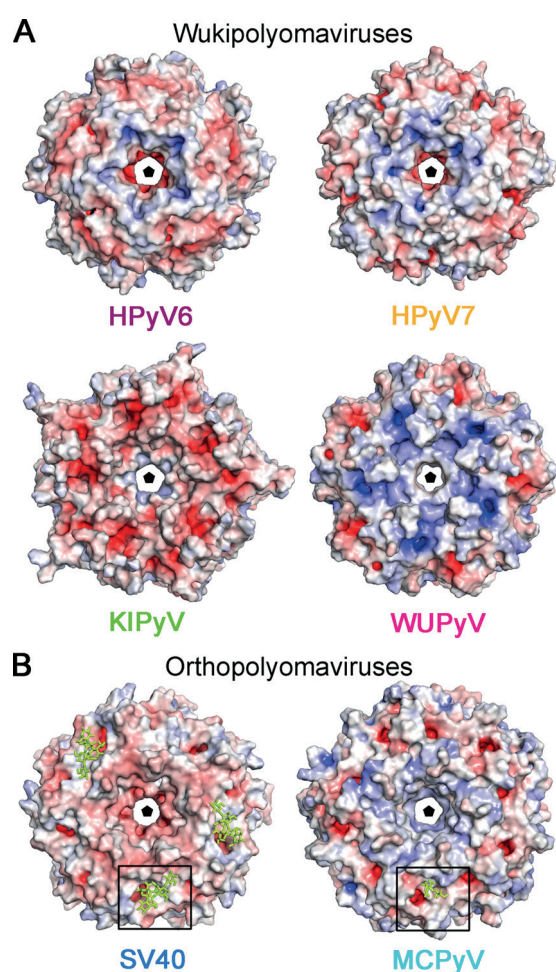


FIG 5 Electrostatic surface potentials of VP1 pentamer from the wuki- and orthopolyomavirus genera. Overall surface representations of HPyV6, HPyV7, and KIPyV (PDB accession no. 3S7V) and WUPyV (PDB accession no. 3S7X) (A) and SV40 VP1 (PDB accession no. 3BWR) and MCPyV (PDB accession no. 4FMI) (B) pentamers are colored according to electrostatic potential (calculated using APBS tool 2.1; 46), with blue and red corresponding to +7 kT and -7 kT, respectively. Views are equivalent in panels A and B and are shown from the top—the outer surface of the virion—along the 5-fold axis of the pentamer. Carbohydrates (GM1 pentasaccharide and Neu5Ac and Gal of α 2,3-sialyllactosamine) in panel B are shown in yellow stick representations, and the glycan binding site is highlighted for clarity with a box.

revealed interactions with tested sialylated compounds (data not shown).

Unique features of the virion surface. In order to obtain some initial clues about the molecular determinants of the receptor specificities and antigenic properties of HPyV6 and HPyV7 VP1, we examined their electrostatic surface potentials and compared these with the electrostatic surface potentials of other polyomavirus VP1 pentamers (Fig. 5). Consistent with their structural conservation, the electrostatic potentials of HPyV6 and HPyV7 VP1 are very similar (Fig. 5A). Both pentamers possess negatively charged patches within the central pore, which are surrounded by concentric positively and negatively charged rings around the 5-fold axis. The similarities in the distribution of surface charges suggest that the two viruses pursue similar strategies for cell surface attachment and perhaps also for viral entry.

The electrostatic surface potentials of the remaining wuki-

polyomavirus family members, WUPyV and KIPyV, differ and are also distinct from those of HPyV6 and HPyV7 (Fig. 5A). Whereas the WUPyV VP1 pentamer surface is mostly electropositive, the corresponding KIPyV residues have a more negative potential. Interestingly, the surface potentials of HPyV6 and HPyV7 VP1 are more similar to that of MCPyV (compare Fig. 5A and B) and feature a negatively charged area, a central pore around the 5-fold axis, surrounded by positively and negatively charged patches (Fig. 5A and B).

DISCUSSION

We have determined the high-resolution structures of VP1 pentamers of the recently identified human polyomaviruses HPyV6 and HPyV7. Our structure analysis reveals essential features and critical differences in surface morphology that are likely important for antigenicity and receptor engagement of these two skin viruses. The region that accommodates sialic acid receptors in other polyomaviruses is obstructed, and consistent with this, cell binding analysis and STD NMR spectroscopy (Fig. 4) as well as experiments employing crystal soaking and glycan array screening (data not shown) did not yield indications of binding of sialic acid by HPyV6 or by HPyV7.

A new sialic acid binding site different in its overall location on the VP1 pentamer from the polyomavirus sialic acid binding sites structurally characterized so far (28–30, 32–35) is possible but, in the light of our experiments, rather unlikely. The findings of our structure-function analysis suggest instead that HPyV6 and HPyV7 do not engage sialylated glycans on cell surfaces during attachment and entry and likely recognize a different receptor type. It is still possible, however, that HPyV6 and HPyV7 bind sialic acids in a different location that is present only in the fully assembled virus, such as in canyons between adjacent pentamers, and not in the free VP1 pentamer.

The analysis of electrostatic surface potentials has in some cases helped to identify potential receptor binding regions in viral proteins, for example, in sialic acid binding adeno- and rotaviruses (53, 54). We note that it is important to keep in mind that sialic acid binding sites need not always display a positive electrostatic potential, as hydrogen bonds rather than salt bridges mediate contacts with the sialic acid carboxylate in at least some polyomaviruses (55) (see also Fig. 5B). As HPyV6, HPyV7, and MCPyV are all shed from skin (8), this similarity might indicate a conserved strategy for receptor engagement. MCPyV is unique among the members of the polyomavirus family because of its sequential engagement of negatively charged glycosaminoglycan (GAG) and sialylated receptors (56). GAG binding sites on proteins are typically lysine/arginine-rich (57, 58); however, the binding site for GAGs on MCPyV VP1 is not known. The electrostatic surface potentials of HPyV6 and HPyV7 VP1 are indeed similar to that of MCPyV, with potential consequences for receptor specificity, including engagement of GAGs, but the typical shallow depression that harbors sialic acid binding sites on MCPyV (32) and other polyomavirus VP1s is clearly not evident in our HPyV6 and HPyV7 VP1 structures. Thus, only a few conclusions concerning specific receptor engagement for HPyV6 and HPyV7 are possible until their receptor class has been identified.

Our analysis of the VP1 pentamer surface properties, such as electrostatic distributions and loop morphology, suggests that a structural-biology approach can improve the phylogenetic polyomavirus classification that is based entirely on sequence similarity

(15). In terms of surface loop length and structure, HPyV6 and HPyV7 differ significantly from the other structurally characterized members of the wukupolyomavirus genus, WUPyV and KIPyV. While the latter two viruses possess strongly charged VP1 pentamers, the electrostatic surfaces of the former two resemble that of the orthopolyomavirus MCPyV more closely. Our analysis thus demonstrates the need for structure-based comparisons to inform understanding of potential receptor binding sites and conserved regions of antigenicity, two characteristics that are largely determined by the electrostatics and loop structure of VP1. It has been shown that, in some cases, structural analysis even allows the identification of evolutionary relationships that are not revealed by sequence analysis (59–61).

Our comprehensive analysis and comparison of VP1 pentamers from different polyomavirus genera lead us to propose that the observed structural diversity explains differences and similarities in the tissue tropism of wukupolyomaviruses compared with other polyomaviruses. Cellular and tissue distributions of receptors and coreceptors are important determinants of viral tropism. The establishment of experimental models and cell culture systems to propagate HPyV6 and HPyV7 as well as other newly identified polyomaviruses is clearly required. Such models should enable functional studies of receptor engagement and cell entry mechanisms of these viruses that are widely circulating within the human population.

ACKNOWLEDGMENTS

We thank members of the Stehle group for critical discussions and the staff at beamlines X06DA of the Swiss Light Source (Villigen, Switzerland) and ID14-4 at ESRF (Grenoble, France) for beam time and assistance with data collection. We acknowledge Remco Sprangers (Max Planck Institute for Developmental Biology, Tübingen, Germany) for assistance in recording the NMR data and the staff of the Flow Cytometry Core Facility, University of Tübingen, Tübingen, Germany, for technical support.

This work was supported by contract research “Glykobiologie/Glykomik” of the Baden-Württemberg Stiftung (T.S.). R.L.G. was supported by NIH grant CA37667.

REFERENCES

- Gross L. 1953. Neck tumors, or leukemia, developing in adult C3H mice following inoculation, in early infancy, with filtered (Berkefeld N), or centrifugated (144,000 X g), Ak-leukemic extracts. *Cancer* 6:948–958. [http://dx.doi.org/10.1002/1097-0142\(195309\)6:5<948::AID-CNCR2820060513>3.0.CO;2-J](http://dx.doi.org/10.1002/1097-0142(195309)6:5<948::AID-CNCR2820060513>3.0.CO;2-J).
- DeCaprio JA, Garcea RL. 2013. A cornucopia of human polyomaviruses. *Nat. Rev. Microbiol.* 11:264–276. <http://dx.doi.org/10.1038/nrmicro2992>.
- Feltkamp MC, Kazem S, van der Meijden E, Lauber C, Gorbalya AE. 2013. From Stockholm to Malawi: recent developments in studying human polyomaviruses. *J. Gen. Virol.* 94:482–496. <http://dx.doi.org/10.1099/vir.0.048462-0>.
- Dalianis T, Hirsch HH. 2013. Human polyomaviruses in disease and cancer. *Virology* 437:63–72. <http://dx.doi.org/10.1016/j.virol.2012.12.015>.
- Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, Andersson B. 2007. Identification of a third human polyomavirus. *J. Virol.* 81:4130–4136. <http://dx.doi.org/10.1128/JVI.00028-07>.
- Gaynor AM, Nissen MD, Whitley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, Sloots TP, Wang D. 2007. Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog.* 3:e64. <http://dx.doi.org/10.1371/journal.ppat.0030064>.
- Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–1100. <http://dx.doi.org/10.1126/science.1152586>.
- Schwalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. 2010. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe* 7:509–515. <http://dx.doi.org/10.1016/j.chom.2010.05.006>.
- van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalya AE, Feltkamp MC. 2010. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromised patient. *PLoS Pathog.* 6:e1001024. <http://dx.doi.org/10.1371/journal.ppat.1001024>.
- Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kuhn J, Hengel H, Ehlers B. 2011. A novel human polyomavirus closely related to the African green monkey-derived lymphotropic polyomavirus. *J. Virol.* 85:4586–4590. <http://dx.doi.org/10.1128/JVI.02602-10>.
- Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, Gordon JI, Wang D. 2012. Identification of MW polyomavirus, a novel polyomavirus in human stool. *J. Virol.* 86:10321–10326. <http://dx.doi.org/10.1128/JVI.01210-12>.
- Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, Brennan DC, Mkakosya RS, Manary MJ, Gordon JI, Wang D. 2013. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. *Virology* 436:295–303. <http://dx.doi.org/10.1016/j.virol.2012.12.005>.
- Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I, Voigt S, Schmuck R, Ehlers B. 2013. Identification of a novel human polyomavirus in organs of the gastrointestinal tract. *PLoS One* 8:e58021. <http://dx.doi.org/10.1371/journal.pone.0058021>.
- Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese T, Faust PL, Lipkin WI. 1 May 2014. Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. *J. Infect. Dis.* <http://dx.doi.org/10.1093/infdis/jiu250>.
- Johne R, Buck CB, Allander T, Atwood WJ, Garcea RL, Imperiale MJ, Major EO, Ramqvist T, Norkin LC. 2011. Taxonomical developments in the family Polyomaviridae. *Arch. Virol.* 156:1627–1634. <http://dx.doi.org/10.1007/s00705-011-1008-x>.
- Kean JM, Rao S, Wang M, Garcea RL. 2009. Seroepidemiology of human polyomaviruses. *PLoS Pathog.* 5:e1000363. <http://dx.doi.org/10.1371/journal.ppat.1000363>.
- van der Meijden E, Bialasiewicz S, Rockett RJ, Tozer SJ, Sloots TP, Feltkamp MC. 2013. Different serologic behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 polyomaviruses found on the skin. *PLoS One* 8:e81078. <http://dx.doi.org/10.1371/journal.pone.0081078>.
- Moens U, Van Ghelue M, Song XB, Ehlers B. 2013. Serological cross-reactivity between human polyomaviruses. *Rev. Med. Virol.* 23:250–264. <http://dx.doi.org/10.1002/rmv.1747>.
- De Gascun CF, Carr MJ. 2013. Human polyomavirus reactivation: disease pathogenesis and treatment approaches. *Clin. Dev. Immunol.* 2013:373579. <http://dx.doi.org/10.1155/2013/373579>.
- Wieland U, Silling S, Hellmich M, Potthoff A, Pfister H, Kreuter A. 13 January 2014. Human polyomaviruses 6, 7, 9, 10 and Trichodysplasia spinulosa-associated polyomavirus in HIV-infected men. *J. Gen. Virol.* <http://dx.doi.org/10.1099/vir.0.061259-0>.
- Schrama D, Buck CB, Houben R, Becker JC. 2012. No evidence for association of HPyV6 or HPyV7 with different skin cancers. *J. Invest. Dermatol.* 132:239–241. <http://dx.doi.org/10.1038/jid.2011.261>.
- Duncavage EJ, Pfeifer JD. 2011. Human polyomaviruses 6 and 7 are not detectable in Merkel cell polyomavirus-negative Merkel cell carcinoma. *J. Cutan. Pathol.* 38:790–796. <http://dx.doi.org/10.1111/j.1600-0560.2011.01765.x>.
- Scola N, Wieland U, Silling S, Altmeyer P, Stucker M, Kreuter A. 2012. Prevalence of human polyomaviruses in common and rare types of non-Merkel cell carcinoma skin cancer. *Br. J. Dermatol.* 167:1315–1320. <http://dx.doi.org/10.1111/j.1365-2133.2012.11141.x>.
- Kreuter A, Silling S, Dewan M, Stucker M, Wieland U. 2011. Evaluation of 4 recently discovered human polyomaviruses in primary cutaneous B-cell and T-cell lymphoma. *Arch. Dermatol.* 147:1449–1451. <http://dx.doi.org/10.1001/archdermatol.2011.330>.
- Liddington RC, Yan Y, Moulai J, Sahlbi R, Benjamin TL, Harrison SC. 1991. Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354:278–284. <http://dx.doi.org/10.1038/354278a0>.
- Stehle T, Gamblin SJ, Yan Y, Harrison SC. 1996. The structure of simian virus 40 refined at 3.1 Å resolution. *Structure* 4:165–182. [http://dx.doi.org/10.1016/S0969-2126\(96\)00020-2](http://dx.doi.org/10.1016/S0969-2126(96)00020-2).
- Schwalter RM, Buck CB. 2013. The Merkel cell polyomavirus minor

- capsid protein. *PLoS Pathog.* 9:e1003558. <http://dx.doi.org/10.1371/journal.ppat.1003558>.
28. Stehle T, Yan Y, Benjamin TL, Harrison SC. 1994. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* 369:160–163. <http://dx.doi.org/10.1038/369160a0>.
 29. Neu U, Woellner K, Gauglitz G, Stehle T. 2008. Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc. Natl. Acad. Sci. U. S. A.* 105:5219–5224. <http://dx.doi.org/10.1073/pnas.0710301105>.
 30. Neu U, Maginnis MS, Palma AS, Stroh LJ, Nelson CD, Feizi T, Atwood WJ, Stehle T. 2010. Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif. *Cell Host Microbe* 8:309–319. <http://dx.doi.org/10.1016/j.chom.2010.09.004>.
 31. Neu U, Wang J, Macejak D, Garcea RL, Stehle T. 2011. Structures of the major capsid proteins of the human Karolinska Institutet and Washington University polyomaviruses. *J. Virol.* 85:7384–7392. <http://dx.doi.org/10.1128/JVI.00382-11>.
 32. Neu U, Hengel H, Blaum BS, Schowalter RM, Macejak D, Gilbert M, Wakarchuk WW, Imamura A, Ando H, Kiso M, Arnberg N, Garcea RL, Peters T, Buck CB, Stehle T. 2012. Structures of Merkel cell polyomavirus VP1 complexes define a sialic acid binding site required for infection. *PLoS Pathog.* 8:e1002738. <http://dx.doi.org/10.1371/journal.ppat.1002738>.
 33. Neu U, Allen SA, Blaum BS, Liu Y, Frank M, Palma AS, Stroh LJ, Feizi T, Peters T, Atwood WJ, Stehle T. 2013. A structure-guided mutation in the major capsid protein retargets BK polyomavirus. *PLoS Pathog.* 9:e1003688. <http://dx.doi.org/10.1371/journal.ppat.1003688>.
 34. Neu U, Khan ZM, Schuch B, Palma AS, Liu Y, Pawlita M, Feizi T, Stehle T. 2013. Structures of B-lymphotropic polyomavirus VP1 in complex with oligosaccharide ligands. *PLoS Pathog.* 9:e1003714. <http://dx.doi.org/10.1371/journal.ppat.1003714>.
 35. Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, Feizi T, Stehle T. 19 March 2014. Crystallographic and glycan microarray analysis of human polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate. *J. Virol.* <http://dx.doi.org/10.1128/JVI.03455-13>.
 36. Stehle T, Harrison SC. 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J.* 16:5139–5148. <http://dx.doi.org/10.1093/emboj/16.16.5139>.
 37. Maginnis MS, Stroh LJ, Gee GV, O'Hara BA, Derdowski A, Stehle T, Atwood WJ. 2013. Progressive multifocal leukoencephalopathy-associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding. *mBio* 4:e00247-13. <http://dx.doi.org/10.1128/mBio.00247-13>.
 38. Kabsch W. 2010. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66:125–132. <http://dx.doi.org/10.1107/S0907444909047337>.
 39. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* 40:658–674. <http://dx.doi.org/10.1107/S0021889807021206>.
 40. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67:235–242. <http://dx.doi.org/10.1107/S0907444910045749>.
 41. Afonine PV, Grosse-Kunstleve RW, Adams PD. 2005. A robust bulk-solvent correction and anisotropic scaling procedure. *Acta Crystallogr. D Biol. Crystallogr.* 61:850–855. <http://dx.doi.org/10.1107/S0907444905007894>.
 42. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501. <http://dx.doi.org/10.1107/S0907444910007493>.
 43. Painter J, Merritt EA. 2006. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* 62:439–450. <http://dx.doi.org/10.1107/S0907444906005270>.
 44. Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53:240–255.
 45. Krissinel E, Henrick K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* 60:2256–2268. <http://dx.doi.org/10.1107/S0907444904026460>.
 46. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 98:10037–10041. <http://dx.doi.org/10.1073/pnas.181342398>.
 47. Buck CB, Pastrana DV, Lowy DR, Schiller JT. 2004. Efficient intracellular assembly of papillomaviral vectors. *J. Virol.* 78:751–757. <http://dx.doi.org/10.1128/JVI.78.2.751-757.2004>.
 48. Nicholson JK, Foxall PJ, Spraul M, Farrant RD, Lindon JC. 1995. 750 MHz ¹H and ¹H-¹³C NMR spectroscopy of human blood plasma. *Anal. Chem.* 67:793–811. <http://dx.doi.org/10.1021/ac00101a004>.
 49. Goujon M, McWilliam H, Li WZ, Valentin F, Squizzato S, Paern J, Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38:W695–W699. <http://dx.doi.org/10.1093/nar/gkq313>.
 50. Nelson CD, Derdowski A, Maginnis MS, O'Hara BA, Atwood WJ. 2012. The VP1 subunit of JC polyomavirus recapitulates early events in viral trafficking and is a novel tool to study polyomavirus entry. *Virology* 428:30–40. <http://dx.doi.org/10.1016/j.virol.2012.03.014>.
 51. Bauer PH, Cui C, Liu WR, Stehle T, Harrison SC, DeCaprio JA, Benjamin TL. 1999. Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse. *J. Virol.* 73:5826–5832.
 52. Mayer M, Meyer B. 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem. Int. Ed.* 38:1784–1788. [http://dx.doi.org/10.1002/\(SICI\)1521-3773\(19990614\)38:12<1784::AID-ANIE1784>3.0.CO;2-Q](http://dx.doi.org/10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q).
 53. Dormitzer PR, Sun ZY, Wagner G, Harrison SC. 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J.* 21:885–897. <http://dx.doi.org/10.1093/emboj/21.5.885>.
 54. Burmeister WP, Guilligay D, Cusack S, Wadell G, Arnberg N. 2004. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J. Virol.* 78:7727–7736. <http://dx.doi.org/10.1128/JVI.78.14.7727-7736.2004>.
 55. Neu U, Bauer J, Stehle T. 2011. Viruses and sialic acids: rules of engagement. *Curr. Opin. Struct. Biol.* 21:610–618. <http://dx.doi.org/10.1016/j.sbi.2011.08.009>.
 56. Schowalter RM, Pastrana DV, Buck CB. 2011. Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry. *PLoS Pathog.* 7:e1002161. <http://dx.doi.org/10.1371/journal.ppat.1002161>.
 57. Xie Q, Lerch TF, Meyer NL, Chapman MS. 2011. Structure-function analysis of receptor-binding in adeno-associated virus serotype 6 (AAV-6). *Virology* 420:10–19. <http://dx.doi.org/10.1016/j.virol.2011.08.011>.
 58. Kern A, Schmidt K, Leder C, Muller OJ, Wobus CE, Bettinger K, Von der Lieth CW, King JA, Kleinschmidt JA. 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J. Virol.* 77:11072–11081. <http://dx.doi.org/10.1128/JVI.77.20.11072-11081.2003>.
 59. Bahar MW, Graham SC, Stuart DI, Grimes JM. 2011. Insights into the evolution of a complex virus from the crystal structure of vaccinia virus D13. *Structure* 19:1011–1020. <http://dx.doi.org/10.1016/j.str.2011.03.023>.
 60. Ravanti J, Bamford D, Stuart DI. 2013. Automatic comparison and classification of protein structures. *J. Struct. Biol.* 183:47–56. <http://dx.doi.org/10.1016/j.jsb.2013.05.007>.
 61. Rissanen I, Grimes JM, Pawlowski A, Mantynen S, Harlos K, Bamford JK, Stuart DI. 2013. Bacteriophage P23–77 capsid protein structures reveal the archetype of an ancient branch from a major virus lineage. *Structure* 21:718–726. <http://dx.doi.org/10.1016/j.str.2013.02.026>.

Mutations in the GM1 Binding Site of Simian Virus 40 VP1 Alter Receptor Usage and Cell Tropism

Thomas G. Magaldi,^a Michael H. C. Buch,^b Haruhiko Murata,^c Kimberly D. Erickson,^d Ursula Neu,^b Robert L. Garcea,^d Keith Peden,^c Thilo Stehle,^{b,e} and Daniel DiMaio^{a,f,g,h}

Department of Genetics, Yale School of Medicine, New Haven, Connecticut, USA^a; Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany^b; Laboratory of DNA Viruses, Division of Viral Products, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland, USA^c; Department of Molecular, Cellular, and Developmental Biology, University of Colorado at Boulder, Boulder, Colorado, USA^d; Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA^e; Department of Therapeutic Radiology, Yale School of Medicine, New Haven, Connecticut, USA^f; Department of Molecular Biophysics and Biochemistry, Yale School of Medicine, New Haven, Connecticut, USA^g; and Yale Cancer Center, New Haven, Connecticut, USA^h

Polyomaviruses are nonenveloped viruses with capsids composed primarily of 72 pentamers of the viral VP1 protein, which forms the outer shell of the capsid and binds to cell surface oligosaccharide receptors. Highly conserved VP1 proteins from closely related polyomaviruses recognize different oligosaccharides. To determine whether amino acid changes restricted to the oligosaccharide binding site are sufficient to determine receptor specificity and how changes in receptor usage affect tropism, we studied the primate polyomavirus simian virus 40 (SV40), which uses the ganglioside GM1 as a receptor that mediates cell binding and entry. Here, we used two sequential genetic screens to isolate and characterize viable SV40 mutants with mutations in the VP1 GM1 binding site. Two of these mutants were completely resistant to GM1 neutralization, were no longer stimulated by incorporation of GM1 into cell membranes, and were unable to bind to GM1 on the cell surface. In addition, these mutant viruses displayed an infection defect in monkey cells with high levels of cell surface GM1. Interestingly, one mutant infected cells with low cell surface GM1 more efficiently than wild-type virus, apparently by utilizing a different ganglioside receptor. Our results indicate that a small number of mutations in the GM1 binding site are sufficient to alter ganglioside usage and change tropism, and they suggest that VP1 divergence is driven primarily by a requirement to accommodate specific receptors. In addition, our results suggest that GM1 binding is required for vacuole formation in permissive monkey CV-1 cells. Further study of these mutants will provide new insight into polyomavirus entry, pathogenesis, and evolution.

Attachment of virus particles to cells is the first step of infection. The cell surface receptors used by viruses are often restricted to certain cell types, and subtle changes in receptor specificity can have a profound effect on cell tropism, pathogenicity, and virulence. Therefore, understanding how viruses interact with their cell surface receptors is essential to combating virus propagation and spread. The *Polyomaviridae* are a family of small, non-enveloped, double-stranded DNA viruses that include simian virus 40 (SV40), mouse polyomavirus (mPyV), and the human viruses BK virus (BKV), JC virus (JCV), and Merkel cell polyomavirus (MCPyV). BKV and JCV are highly prevalent in the human population and can cause serious disease in immunocompromised individuals. MCPyV is believed to be the causative agent of Merkel cell carcinoma, a rare form of skin cancer (13).

Polyomavirus capsids consist of 360 molecules of the major capsid protein VP1, which are organized into 72 pentamers and form the outer shell of the icosahedral capsid (24, 44). The pentamer surface that faces the interior of the capsid is associated with either of the two minor capsid proteins, VP2 or VP3, which are closely related to each other. After virions bind productively to a cell surface receptor, they are internalized and trafficked to the endoplasmic reticulum (ER), where capsid disassembly initiates (20, 37). Eventually, in a poorly understood process, the viral genome is transported from the ER into the nucleus, where viral gene expression and replication of the genome occur (16).

The cell types infected by polyomaviruses *in vitro* differ widely. For example, although SV40 and mPyV can infect multiple cell types, infection by JCV appears more restricted (2). Moreover, polyomaviruses use different routes of entry. SV40, mPyV, and

BKV enter cells via a clathrin-independent process, while JCV uses clathrin-mediated endocytosis (9, 15, 37, 38). It is likely that the cell surface receptors used for infection contribute to these differences in tropism and entry. Polyomaviruses bind to sialylated oligosaccharides on the surfaces of cells. VP1 pentamers bind directly to the oligosaccharide portion of gangliosides, glycosphingolipids that typically contain one or more sialic acids (35, 41). SV40 VP1 exclusively uses the ganglioside GM1 as a receptor for infection, whereas mPyV uses GD1a and GT1b, and BKV uses GD1b and GT1b (26, 49). Although the affinity of the interaction between a single binding site and monomeric GM1 is low (K_d on the order of 5 mM), multivalent binding of several VP1 pentamers on the capsid surface to multiple GM1 molecules in the plasma membrane likely increases binding due to avidity effects. Monkey cells synthesize both α -5-*N*-glycol-neuraminic acid (NeuNGc) and α -5-*N*-acetyl-neuraminic acid (NeuNAc), either of which can serve as an SV40 receptor (6), but human cells synthesize only NeuNAc (52). The higher affinity of SV40 VP1 for NeuNGc-GM1 may account for SV40's ability to infect monkey cells more efficiently than human cells. Addition of GM1 to ganglioside-deficient cells results in its incorporation into the plasma membrane

Received 15 February 2012 Accepted 3 April 2012

Published ahead of print 18 April 2012

Address correspondence to Daniel DiMaio, daniel.dimaio@yale.edu.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.00371-12

and stimulates infection by SV40 and SV40-derived viral vectors (28, 49). In contrast, treating SV40 preparations with GM1 neutralizes infection (49). JCV can bind to LSTc, a linear pentasaccharide present on glycolipids and glycoproteins, and uses it as an attachment receptor for infection (34). The use of LSTc rather than a ganglioside may explain the difference in the route of entry employed by JCV and its restricted cell tropism. In addition to their role at the cell surface, gangliosides mediate sorting of intracellular polyomavirus particles from the endocytic trafficking vesicles to the ER (39).

The crystal structure of the SV40 VP1 pentamer bound to GM1 pentasaccharide revealed that GM1 binds to a shallow V-shaped groove constructed from the BC, DE, and HI loops of three VP1 monomers. In this groove, VP1 directly interacts with the terminal NeuNAc and terminal galactose on the two branches of GM1. VP1 also makes minor contacts with the *N*-acetylgalactosamine of GM1. Because there is a rigid spacer in the SV40 VP1 groove, polysaccharides that bind to VP1 must be aligned in a Y shape to fit into the groove (36). This requirement is believed to narrow the specificity of SV40 to GM1, which has a dominant Y shape in solution (40).

As SV40, BKV, and JCV share over 70% nucleotide sequence similarity with each other, it is believed that these viruses evolved from a common ancestor (21). The VP1 sequences of SV40, BKV, and JCV are largely conserved. The structures of SV40 and JCV in complex with their respective receptors, GM1 and LSTc, show that both bind sialic acid in the same orientation and with similar contacts. A small number of additional residues outside the sialic acid binding pocket also contact the oligosaccharide and enable each virus to specifically engage distinct oligosaccharide conformations and linkages (34, 36). These residues are located in the BC, DE, and HI loops, the most divergent regions of VP1, and define the context in which sialic acid must be placed in order for it to interact with SV40 or JCV VP1. BKV also binds sialylated oligosaccharides, including the gangliosides GD1b and GT1b, but it is not known how BKV achieves specificity (26). In addition to the amino acids in the receptor-binding regions of VP1, there are numerous other differences scattered throughout VP1 proteins from different polyomaviruses. It is not known if amino acids in the SV40 binding site are sufficient to specify physiologically relevant differences in receptor usage and cell tropism or if divergent amino acids elsewhere in the VP1 structure influence the conformation of the oligosaccharide binding site and indirectly contribute to determining the specific receptors used for infection and the specific cell types that are infected.

Amino acid substitutions in VP1 can alter receptor specificity and have a profound effect on pathogenicity. A highly pathogenic strain of mPyV possesses amino acid substitutions in VP1 that decrease its affinity for its receptors, allowing the virus to spread rapidly in the host (4, 14). JCV variants with mutations in the receptor-binding site of VP1 are frequently found in the blood and cerebrospinal fluid of patients with progressive multifocal leukoencephalopathy (PML), the disease caused by this virus. These variants display a decreased ability to bind to sialic acid and altered binding to ganglioside and cells, suggesting that these mutations have changed the specificity of JCV VP1 for its cell surface receptor (18, 34, 46).

Here, we used SV40 as a model to study how differences in VP1 might determine the receptors used for polyomavirus infection and cell tropism. We hypothesized that it would be possible to

change the receptor used by SV40 by introducing mutations in the GM1 binding site on VP1 and that changing the interaction with its receptor would alter the tropism of the virus. To address this hypothesis, we conducted two sequential screens to select for viable SV40 mutants with altered receptor usage, and we isolated two mutant viruses that no longer appear to use GM1 for infection. One of these viruses, which contains mutations at residues in the GM1 binding site of VP1, also displayed altered cell tropism and infected cells with low cell surface GM1 more efficiently than wild-type (WT) virus, suggesting it uses a receptor other than GM1. Strikingly, these mutations also inhibited the ability of SV40 to induce vacuolization in permissive monkey cells, a hallmark of SV40 infection, but did not impair virus production.

MATERIALS AND METHODS

Cells. HeLa S3, CV-1, 293T, LLC-MK2, and Vero cells were purchased from the American Type Culture Collection (ATCC) (Manassas, VA). Normal diploid human foreskin fibroblasts (HFFs) were obtained from the Yale Skin Diseases Research Center (YSDRC). Human embryonic kidney (HEK) 293 cells were provided by Tae Hoon Kim (Yale University), and 293TT cells were provided by Chris Buck (National Cancer Institute). All the cells listed above were grown in Dulbecco's minimal essential medium (DMEM) with 10% fetal bovine serum (FBS), standard antibiotics, 10 mM L-glutamine, and 10 mM HEPES, pH 7.2 (with standard supplements, DMEM10). Hybridomas expressing polyclonal antibody (PAb) 108 mouse anti-large T antigen (purchased from ATCC), PAb 597 mouse anti-VP1 (obtained from Edward Harlow, Harvard Medical School), and PAb BH3-FG mouse anti-VP1 were cultured in DMEM plus 20% FBS and standard supplements.

To generate HeLa S3 cells expressing short hairpin RNAs (shRNAs) against GM3 synthase (shGM3) and glucosylceramide synthase (shGC), cells were infected with concentrated retroviral stocks of pSiren encoding the appropriate shRNAs (see below) and selected in DMEM10 and 1 μ g/ml puromycin for 3 days. The cells were maintained in medium with 0.5 μ g/ml puromycin.

Reagents. Monosialoganglioside GM1 from bovine brain was purchased from Sigma-Aldrich Corp. (St. Louis, MO). Alexa Fluor 488 donkey anti-mouse and donkey anti-rabbit IgG (H+L) and Alexa Fluor 488 cholera toxin B (CTXB) were purchased from Invitrogen (Carlsbad, CA). Rabbit anti-VP1 antibodies were purchased from Abcam (Cambridge, MA). RNeasy kits and PCR purification kits were purchased from Qiagen (Valencia, CA). iScript cDNA synthesis and iQ SYBR green Supermix kits were purchased from Bio-Rad (Hercules, CA). The GeneMorph II domain mutagenesis kit was purchased from Agilent (Santa Clara, CA). RQ1 DNase was purchased from Promega (Madison, WI), and PCR grade proteinase K was obtained from Roche Applied Sciences (Indianapolis, IN).

Viruses. SV40 776 genome present in either the pBR322 or the pHM1 vector was excised with BamHI and KpnI, respectively, and religated into circles. The SV40 genome in pHM1 vector is designated pSV776.1. The ligation reaction mixture was then ethanol precipitated and transfected into 293TT or CV-1 cells with either TransIT 293 transfection reagent (Mirus, Madison, WI) or Lipofectamine (Invitrogen). Primary virus stocks were generated by three rounds of freeze/thaw lysis at day 4 (293TT) or 7 (CV-1) posttransfection. Subsequent virus stocks were generated by infecting CV-1 or 293TT cells with primary virus, and final high-titer virus stocks were prepared by using a modified Pava1 protocol described previously (17, 33). Mutations were introduced into VP1 of pSV776.1 by a standard quick-change mutagenesis protocol.

shRNAs against GM3 synthase and glucosylceramide synthase were designed using Invitrogen Block-iT RNAi Designer and cloned into the pSiren retroviral vector (Clontech, Mountain View, CA). The sequences of the shRNAs targeting GM3 synthase and glucosylceramide synthase that were cloned between the BamHI and EcoRI sites of pSiren are as follows: shGM3-1, 5'-GCACCAGTTGAGGGATATTCAC

GAATGAATATCCCTCAACTGGTGC-3'; shGM3-2, 5'-GCACTACTT CGACAGTCAATGCGAACATTGACTGTGCGAAGTAGTGC-3'; shGC-1, 5'-GCAACTGACAAACAGCCTTATCGAAATAAGGCTGTTTGTGTCAGT TGC-3; shGC-2, 5'-GCTTACTGACATGGTGAATCACGAATGATTCA CCATGTGACGTAAGC-3'. The pSiren vector expressing a nonspecific shRNA (shControl) was used as a control. Retrovirus stocks were generated in 293T cells as described previously. Virus was harvested at 48 and 72 h posttransfection and concentrated with Peg-It (System Biosciences, Mountain View, CA) as described previously (28, 53).

SV40 infection and flow cytometry of large T antigen. Cells were plated in triplicate and infected with the indicated number of virions in DMEM with 10% FBS. At the indicated times postinfection, the cells were trypsinized and washed with PBS. The cells were then fixed by adding ice-cold methanol dropwise to the cell pellet while vortexing gently, and the fixation continued for at least 20 min on ice. After centrifugation, methanol was removed, and the cells were blocked for 5 min in phosphate-buffered saline (PBS) plus 0.5% bovine serum albumin (BSA). The cell pellet was then resuspended in 100 μ l of a 1:1 mixture of 5% normal donkey serum in PBS (NDS-PBS) and Pab 108 monoclonal mouse anti-large T antigen supernatant and incubated for 1 h at 37°C. The cells were then washed twice with PBS plus 0.5% BSA and resuspended in 100 μ l NDS-PBS containing a 1:500 dilution of Alexa Fluor 488 donkey anti-mouse IgG and incubated at 37°C for 30 min. Following two washes in PBS plus 0.5% BSA, the cell pellet was resuspended in 200 μ l PBS and kept on ice until analysis. The fraction of cells expressing large T antigen was measured using a 488-nm excitation and 530-nm emission filter on a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA) and plotted on a 650-nm versus 530-nm emission filter two-dimensional (2D) density plot. The protocol used for measuring VP1 expression was identical to the protocol used for measuring large T antigen expression, except that the cells were fixed in 10% formaldehyde and permeabilized in 10% formaldehyde with 0.2% Triton X-100. Pab 597 and Pab BH3-FG mouse anti-VP1 were used to measure VP1 expression. For all infection assays, the percentages of infected cells for each condition were corrected for the possibility of multiple infections per cell using the Poisson distribution.

GM1 supplementation assay. HeLa S3 cells were plated in triplicate and were either left untreated or treated with 10 μ M GM1 overnight in DMEM with 1% FBS. The cells were then washed with DMEM10 and infected with the indicated number of encapsidated WT and mutant genomes. Infection was measured at the indicated times postinfection as described above.

GM1 neutralization assay. Equal numbers of encapsidated WT and mutant genomes were treated with 7 μ M GM1 in 1 ml of DMEM with 1% FBS for 30 min at 37°C. HeLa S3 or CV-1 cells plated in triplicate were then infected with the virus for 3 h. The cells were washed with PBS and fed with DMEM10, and infection was measured at the indicated times postinfection as described above.

Generation of mutant libraries. The generation of the VP1 mutant library used in the first screen was described previously (32). The library contained approximately 23,000 members with an average of 1 nucleotide substitution per VP1 gene. Primary virus library stocks were generated by digesting 18 μ g of the pSV776.1 library with KpnI to liberate SV40 genomes. Digestion products were then purified and transfected into four 25-cm plates of 293TT cells with 293 TransIT transfection reagent (Mirus) according to the manufacturer's protocol. The number of infectious units present in primary virus stocks was then determined by infecting CV-1 cells with serial dilutions of virus and measuring the percentage of cells expressing large T antigen as described above.

For the library used in the second screen, mutations were introduced into A70L VP1 by using the GeneMorph II domain mutagenesis kit according to the manufacturer's protocol. Briefly, the A70L VP1 gene was amplified by error-prone PCR with the VP1 primers 5'-GTTCTAGGAG TTA AAACTGGAGTAGAC-3' and 5'-TCACTGCATTCTAGTTGTGGT TTGTCC-3'. The PCR conditions for this reaction were one cycle at 95°C for 2 min; 20 cycles at 95°C for 30 s, 50°C for 20 s, and 72°C for 90 s

(amplification step); and one cycle at 72°C for 10 min. Parallel PCRs were performed with 25 and 30 cycles of the amplification step. The products of the PCR were pooled, electrophoresed on a 1% agarose gel, and purified. The PCR products were then digested with AccI and BamHI, PCR purified, and ligated into the BamHI/AccI sites of pSV776.1. Ligations were transformed into SURE electroporation-competent cells (Agilent) according to the manufacturer's protocol. The resulting library consisted of approximately 380,000 bacterial transformants. The VP1 genes of 34 isolated clones were sequenced; 29 of 34 clones contained additional mutations with an average of 4.4 nucleotide changes per gene. Bacterial transformants were scraped in LB broth and pooled, and DNA was isolated by midi prep. Infectious virus was generated by digesting 60 μ g of the A70L pSV776.1 library with KpnI to liberate SV40 genomes, religating these genomes into circles, and ethanol precipitation. DNA was then transfected into six 25-cm dishes of 293TT cells as described above. Primary virus stocks were generated by three rounds of freeze/thaw at day 4 posttransfection.

Genetic screens. To eliminate individual capsids containing multiple different VP1 molecules ("mixed capsids"), the primary library was cleared by infecting CV-1 cells with 2.1×10^6 infectious units (IU) of primary library at a low multiplicity of infection (MOI) (<0.3). At 67 h postinfection (p.i.), virus was harvested by freeze/thaw, and the titer was determined in CV-1 cells as described above. Three separate pools of 5×10^5 IU of virus were then treated with 5 μ M GM1 in DMEM10 for 15 min at 37°C. The cells were infected with the virus overnight. At 3 days postinfection, virus was harvested by freeze/thaw. A fraction of this virus was also treated with GM1 and used to infect CV-1 cells as described above. This process was repeated two times for a total of three rounds of selection. To isolate virus from these pools, a fraction of virus was again treated with GM1 and used to infect CV-1 cells for 6 days. Viral genomes were then harvested by using the Hirt preparation protocol. Briefly, infected CV-1 cells were treated with 2 ml of Hirt solution (0.6% SDS, 10 mM EDTA) for 30 min at room temperature. The cells were then scraped in Hirt solution and treated with 600 μ l of 5 M NaCl at 4°C overnight. Cellular DNA was removed by centrifugation at 15,000 rpm for 20 min. The supernatant was collected, supplemented with 100 μ l of 10% SDS, and treated with 50 μ g/ml of proteinase K for 6 h at 37°C. Viral genomes were then isolated by sequential phenol extraction, phenol-chloroform extraction, isopropanol precipitation, and ethanol precipitation. VP1 genes were excised from Hirt genomes by digestion with AccI and BamHI. VP1 fragments were gel purified and ligated into pSV776.1, and the ligations were then transformed into SURE cells.

In the second screen, 293TT and CV-1 cells were infected with 3.5×10^6 infectious units of the library at an MOI of 0.3 to clear the library. At 2.5 (CV-1) and 3 (293TT) days postinfection, virus was harvested by freeze/thaw and pooled; 5.6×10^5 IU of virus was then treated with 12 μ M GM1 at 37°C for 30 min or left untreated and used to infect two separate dishes of 293TT cells. At 4.5 days postinfection, virus was harvested by freeze/thaw. A fraction of this virus was treated with GM1 or left untreated and used to infect 293TT cells as described above. This process was repeated two times for a total of three rounds of selection. To harvest virus genomes, virus isolated after three rounds of selection was treated with GM1 or left untreated and added to 293TT cells. At 5 days postinfection, virus genomes were extracted by the Hirt protocol, and the VP1 gene was cloned into pSV776.1 as described above.

Titering virus by qPCR. The protocol used for titering encapsidated genomes was adapted from reference 48. Briefly, 5 μ l of virus stocks was treated with 4 μ l of RQ1 DNase (Promega) in 100 μ l of DNase buffer (50 mM Tris-HCl, pH 7.6, 10 mM MgCl₂) for 1 h at 37°C. After inactivating the DNase at 75°C for 30 min, samples were supplemented with 13 μ l of 10 \times PK buffer (100 mM Tris-HCl, pH 8.0, 100 mM EDTA, 2.5% SDS) and treated with 50 μ g of proteinase K (Roche) for 1 h at 37°C. DNA was isolated from the samples by using a PCR purification kit (Qiagen). The extracted DNA was then diluted 1/100 in H₂O. The number of encapsidated genomes isolated was measured by quantitative PCR (qPCR).

Briefly, PCR assays were set up in triplicate using 10 μ l iQ SYBR green Supermix, 8.2 μ l of H₂O, 0.8 μ l of 5 μ M primer set, and 1 μ l of diluted viral DNA, and reactions were run on a Bio-Rad MyiQ Single-Color Real-Time PCR detection system. The primers used for quantitative PCR (5'-GAGGAGGTTAGGGTTTATGAGGA-3' and 5'-CATCAATGTATCTTATCATGTCTGGA-3') annealed to a region of VP1 with no mutations. The number of genomes present in each virus stock was calculated by comparing the critical threshold values of virus stocks to critical threshold values of 10-fold serial dilutions (10⁹ to 10³ genomes/ μ l) of pSV776.1 run on the same plate.

qRT-PCR. Quantitative reverse transcriptase PCR (qRT-PCR) and primer design were performed as described previously (19). Briefly, total RNA was harvested using the RNeasy kit with DNase treatment. One microgram of RNA was used as a template for cDNA synthesis using an iScript cDNA synthesis kit. qRT-PCR was performed by using iQ SYBR green Supermix with 40 ng cDNA per 20- μ l reaction mixture and the Bio-Rad MyiQ Single-Color Real-Time PCR detection system. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) transcripts were detected using primers 5'-CAGCCTCAAGATCATCAGCA-3' and 5'-TGTGGTCATGAGTCCTTCCA-3'. GM3 synthase transcripts were detected using the primers 5'-CTGCCTTGACATCCTTCAGT-3' and 5'-CGATTGTGGGACGTTCTTA-3'.

VP1 pentamer expression and purification. Recombinant WT and mutant VP1 proteins were truncated after amino acid position 306 and expressed from the pGEX-4T2 plasmid (GE Healthcare, Piscataway, NJ) as glutathione S-transferase (GST) fusion proteins in *Escherichia coli* BL21(DE3) cells (Agilent). Protein expression and purification were performed as described previously (5), with some modifications. Briefly, cells were grown at 37°C in 2 \times yeast extract and tryptone (YT) broth supplemented with ampicillin (100 μ g/ml) to an optical density at 600 nm of \sim 0.2. The culture was cooled to 25°C, protein expression was induced with the addition of 0.2 mM isopropyl- β -D-thiogalactopyranoside (IPTG), and the culture was incubated overnight at 25°C (>16 h). Cell pellets were resuspended in buffer L (50 mM Tris, pH 8.0, 200 mM NaCl, 1 mM EDTA, 5% glycerol, 5 mM dithiothreitol [DTT]) supplemented with protease inhibitors (Complete Protease; Roche). The cells were lysed by incubating them with 0.1% deoxycholate and 0.5 mg/ml lysozyme for 30 min on ice, followed by DNase I (40 units supplemented with 5 mM MnCl₂) treatment for 1 h, and finally passed through a French press two times at 1,000 lb/in². The cell lysate was clarified by centrifugation, and the supernatant was applied to a glutathione-Sepharose column (GE Healthcare) equilibrated with buffer L. The flowthrough was collected by gravity filtration and reapplied to the column. The column was then washed with 5 to 10 volumes of buffer L until no protein was detectable in the wash. The protein was cleaved from the GST tag, still immobilized on the glutathione-Sepharose column, by overnight incubation at 4°C with thrombin. Peak fractions were collected, and protein concentrations were determined by a Bradford assay. The eluted protein was further purified and concentrated on a phosphocellulose cation-exchange column. Protein purity was confirmed by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining, and pentamer formation was confirmed by negative staining with uranyl acetate and electron microscopy.

VP1 pentamer binding. HeLa S3 cells were treated with 20 μ M GM1 or left untreated as described above. The cells were then washed with DMEM10 and treated with 0.25% trypsin for 15 min at 37°C. To neutralize the trypsin, the cells were washed with DMEM10 and chilled on ice for 10 min in PBS. The cells were then treated with 1 μ g of pentamers in DMEM without serum for 1 h on ice. The cells were washed with ice-cold PBS and fixed with 1% paraformaldehyde diluted in PBS. The cell pellet was then resuspended in 100 μ l of PBS containing a 1:400 dilution of rabbit anti-VP1 and incubated at room temperature for 1 h. The cells were washed two times with PBS and resuspended in 100 μ l of PBS containing a 1:500 dilution of Alexa Fluor 488 donkey anti-rabbit IgG for 30 min. Following two washes in PBS, the cell pellet was resuspended in 200 μ l of PBS and kept on ice until analysis. VP1 binding was measured using a

488-nm excitation and 530-nm emission filter on a FACSCalibur flow cytometer.

Cholera toxin B binding. Cells were mock treated or treated with GM1 as described above. After removing unincorporated GM1, the cells were trypsinized, neutralized, and washed once with ice-cold PBS. The cells were then resuspended in 100 μ l of ice-cold DMEM with 1 μ g/ml Alexa Fluor 488-labeled CTXB and incubated at 4°C for 30 min. The cells were then washed with cold PBS, resuspended in PBS, and kept on ice until analysis. CTXB binding was analyzed using a 488-nm excitation and 530-nm emission filter on a FACSCalibur flow cytometer.

Virus production assay. WT and A70L/F75L/H129Q viruses were excised from their vector with KpnI and religated into circles overnight. The ligation was then ethanol precipitated and transfected into CV-1 cells in triplicate as described above. At 80 h posttransfection, virus was harvested by three rounds of freeze/thaw. The amount of virus produced was determined by qPCR as described above.

Molecular modeling. The structure of SV40 VP1 in complex with GM1 was uploaded into COOT (11), and residues A70, F75, H129, and N138 were mutated accordingly. All possible rotamers were evaluated, and in each case, the most plausible rotamer was chosen. In addition, energy minimization calculations were carried out for each of the two most likely rotamers of leucine, valine, and threonine in the mutants A70L, A70V, and A70T, as well as the most abundant rotamers for F75L and N138Y. Water, small molecules, and carbohydrate ligands were deleted from the structures, which were then subjected to simulated annealing in PHENIX (1) from 500 to 300 K in 20 K steps. To assess the effects of the mutations, all residues within a distance of 7 Å from the mutated residues in the resulting Protein Data Bank (PDB) files were superimposed over the ones before energy minimization in COOT. The root mean square deviation (RMSD) values of the superimpositions were between 0.21 and 0.32 for all atoms of one binding site, indicating only very small structural changes in the area of the binding pocket surrounding residue 70 upon mutation. Images were created using PYMOL (43).

Statistical analysis. A two-tailed *t* test was used for all statistical analysis of data.

RESULTS

Design of a screen for SV40 mutants that are resistant to GM1 neutralization. Because GM1 binds to the major SV40 capsid protein, VP1, we reasoned that mutations in VP1 might eliminate the requirement for GM1 in infection. Therefore, we used libraries of VP1 mutants to isolate viruses that no longer used GM1 for infection. We previously constructed a library of viable SV40 mutants with VP1 mutagenized by error-prone PCR and used this library to identify the VP1 epitope recognized by neutralizing monoclonal antibodies (31). The library contains approximately 23,000 members, with an average of one nucleotide substitution per mutant. To identify SV40 mutants that no longer required GM1 for infection, we took advantage of the ability of purified GM1 to neutralize WT SV40. Treatment of SV40 with GM1 at 37°C prior to infection of CV-1 cells, a monkey cell line permissive for SV40 infection, inhibits infection (49). To identify viruses that use a different receptor, we selected mutant viruses that escaped neutralization by GM1. We note that the experiments described here used NeuNAc-GM1 (designated GM1) purchased from a commercial source because NeuNGc-GM1 is not readily available.

The design of this genetic screen is shown in Fig. 1A. CV-1 cells were first infected with primary stocks of the virus library at an MOI of <0.3 infectious units/cell to minimize the number of coinfecting cells that produced capsids comprised of more than one type of VP1 ("mixed capsids"). After one round of infection, replication, and assembly, the resulting virus was harvested by freeze/thaw and used for the subsequent screen. We treated three

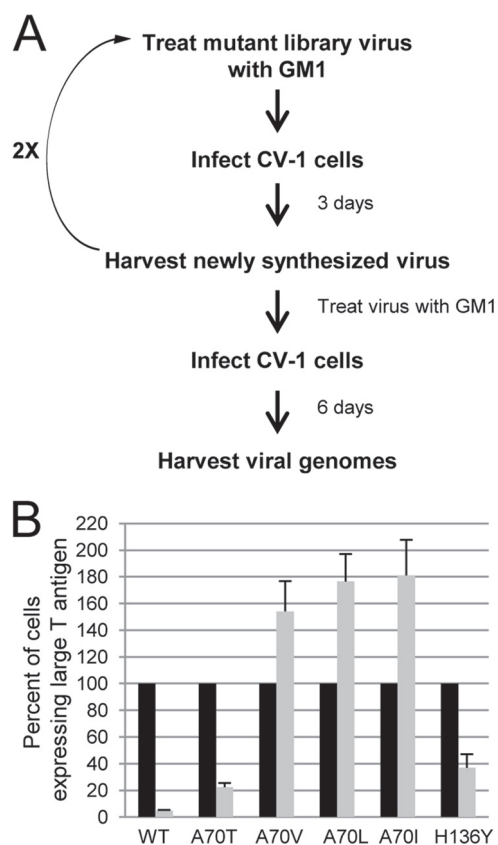


FIG 1 Design of a screen to select for SV40 mutant viruses that are resistant to GM1 neutralization. (A) Schematic outline of the screen to isolate GM1 neutralization-resistant mutant viruses. (B) The WT and indicated mutant viruses were left untreated (black bars) or treated with 7 μ M GM1 for 30 min at 37°C (gray bars). CV-1 cells were then infected with 1,000 virions/cell for 3 h. The fraction of cells expressing large T antigen was measured by immunostaining and flow cytometry at 24 h p.i. and is displayed as percent infection relative to cells infected with the respective untreated virus. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in at least five independent experiments. The error bars indicate standard deviations.

separate pools of mutant virus stocks with NeuNAc-GM1 for 15 min at 37°C and then infected CV-1 cells. After 3 days, virus was harvested, similarly treated with GM1, and used to infect CV-1 cells as before. This process was repeated two times for a total of three rounds of selection. To test whether the selection enriched for viruses that were resistant to GM1 neutralization, WT SV40 and mutant viruses harvested after three rounds of selection were treated with GM1 or left untreated. CV-1 cells were then infected with these virus preparations, and at 24 h p.i., the efficiency of infection was measured by immunostaining and flow cytometry to determine the fraction of cells expressing large T antigen, the major early gene product of SV40. Infection by GM1-treated WT SV40 was substantially inhibited compared with untreated virus, but the pools of mutant virus were not inhibited by GM1 treatment (data not shown), indicating that the screen enriched for viruses that were resistant to GM1 neutralization.

Viruses with mutations at residue 70 of VP1 are resistant to GM1 neutralization. In order to isolate individual GM1-resistant mutant viruses, viruses harvested after three rounds of screening were treated with GM1 and used to infect CV-1 cells. Six days

postinfection, SV40 genomes were isolated, and the *AccI*-to-*Bam*HI fragment of the VP1 gene was cloned into the WT SV40 backbone in the plasmid pSV776.1. At least 10 clones from each pool were sequenced (Table 1). Two WT genomes and five different mutants were identified from 31 complete VP1 sequences. All of the mutant viruses contained either an alanine-to-valine (27/31) or an alanine-to-threonine (2/31) substitution at residue 70 of VP1; most mutants also contained mutations at other positions. We focused subsequent studies on mutant viruses containing the single substitutions A70V and A70T.

To determine whether these mutants were resistant to GM1 neutralization, full-length mutant virus genomes were excised from the vector and transfected into CV-1 cells. After 7 days, virus was harvested and treated with DNase to digest DNA that was not packaged into virus particles. Encapsidated viral genomes were quantitated by quantitative PCR (qPCR) to assess the relative numbers of virions in the preparation. Equal numbers of WT, A70V, A70T, and H136Y (32) virions were treated with GM1 or left untreated and added to CV-1 cells. At 24 h p.i., the efficiency of infection was determined by flow cytometry for large T antigen (Fig. 1B). Infection by WT SV40 was reduced 20-fold by GM1 treatment. Strikingly, infection by the A70V mutant virus not only was resistant to neutralization, but was actually slightly stimulated by GM1. Infection by the A70T and H136Y mutant viruses was reduced by 5-fold and 2.5-fold, respectively, by GM1, indicating that the mutants were partially resistant to GM1 neutralization. Because only the VP1 fragment from the mutant genomes was cloned into the WT SV40 genome, the phenotype of the mutants can unequivocally be assigned to the mutations present in VP1. To assess the consequences of other hydrophobic amino acids at residue 70, leucine and isoleucine substitutions were introduced at this position, and mutant viruses were generated. These substitutions also resulted in a complete block to GM1 neutralization, similar to the A70V substitution (Fig. 1B). Examination of the structure of the WT SV40 VP1 pentamer bound to GM1 (Fig. 2) revealed that the side chain of the alanine at position 70 directly interacts with the terminal galactose of GM1 (a distance of 4 Å). Larger side chains at position 70 would result in unfavorable contacts with the galactose (see Fig. 10A). Taken together, these data suggested that inserting a larger amino acid at residue 70 of VP1 altered the ability of SV40 to interact with GM1 and therefore decreased the ability of GM1 to neutralize infection.

Mutations at A70 reduce the ability of SV40 to use GM1 for infection. We next tested the ability of the A70 mutant viruses (e.g., A70V, A70T, A70L, and A70I) to use GM1 for infection by

TABLE 1 *AccI*-to-*Bam*HI fragments of VP1 excised from harvested genomes, cloned, and sequenced

VP1 sequence	No. of isolates			Total
	Pool 1	Pool 2	Pool 3	
WT	0	0	2	2
A70V	3	5	3	11
A70T	1	1	0	2
A70V/A255S	3	0	3	6
A70V/S261T	4	3	2	9
A70V/I120V	0	1	0	1
Total	11	10	10	31

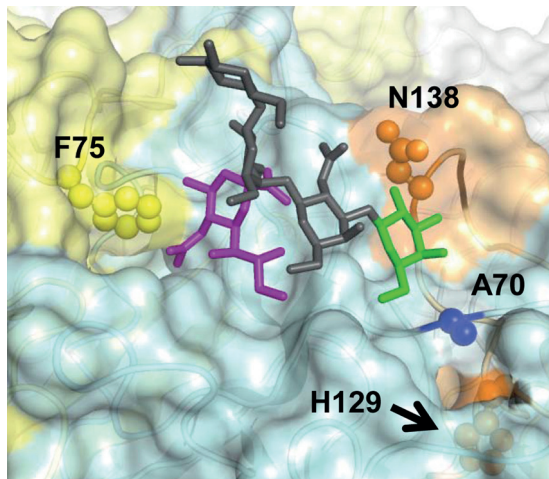


FIG 2 Predicted structure of an SV40 VP1 pentamer bound to GM1. Shown is a PyMol rendering of one GM1 binding site in a VP1 pentamer. VP1 monomers are shown as ribbon tracings, with the surface of each monomer highlighted in a different color. The monomer with the most interactions with GM1 is colored blue, while the clockwise and counterclockwise monomers are highlighted in yellow and orange, respectively. Monomers that do not contribute to this GM1 binding site are shown in gray. Side chains of amino acids mutated in GM1-resistant mutants are depicted in ball-and-stick mode and are colored according to the monomer in which they reside. The arrow points to H129, a residue that is not located at the VP1 surface. GM1 is depicted in stick representation, with its α -5-*N*-acetylneuraminic acid, glucose/galactose/*N*-acetylgalactosamine, and galactose moieties shown in magenta, dark gray, and green, respectively.

conducting a GM1 supplementation assay in HeLa S3 cells. These cells have significantly lower levels of cell surface GM1 than CV-1 cells (as assessed by cell surface binding of fluorescently labeled cholera toxin B [CTXB], a bacterial toxin subunit that binds to GM1 [7]) (Fig. 3A). HeLa S3 cells were incubated with GM1 overnight to allow incorporation of GM1 into the plasma membrane (Fig. 3A). The cells were then washed to remove unincorporated GM1 and infected with equal numbers of WT and mutant virions. Infection efficiency was determined at 48 h p.i. by flow cytometry. Addition of GM1 to HeLa S3 cells stimulated infection by WT virus 4.5-fold, but infection by the A70 mutant viruses increased only 2- to 3-fold relative to untreated cells, suggesting that these mutants have a partial defect in the ability to use GM1 for infection (Fig. 3B). The A70V, A70I, and A70L mutant viruses, which were completely resistant to GM1 neutralization, displayed the least stimulation by GM1. The fact that addition of GM1 to cells stimulates infection by the A70 mutant viruses suggested that these mutants can still interact with GM1.

We next measured the effects of these mutations on the efficiency of infection. CV-1 cells were infected with equal numbers of WT and mutant virions. At 24 h p.i., infection efficiency was measured by flow cytometry for large T antigen. Figure 3C shows that the A70V, A70I, and A70L mutant viruses had a 3- to 4-fold reduction in their ability to infect CV-1 cells compared with the WT, while A70T and H136Y mutants showed a modest reduction in activity. These data indicated that more A70 mutant virus particles are needed to infect the same number of CV-1 cells as WT virus. Since these mutants are resistant to GM1 neutralization and are impaired in their ability to be stimulated by the addition of GM1 to cells, this decrease in infectivity likely represents a de-

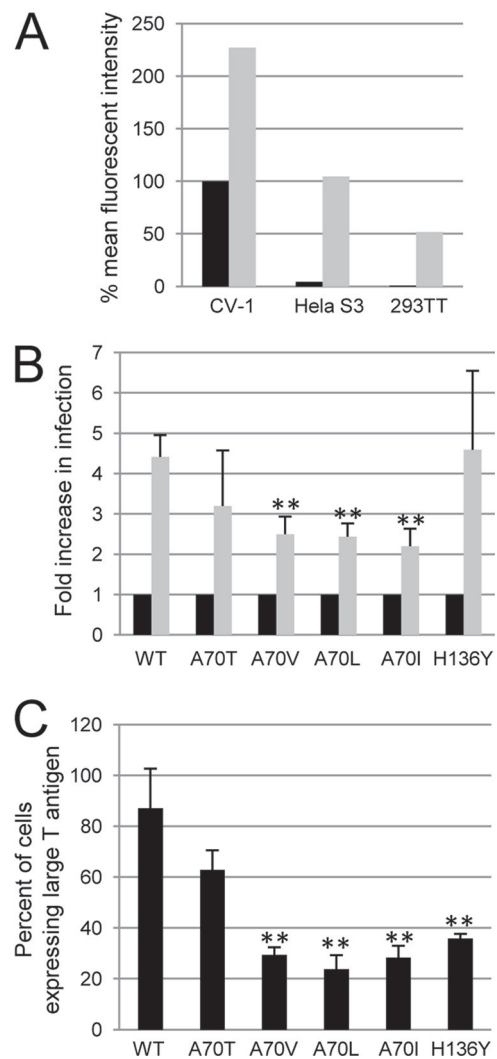


FIG 3 SV40 A70 VP1 mutant viruses are partially defective in their ability to be stimulated by addition of GM1 to cells and are less infectious than WT virus. (A) Cells were treated with GM1 (gray bars) overnight or left untreated (black bars). To assess cell surface GM1 levels, binding of fluorescently labeled CTXB to the indicated cell lines was measured by flow cytometry. CTXB binding is displayed as the percent mean fluorescence intensity relative to untreated CV-1 cells. (B) HeLa S3 cells were left untreated (black bars) or treated with GM1 overnight (gray bars) and then infected with 1,000 virions/cell of WT and the indicated mutant SV40 virus. The fraction of cells expressing large T antigen was measured by immunostaining and flow cytometry at 48 h p.i. and is displayed as the fold increase in infection relative to the respective untreated cells. The averages of three independent experiments are presented. The asterisks indicate a significant difference between the fold increase in infection by mutant virus and the fold increase in infection of WT virus (**, $P < 0.009$). (C) CV-1 cells were infected with 1,000 virions/cell in the absence of GM1 treatment. Infection was measured as in Fig. 1B and is depicted as the percentage of cells expressing large T antigen. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in two independent experiments. The asterisks indicate a significant difference between infection of cells with WT SV40 and the indicated mutant virus (**, $P < 0.005$). The error bars indicate standard deviations.

crease in the ability of the mutant viruses to use GM1 for infection of CV-1 cells.

A screen for SV40 mutants that no longer use GM1. The results reported above described the isolation of SV40 mutants re-

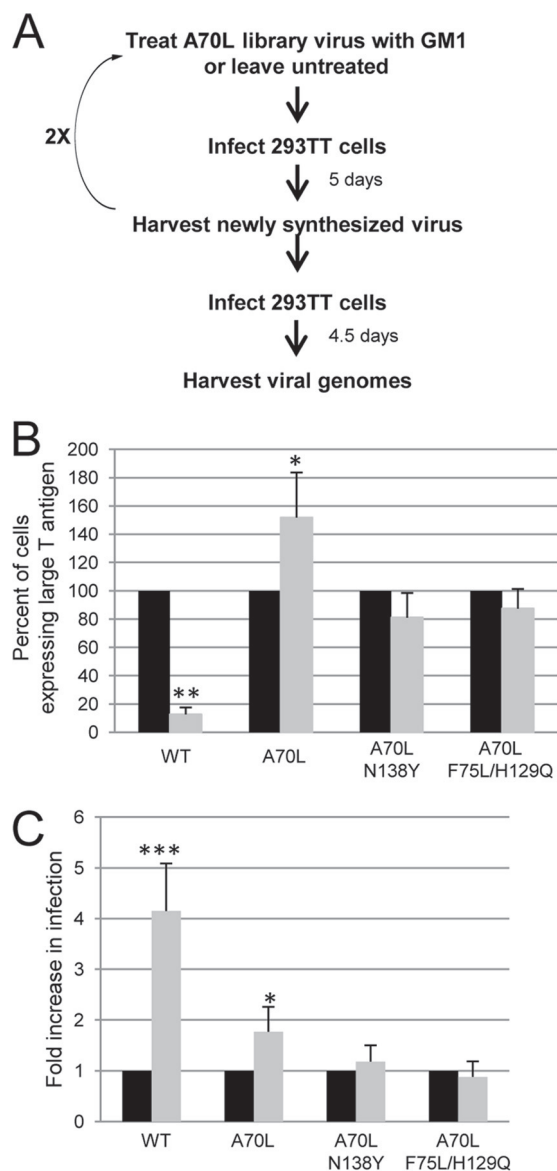


FIG 4 Design of a screen to isolate mutant viruses that no longer use GM1. (A) Outline of screen conducted to isolate SV40 mutant viruses that do not use GM1 for infection. (B) Equal numbers of WT SV40 or the indicated mutant viruses were left untreated (black bars) or treated with 7 μ M GM1 (gray bars). HeLa S3 cells were then infected with 25,000 virions/cell, and infection was measured 48 h p.i. as in Fig. 1B. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in six independent experiments. The asterisks indicate a significant difference in the fold change in infection of cells with GM1-treated virus relative to cells infected with the respective untreated virus (*, $P < 0.04$; **, $P < 0.002$). The error bars indicate standard deviations. (C) HeLa S3 cells were treated with GM1 and infected with 5,000 genomes/cell as in Fig. 3B. Infection was measured and is presented as in Fig. 3B. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in five independent experiments. The asterisks indicate a significant difference in the fold change in infection of cells treated with GM1 relative to untreated cells (*, $P < 0.03$; ***, $P < 0.0004$).

sistant to GM1 neutralization. However, because these mutants were still stimulated by the addition of GM1 to cells, albeit at a reduced level, and were impaired in the ability to infect CV-1 cells, which have high levels of GM1, we concluded that the mutant

viruses continue to use GM1 for infection. Therefore, we conducted a second screen to isolate mutant viruses that were incapable of using GM1 and were forced to use an alternative receptor. In order to isolate such viruses, we screened a library that was already partially defective for recognizing GM1 and conducted the screen in permissive cells with low levels of cell surface GM1. This new library of viable SV40 VP1 mutant viruses was constructed on the background of the A70L VP1 substitution, which impairs the interaction with GM1, as described above. The library contained approximately 380,000 members with an average of approximately four nucleotide changes per VP1 gene. Prior to selection, the library was first passaged at low multiplicity in permissive cells to eliminate mixed capsids, as described above. We screened the library in the human cell line 293TT, which has low levels of GM1 compared with CV-1 cells and constitutively expresses SV40 large T antigen to allow high-level viral DNA replication and virus production (Fig. 3A). Although 293TT cells are poorly infectible by SV40 in the absence of GM1, incubation of these cells with GM1 significantly increased infection (data not shown), suggesting that the inefficient infection of the cells was the result of low levels of endogenous GM1 and not an intracellular block. The design of this screen is outlined in Fig. 4A. We first incubated the mutant A70L virus library with GM1 to select against potential revertants, or we used the library without prior GM1 treatment. We then infected 293TT cells with either untreated or GM1-treated virus. After four rounds of serial screening in 293TT cells, viral genomes were harvested, and the VP1 gene was cloned into pSV776.1.

As predicted, all of the mutants recovered from cells retained the A70L substitution, and the vast majority contained additional amino acid substitutions. Ten unique mutant viruses were identified from 13 sequenced clones isolated from the untreated pool of virus (Table 2), and three unique mutants were identified from 12 clones sequenced from the GM1-treated pools. Notably, several of the mutant viruses isolated from this screen contained substitutions at VP1 residues that either interact directly with GM1 or lie immediately adjacent to direct interactors (Fig. 2; see Fig. 11) (36). For example, the N138 side chain directly contacts the terminal galactose, whereas F75 is part of the hydrophobic pocket that accommodates the N-acetyl group of NeuNAc. The H129 side

TABLE 2 AclI-to-BamHI fragments of VP1 excised from harvested genomes, cloned, and sequenced

Sequence	No. of isolates	
	Pool 1 (no GM1)	Pool 2 (GM1)
A70L	2	2
A70L/T273P	2	0
A70L/F75L/H129Q	1	0
A70L/T242A/T273P	1	0
A70L/Q74R	1	0
A70L/Q54R/V184G	1	0
A70L/Q54R/V184G/Y285S	1	0
A70L/N138Y	1	6
A70L/E83A/N138K	1	0
A70L/Q317H	1	0
A70L/A13V/N138Y	1	0
A70L/E160D/V184F	0	2
A70L/N138K	0	2
Total	13	12

chain is in the vicinity of the GM1 binding site but does not form direct contacts with the ganglioside. Because the A70L/N138Y and A70L/F75L/H129Q mutant viruses contained mutations at residues that directly interact with GM1, we investigated their phenotypes further.

A70L/N138Y and A70L/F75L/H129Q mutant viruses are neither neutralized nor stimulated by GM1. To determine whether the mutant viruses isolated from this screen were still resistant to GM1 neutralization, we treated virus with GM1 or left virus untreated, as described above, and then infected HeLa S3 cells. The A70L/N138Y and A70L/F75L/H129Q viruses were resistant to GM1 neutralization and, unlike the A70L virus, treating these mutant viruses with GM1 did not stimulate infection (Fig. 4B). When GM1 was added to HeLa S3 cells, infection with the WT and A70L viruses was stimulated, whereas infection by the A70L/N138Y and A70L/F75L/H129Q mutant viruses was not (Fig. 4C). Similar results were obtained when GM1 was added to 293TT cells (data not shown). These data demonstrated that the A70L/N138Y and A70L/F75L/H129Q mutant viruses no longer used GM1 for infection. Infection by the other mutant viruses isolated from this screen was stimulated by the addition of GM1 to cells to an extent comparable to that of A70L mutant virus (data not shown), implying that they still used GM1 for infection. These mutants were not studied further.

Mutant viruses with an A70L mutation cannot bind to GM1 on the cell surface. Because mutant viruses with amino acid substitutions at residue 70 are resistant to GM1 neutralization and have a defect in their ability to be stimulated by adding GM1 to cells, we hypothesized that these mutants were unable to bind to GM1. To measure the binding of VP1 to GM1 on the cell surface, we purified recombinant VP1 pentamers with a C-terminal truncation after position 306, which prevents formation of higher-order VP1 structures, and confirmed pentamer formation by electron microscopy (data not shown). HeLa S3 cells were treated with GM1 overnight or left untreated. After removal of unincorporated GM1, the cells were harvested with trypsin, which eliminates background pentamer binding without affecting GM1 surface levels (data not shown), and then incubated with WT, A70L, A70L/N138Y, and A70L/F75L/H129Q recombinant VP1 pentamers on ice for 1 h. The cells were then washed to remove unbound pentamers, and bound VP1 was detected by immunostaining and flow cytometry. In the absence of GM1, neither WT nor mutant VP1 pentamers bound cells (Fig. 5, black lines). Addition of GM1 resulted in a dramatic increase in the binding of WT pentamers to cells (Fig. 5, blue lines) but did not stimulate binding by the mutant pentamers. GM1 addition also stimulated binding of intact WT capsids but did not stimulate binding of any of the three types of mutant capsids (data not shown). Thus, the introduction of the A70L substitution prevented VP1 pentamers or capsids from binding to GM1 on the cell surface, as measured by flow cytometry, consistent with the biological evidence that these viruses are impaired in their interaction with GM1.

The A70L/F75L/H129Q VP1 mutant virus has altered cell tropism. Although GM1 binding was eliminated by the VP1 mutations, the mutant viruses were still able to infect cells. To test whether the tropism of these viruses was changed, we infected a panel of cell lines with an equivalent number of WT and mutant virions. Infection efficiency was then assessed by flow cytometry for large T antigen (in CV-1 and HeLa S3 cells and HFFs) or VP1 (in 293TT cells, which constitutively express T antigen). The A70L

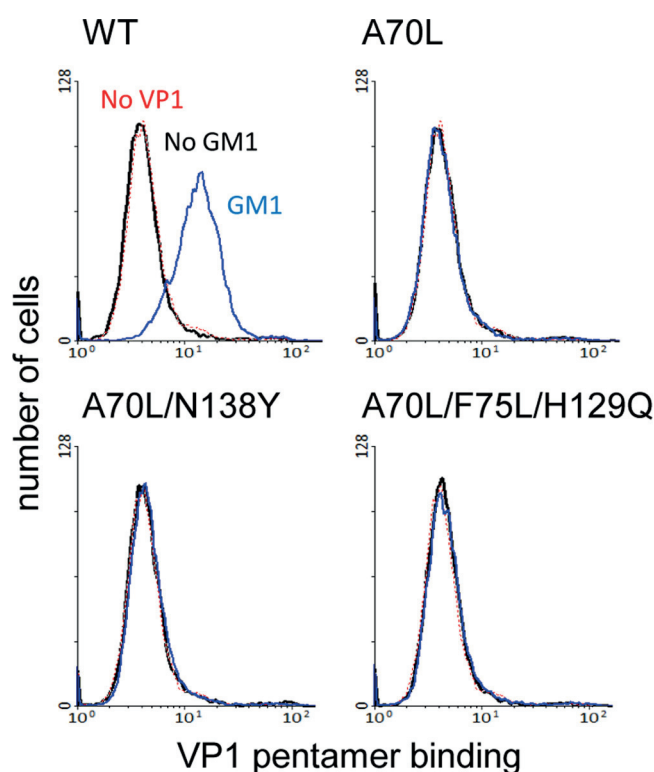


FIG 5 A70L mutant VP1 pentamers cannot bind to GM1 on the surface of cells. HeLa S3 cells were left untreated (black lines) or treated with GM1 overnight (blue lines) and harvested with trypsin. The cells were then incubated with 1 μ g of WT or mutant VP1 pentamers for 1 h on ice, washed, and fixed. Cells treated with GM1 and incubated without VP1 are depicted as red lines. Bound pentamer was detected by immunostaining and flow cytometry of VP1. The results of a typical experiment are presented. Similar results were obtained in three independent experiments.

mutant virus infected all cell lines 2- to 3-fold less efficiently than WT virus (Fig. 6A). Like A70L, the double- and triple-mutant viruses infected monkey CV-1 cells 3-fold less than WT virus. In contrast, infection by the A70L/F75L/H129Q mutant was increased 3.5- to 4-fold compared with WT virus in 293TT and HeLa S3 cells, both of which display lower levels of cell surface GM1 than CV-1 cells. Infection of HeLa S3 and 293TT cells by the A70L/N138Y mutant virus was restored to the level seen for WT virus in these cells. These results indicated that the substitutions altered the cell tropism of SV40 and allowed the mutant virus to infect cells with low levels of GM1 more efficiently than WT virus. This change in cell tropism appeared to be cell type specific rather than species specific, because the mutant viruses failed to infect primary HFFs. In addition, neither the WT nor the mutant viruses were able to infect primary human keratinocytes (data not shown).

To further analyze this difference in tropism, we compared the abilities of WT and mutant viruses to infect additional cell lines, including kidney cells, a cell type infected by SV40 *in vivo*. HEK 293 (human), Vero (African green monkey kidney), and LLC-MK2 (rhesus macaque kidney) cells were infected with equal numbers of virions, and infection efficiency was measured by flow cytometry. As expected, infection by the A70L mutant was reduced in all three cell lines compared with the WT (Fig. 6B). However, although the A70L/N138Y and A70L/F75L/H129Q mutant

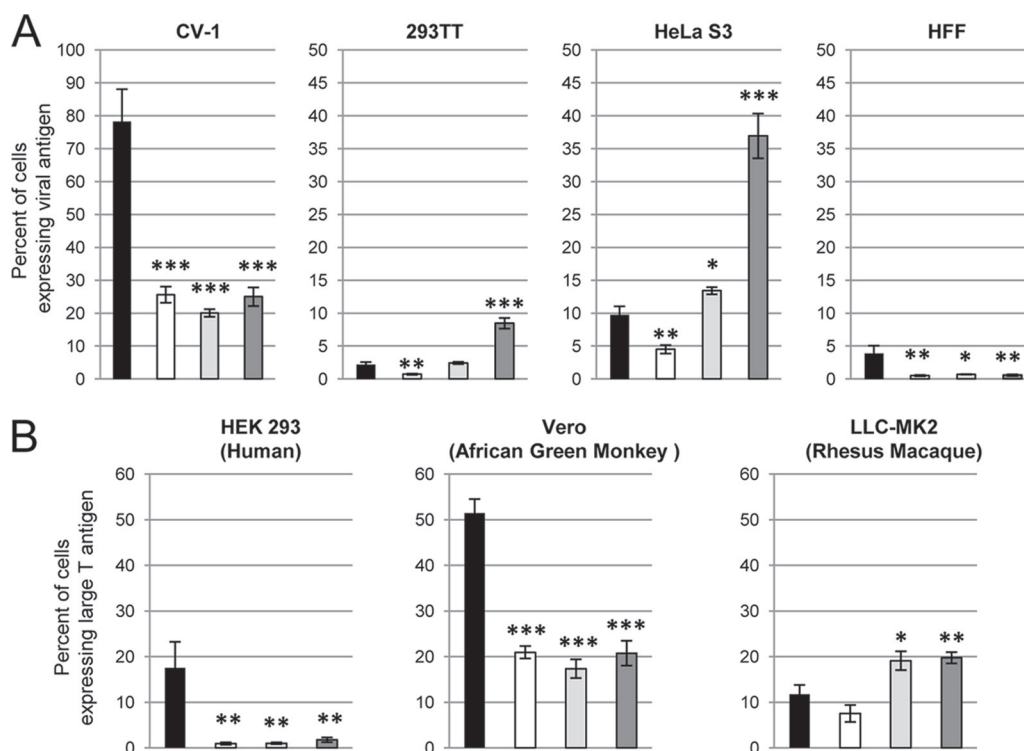


FIG 6 The A70L/F75L/H129Q mutant has altered tropism. (A) The indicated cells were infected with 25,000 virions/cell of WT (black bars), A70L (white bars), A70L/N138Y (light-gray bars), and A70L/F75L/H129Q (dark-gray bars) virus. Infection was measured by immunostaining and flow cytometry of large T antigen (CV-1 and HeLa S3 cells and HFFs) or VP1 (293TT cells) 24 h p.i. and is presented as in Fig. 3C. Similar results were obtained in at least three independent experiments. The asterisks indicate a significant difference between infection of cells with WT SV40 and the indicated mutant virus (*, $P < 0.02$; **, $P < 0.01$; ***, $P < 0.0009$). The error bars indicate standard deviations. (B) HEK 293, Vero, and LLC-MK2 cells were infected with 25,000 virions/cell of the WT or indicated mutant viruses. At 24 h p.i., infection was measured and is displayed as in panel A. Similar results were obtained in three independent experiments. The asterisks indicate a significant difference between infection of cells with WT SV40 and the indicated mutant virus (*, $P < 0.02$; **, $P < 0.01$; *** $P < 0.0003$).

viruses did not infect HEK 293 cells and infection was reduced 2.5-fold in Vero cells compared with WT virus, the mutants infected LLC-MK2 cells almost twice as well as WT virus. Taken together, these data demonstrated that altering the GM1 binding site of SV40 changed the tropism of the A70L/F75L/H129Q mutant virus, presumably by allowing it to utilize a receptor other than GM1.

All three substitutions in A70L/F75L/H129Q VP1 are required for the change in cell tropism. One or both of the additional substitutions in A70L/F75L/H129Q VP1 must contribute to the altered receptor specificity and change in tropism. In order to determine which amino acid substitution(s) altered receptor usage, we introduced F75L or H129Q separately into A70L VP1 and measured how these individual changes affected GM1 usage. Like the A70L/F75L/H129Q virus, both double-mutant viruses were completely resistant to GM1 neutralization (data not shown). Infections by the A70L and A70L/F75L mutant viruses were stimulated to similar extents by the addition of GM1 to HeLa S3 cells, while the A70L/H129Q mutant virus was not stimulated (Fig. 7A). These data suggest that H129Q further interferes with the ability of SV40 to use GM1 for infection. We next measured the effects of these mutations on the tropism of SV40. CV-1 and HeLa S3 cells were infected with equal numbers of virions, and infection efficiency was measured by flow cytometry for large T antigen. All of the mutants infected CV-1 cells less efficiently than WT virus. However, although the A70L mutant displayed a reduced ability

to infect HeLa S3 cells, the A70L/F75L and A70L/H129Q viruses infected these cells as efficiently as WT virus, and infection by the triple-mutant virus was increased 3.5-fold compared with the WT (Fig. 7B). Although the F75L and H129Q substitutions can alter the tropism of the A70L virus, both of these substitutions appear to be needed for optimal interactions with a novel receptor and maximum infection of HeLa S3 cells.

Mutants with altered cell tropism still require gangliosides for infection. Because the A70L/F75L/H129Q and A70L/N138Y mutant viruses are no longer stimulated by GM1 and the A70L/F75L/H129Q virus can infect cell lines with low levels of GM1 more efficiently than WT virus, we hypothesized that these mutants were using an alternative molecule for infection. Since most polyomaviruses use gangliosides, we tested whether these mutant viruses utilized gangliosides other than GM1. We used RNA interference to repress levels of GM3 synthase mRNA, an enzyme that is essential for the synthesis of GM3, a simple ganglioside from which the larger gangliosides, including GM1, GD1b, and GT1b, are derived (54). A control shRNA (shControl) or an shRNA against GM3 synthase was stably expressed in HeLa S3 cells using retrovirus transduction, and repression of GM3 synthase mRNA levels in these cells was confirmed by qRT-PCR (Fig. 8A). As expected, cells knocked down for GM3 synthase displayed significantly lower levels of surface GM1 than control cells, as assessed by CTXB binding (data not shown). We infected these cells with equal numbers of WT and mutant virions and measured

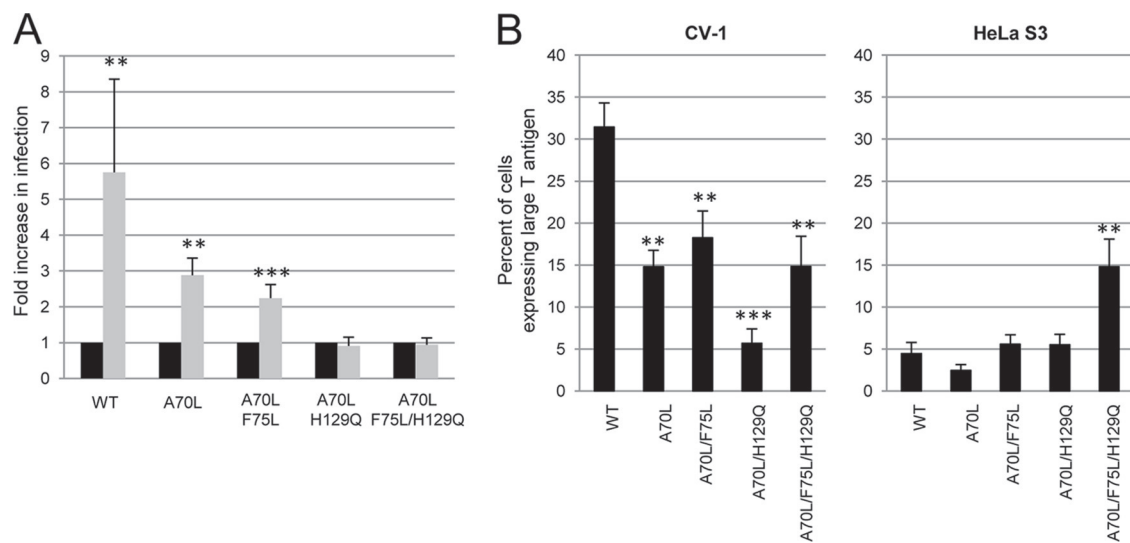


FIG 7 All three mutations are required for the change in A70L/F75L/H129Q mutant virus tropism. (A) HeLa S3 cells were left untreated (black bars) or treated with GM1 (gray bars) and infected with 5,000 virions/cell of WT and mutant virus as in Fig. 3B. Infection was measured and is presented as in Fig. 3B. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in three independent experiments. The asterisks indicate a significant difference in the fold change in infection of cells treated with GM1 relative to untreated cells (**, $P < 0.01$; ***, $P < 0.0003$). The error bars indicate standard deviations. (B) CV-1 and HeLa S3 cells were infected with 5,000 WT or mutant virions/cell. Infection was measured at 24 h p.i. and is presented as in Fig. 6A. Similar results were obtained in three independent experiments. The asterisks indicate a significant difference between infection of cells with WT SV40 and the indicated mutant (**, $P < 0.008$; ***, $P < 0.0002$).

infection by flow cytometry. Interestingly, the ability of WT and mutant viruses to infect cells expressing shGM3-2 was inhibited greater than 15-fold compared with the cells expressing shControl (Fig. 8B). A second shRNA against GM3 synthase also inhibited WT and mutant infection, as did shRNAs against glucosylceramide synthase, another enzyme essential for ganglioside synthesis (data not shown). These results implied that the mutant viruses utilized a ganglioside other than GM1. The shRNA against GM3 synthase had no effect on infection by human papillomavirus type 16 pseudovirus (data not shown), which does not use gangliosides for infection.

To confirm that the block to infection was the result of repression of ganglioside synthesis, we conducted a rescue experiment. Incubation of cells expressing shGM3-2 with GM1 overnight not only restored WT and A70L SV40 infection, but also stimulated infection 7-fold and 2-fold, respectively, compared with cells expressing control shRNA (Fig. 8B). These data indicated that the block to WT and A70L SV40 infection was specific to inhibition of ganglioside synthesis and not the result of an off-target effect of the shRNA. In contrast, even after GM1 treatment, the ability of A70L/N138Y and A70L/F75L/H129Q mutants to infect knock-down cells was still dramatically reduced relative to cells expressing shControl, as expected for viruses that do not interact with GM1. Taken together, these data demonstrated that WT SV40 requires endogenous gangliosides for infection and that the SV40 A70L/N138Y and A70L/F75L/H129Q mutant viruses require gangliosides other than GM1.

Substitutions at A70 in VP1 inhibit the ability of SV40 to induce vacuolization in CV-1 cells. One of the hallmarks of SV40 infection of some strains of permissive monkey cells is marked vacuolization at late times after infection. We previously reported that following infection by the H136Y mutant, the number of infected cells with vacuoles was significantly reduced compared

with cells infected with WT SV40 (32). In order to determine whether there was a consistent correlation between the ability of SV40 to use GM1 and form vacuoles, we infected CV-1 cells with a panel of mutant viruses at an MOI of 0.25 and examined the cells 4 days later. Figure 9A shows that the GM1 neutralization-resistant mutant viruses also had a significant defect in the ability to induce the formation of vacuoles. We observed an occasional cell containing vacuoles following infection with A70T, but no vacuolization occurred in cells infected with A70V, A70I, or any mutant virus containing the A70L substitution, including A70L/F75L/H129Q (Fig. 9A and data not shown). These results suggested that the more resistant a virus is to GM1 neutralization, the greater the defect in vacuolization. Although little vacuolization is seen in CV-1 cells infected with the defective mutants, the cells eventually lysed, and cells transfected with mutant virus DNA produced at least as many virions as cells transfected with the WT viral genome (Fig. 9B and data not shown). Taken together, these data indicated that there is a correlation between the ability of SV40 to interact with GM1 and its ability to induce the formation of vacuoles.

DISCUSSION

Altering the receptors used by viruses for cell attachment and entry can have dramatic effects on virus tropism, pathogenicity, and virulence. In this report, we demonstrated that amino acid substitutions in the receptor-binding site of SV40 VP1 altered receptor usage and tropism. We first selected viruses that were resistant to GM1 neutralization by treating a library of viable SV40 VP1 mutants with GM1 prior to infection of CV-1 cells. Mutant viruses in which the alanine at residue 70 was replaced with larger hydrophobic amino acids were resistant to GM1 neutralization. VP1 pentamers containing the A70L mutation failed to bind to GM1 on the surfaces of cells, directly demonstrating that the mutation interfered with GM1 binding. Additional biophysical studies

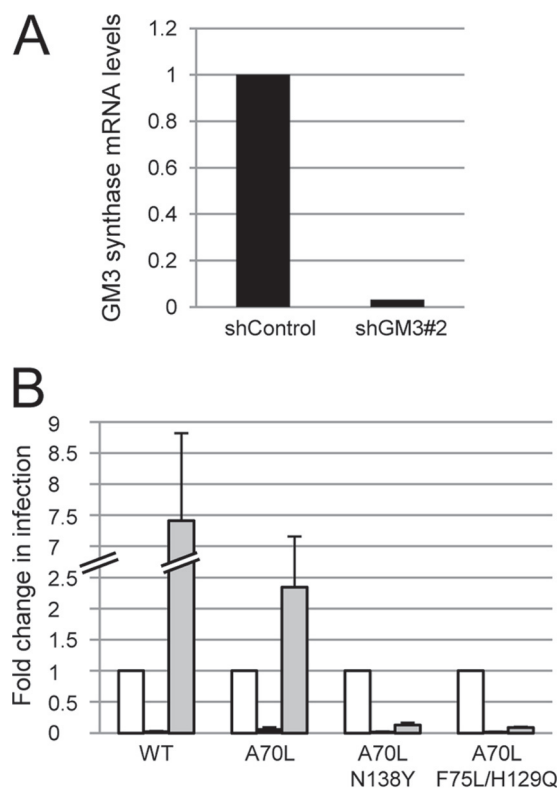


FIG 8 A70L/N138Y and A70L/F75L/H129Q mutant viruses require gangliosides for infection. (A) GM3 synthase mRNA levels in cells expressing a control shRNA and an shRNA against GM3 synthase were measured by qRT-PCR and normalized to GAPDH mRNA levels. GM3 synthase expression is reported as fold repression of the level of mRNA in shGM3-2 cells relative to shControl cells. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in two independent experiments. (B) shGM3-2 cells were left untreated (black bars) or treated with GM1 overnight (gray bars). These cells and untreated shControl cells (white bars) were then infected with 25,000 virions/cell of WT SV40 and the indicated mutant viruses. The fraction of cells expressing large T antigen was measured by immunostaining and flow cytometry at 48 h p.i. and is displayed as fold change in infection relative to untreated shControl cells. The results of a typical experiment are presented and represent the average of triplicate samples. Similar results were obtained in three independent experiments. The error bars indicate standard deviations.

showed that the A70L mutation diminishes but does not eliminate GM1 binding (unpublished results), consistent with the ability of GM1 addition to weakly stimulate infection by the mutant virus. In the wild-type VP1 pentamer, A70 directly interacts with the terminal galactose moiety of GM1 via van der Waals interactions (36). Molecular modeling and energy minimalization calculations predict that each of the allowed conformations of the mutant leucine side chain would result in steric clashes with the terminal galactose of GM1 (Fig. 10A), accounting for the binding defect and the phenotype of the mutant. The calculations further show that mutation of A70, N138, and F75 leads only to minor structural changes in the area surrounding these residues.

It is interesting that the mutant viruses that are fully resistant to GM1 neutralization can still be stimulated by GM1, although to a lesser extent than WT SV40. In fact, treatment of A70V, A70I, and A70L viruses with GM1 prior to infection actually increased infection, presumably because excess free GM1 is incorporated into cell membranes, where it can serve, albeit poorly, as a cell surface SV40

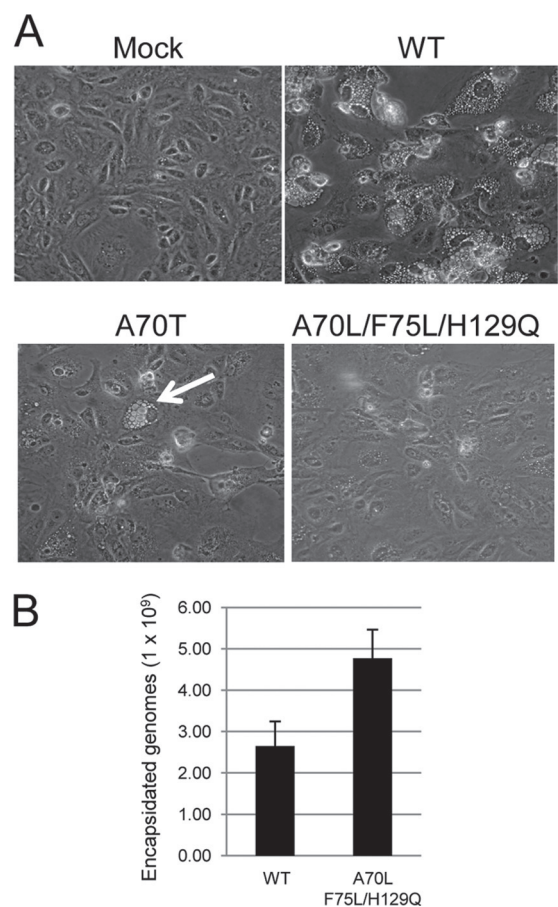


FIG 9 A70 mutant viruses have a defect in the ability to induce vacuolization of CV-1 cells. (A) A confluent dish of CV-1 cells was either mock infected or infected with SV40 at an MOI of 0.25. Phase-contrast micrographs were taken at $\times 400$ magnification at day 4 postinfection. The arrow in the A70T image indicates cells with vacuoles. (B) WT and A70L/F75L/H129Q genomes were excised from pSV766.1, religated into circles, and transfected into CV-1 cells. Eighty hours later, virus was harvested by freeze/thaw. Encapsulated genomes were then isolated from virus stocks and measured by qRT-PCR. Data are depicted as the total number of encapsulated genomes produced, and the results represent the average of triplicate samples. The error bars indicate standard deviations.

receptor. This stimulatory effect might be explained by the high valency of the virus bound to multiple molecules of GM1 in the membrane. An analysis of binding of CTXB, a bacterial toxin, to GM1 supports this suggestion. The avidity of CTXB for membrane-associated GM1 is 3 orders of magnitude higher than its affinity for monomeric GM1 in solution (12, 22, 42, 50). Thus, high avidity for cell surface GM1 may account for the ability of GM1 addition to stimulate infection by mutant viruses that display lower affinity for monomeric GM1. A similar scenario appears to occur with poliovirus mutants that are resistant to neutralization by soluble receptor (8).

To isolate mutants that were incapable of using GM1, a library constructed in an A70L mutant backbone was screened in 293TT cells, a cell line with low levels of cell surface GM1. Two mutants recovered from this screen, A70L/N138Y and A70L/F75L/H129Q, were not stimulated when GM1 was added to the cells and remained defective for GM1 neutralization and binding to GM1 on the cell surface, indicating that these mutants were no longer using

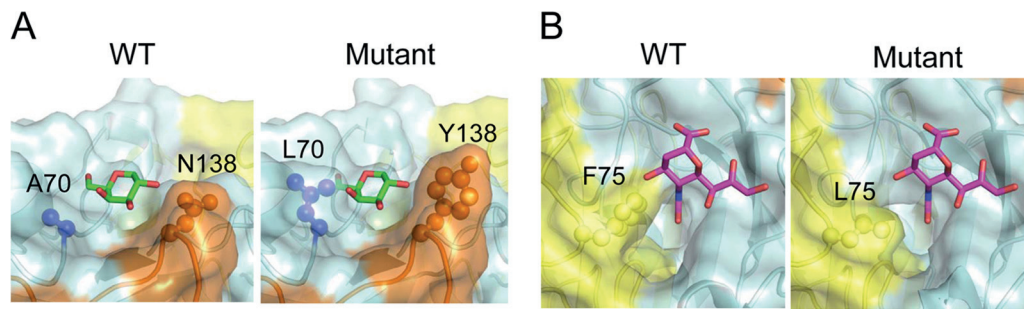


FIG 10 Potential clashes between GM1 and mutations in the VP1 binding site. (A) PyMol rendering of the galactose binding region in WT (left) and mutant (right) VP1 pentamers. The structure of the mutant pentamer was determined by energy minimizations and molecular modeling as described in Materials and Methods. The surfaces of the main monomer and the counterclockwise monomer are colored in blue and orange, respectively. The side chains of residue 70 present in the main monomer and residue 138 present in the counterclockwise monomer are depicted in ball-and-stick mode and are colored according to the monomer in which they reside. The terminal galactose moiety is displayed in stick representation, with carbon atoms labeled in green and oxygen atoms labeled in red. (B) PyMol rendering of the hydrophobic pocket for the *N*-acetyl group of neuraminic acid in WT (left) and mutant (right) pentamers, as determined in panel A. The surfaces of the main monomer and the clockwise monomer are colored in blue and yellow, respectively. The side chain of residue 75 present in the clockwise monomer is depicted in ball-and-stick mode and is colored according to the monomer in which it resides. NeuNAc is displayed in stick representation, with carbon atoms shown in magenta, oxygen atoms in red, and nitrogen atoms in blue.

GM1 for infection. Moreover, A70L/F75L/H129Q virus exhibited a striking change in cell tropism, infecting HeLa S3 and 293TT cells, cell lines with low levels of GM1, up to 4-fold better than the WT. Since A70L/F75L/H129Q virus no longer used GM1 for infection, it appears that the mutant used some other molecule as a receptor. Because genetic repression of ganglioside synthesis inhibited infection of the mutant virus, it is likely that the receptor used by this mutant is a ganglioside. Preliminary data suggest that this putative ganglioside receptor is not GT1b, GD1b, GD1a, or GM2, because supplementation of HeLa S3 cells with those gangliosides failed to stimulate infection by the mutant (unpublished results). Future experiments are required to identify the novel receptor utilized by the GM1 binding-defective SV40 mutants. We hypothesize that the triple-mutant virus infects certain cells better than WT SV40 because this receptor is more abundant on these cells than GM1. This change in tropism is not species specific, because the mutant viruses fail to infect other human cells.

A70L/F75L/H129Q and A70L/N138Y mutants both contained additional substitutions at residues that directly interact with GM1. Like A70, N138 directly interacts with the terminal galactose of GM1 via van der Waals forces. Modeling and energy minimization predict that the bulky tyrosine side chain of the mutant will approach the hydroxyl groups of the terminal galactose moiety and impinge upon the binding pocket (Fig. 10A). H129 does not directly interact with GM1, but it lies adjacent to A70 in the quaternary pentamer structure ($C\beta$ - $C\beta$ distance, 4.8 Å) (Fig. 2). The H129Q mutation presumably perturbs the position of A70, which directly contacts GM1, thus further impairing GM1 usage. F75, which is invariant in BKV and JCV VP1 proteins (Fig. 11), is one of several residues that line a deep hydrophobic cavity that accommodates the acetyl group of NeuNAc (Fig. 10B) (36). This pocket is substantially larger in SV40 pentamers than in the VP1 pentamers of other polyomaviruses (34, 36, 45), a difference that may explain the preferential binding of SV40 to the monkey NeuNGc-GM1, which has a hydroxyl group attached to the methyl group and is believed to be the natural receptor for SV40 (6). Modeling and energy minimalization reveal that the replacement of the aromatic side chain at position 75 with leucine enlarges this hydrophobic cavity (Fig. 10B). We hypothesize that the

cavity may now be able to accommodate a larger carbohydrate moiety, thereby expanding the repertoire of molecules that can serve as receptors. F75L and H129Q each individually increased the infectivity of the A70L mutant in HeLa S3 cells, but the most pronounced change in tropism was exhibited by the triple mutant, suggesting that both of these substitutions are required for optimal usage of an as-yet-unidentified receptor. The requirement for all three substitutions in the A70L/F75L/H129Q mutant virus for its change in tropism highlights the power of the genetic approach used here. Because the phenylalanine at position 75 is invariant and the histidine at position 129 is not directly located in the receptor-binding site, it was not possible to predict the constellation of substitutions that would alter ganglioside usage and SV40 tropism.

Structural and biological studies by Benjamin and colleagues demonstrated that amino acid substitutions in the glycan binding region of mPyV VP1 were sufficient to alter receptor specificity and cell tropism. By preventing recognition of branched-chain oligosaccharide pseudoreceptors without affecting binding to the functional straight-chain receptor, the G91E substitution in mPyV VP1 increases virus spread and broadens virus tropism in the murine host, which results in increased pathogenicity (4, 14). Moreover, the V296A substitution, which is found in the LID laboratory strain of mPyV that also contains the G91E mutation, further increases virus spread and expands tropism by reducing avidity for the functional receptor, resulting in an extremely virulent virus (3). In contrast, the A70L/F75L/H129Q substitutions in SV40 VP1 appear to alter cell tropism by reducing affinity for the functional GM1 receptor and by allowing recognition of a new receptor. Thus, a genetic analysis of the receptor usage of SV40, which is more closely related to clinically relevant BKV and JCV than is mPyV, highlights an alternative mechanism by which polyomavirus tropism can be altered by substitutions in the glycan binding site.

Most of the amino acid differences among the VP1 sequences from SV40, BKV, and JCV are located in the BC and DE loops, two of three regions in VP1 that constitute the oligosaccharide binding site (Fig. 11). Our results show that a limited number of amino acid substitutions at the GM1 binding site are sufficient to deter-

viruses with their receptors and how these interactions can be altered is essential to combating virus infection and spread.

The ability to induce vacuolization in permissive monkey cells is a signature characteristic of SV40. In fact, when the virus was first isolated from contaminated polio vaccine stocks in 1960, the induction of vacuolization in African green monkey cells was one of the first phenotypes of the virus to be recognized (47), and an alternative early name for SV40 was simian vacuolating virus. Since its discovery, SV40 has become one of the most intensely studied animal viruses, but little is known about how and why SV40 induces vacuolization. Previously, we showed that the H136Y SV40 mutant failed to induce robust vacuolization in CV-1 cells (31). This mutant was also resistant to GM1 neutralization. Interestingly, the GM1-defective mutants isolated from the two screens conducted here also failed to induce vacuolization. Moreover, there appears to be a gradation in the defect, as rare cells infected with the partially GM1-resistant A70T and H136Y viruses are able to form vacuoles. These data suggest that the interaction of VP1 with GM1 in cells is required for vacuolization, but the molecular basis for this effect is unknown. Thus, more than 50 years later, there is much to be learned about SV40.

ACKNOWLEDGMENTS

We thank Bärbel Blaum for her insight and helpful discussion on concepts and experiments outlined in the manuscript, Genaro Pimienta-Rosales for assistance with PyMol, and Jan Zulkowski for assistance in preparing the manuscript.

T.G.M. was supported in part by a training grant from the National Institutes of Health (T32-HD007149). This work was supported by grants from the National Cancer Institute to D.D. (P01 CA016038) and to R.L.G. (R01 CA37667). T.S., M.H.C.B., and U.N. acknowledge support from the Deutsche Forschungsgemeinschaft (SFB-685).

REFERENCES

- Adams PD, et al. 2011. The Phenix software for automated determination of macromolecular structures. *Methods* 55:94–106.
- Ashok A, Atwood WJ. 2006. Virus receptors and tropism. *Adv. Exp. Med. Biol.* 577:60–72.
- Bauer PH, et al. 1995. Genetic and structural analysis of a virulence determinant in polyomavirus VP1. *J. Virol.* 69:7925–7931.
- Bauer PH, et al. 1999. Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse. *J. Virol.* 73:5826–5832.
- Bird G, O'Donnell M, Moroianu J, Garcea RL. 2008. Possible role for cellular karyopherins in regulating polyomavirus and papillomavirus capsid assembly. *J. Virol.* 82:9848–9857.
- Campanero-Rhodes MA, et al. 2007. N-glycolyl GM1 ganglioside as a receptor for simian virus 40. *J. Virol.* 81:12846–12858.
- Chinnapen DJ, Chinnapen H, Saslowsky D, Lencer WI. 2007. Rafting with cholera toxin: endocytosis and trafficking from plasma membrane to ER. *FEMS Microbiol. Lett.* 266:129–137.
- Colston E, Racaniello VR. 1994. Soluble receptor-resistant poliovirus mutants identify surface and internal capsid residues that control interaction with the cell receptor. *EMBO J.* 13:5855–5862.
- Eash S, Querbes W, Atwood WJ. 2004. Infection of Vero cells by BK virus is dependent on caveolae. *J. Virol.* 78:11583–11590.
- Eddy BE, Borman GS, Grubbs GE, Young RD. 1962. Identification of the oncogenic substance in rhesus monkey kidney cell culture as simian virus 40. *Virology* 17:65–75.
- Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60:2126–2132.
- Ewers H and A Helenius. 2011. Lipid-mediated endocytosis. *Cold Spring Harbor Perspect. Biol.* 3:a004721. doi:10.1101/cshperspect.a004721.
- Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–1100.
- Freund R, Garcea RL, Sahli R, Benjamin TL. 1991. A single-amino-acid substitution in polyomavirus VP1 correlates with plaque size and hemagglutination behavior. *J. Virol.* 65:350–355.
- Gilbert J, Benjamin T. 2004. Uptake pathway of polyomavirus via ganglioside GD1a. *J. Virol.* 78:12259–12267.
- Goodwin EC, et al. 2011. BiP and multiple DNAJ molecular chaperones in the endoplasmic reticulum are required for efficient simian virus 40 infection. *MBio* 2:101–111.
- Goodwin EC, Naeger LK, Breiding DE, Androphy EJ, DiMaio D. 1998. Transactivation-competent bovine papillomavirus E2 protein is specifically required for efficient repression of human papillomavirus oncogene expression and for acute growth inhibition of cervical carcinoma cell lines. *J. Virol.* 72:3925–3934.
- Gorelik L, et al. 2011. Progressive multifocal leukoencephalopathy (PML) development is associated with mutations in JC virus capsid protein VP1 that change its receptor specificity. *J. Infect. Dis.* 204:103–114.
- Johung K, Goodwin EC, DiMaio D. 2007. Human papillomavirus E7 repression in cervical carcinoma cells initiates a transcriptional cascade driven by the retinoblastoma family, resulting in senescence. *J. Virol.* 81:2102–2116.
- Kartenbeck J, Stukenbrok H, Helenius A. 1989. Endocytosis of simian virus 40 into the endoplasmic reticulum. *J. Cell Biol.* 109:2721–2729.
- Krumbholz A, Bininda-Emonds OR, Wutzler P, Zell R. 2009. Phylogenetics, evolution, and medical importance of polyomaviruses. *Infect. Genet. Evol.* 9:784–799.
- Kuziemko GM, Stroh M, Stevens RC. 1996. Cholera toxin binding affinity and specificity for gangliosides determined by surface plasmon resonance. *Biochemistry* 35:6375–6384.
- Li F, Li W, Farzan M, Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309:1864–1868.
- Liddington RC, et al. 1991. Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354:278–284.
- Liu Y, et al. 2010. Altered receptor specificity and cell tropism of D222G hemagglutinin mutants isolated from fatal cases of pandemic A(H1N1) 2009 influenza virus. *J. Virol.* 84:12069–12074.
- Low JA, Magnuson B, Tsai B, Imperiale MJ. 2006. Identification of gangliosides GD1b and GT1b as receptors for BK virus. *J. Virol.* 80:1361–1366.
- Luo C, Hirsch HH, Kant J, Randhawa P. 2012. VP-1 quasispecies in human infection with polyomavirus BK. *J. Med. Virol.* 84:152–161.
- Magaldi TG, et al. 2012. Primary human cervical carcinoma cells require human papillomavirus E6 and E7 expression for ongoing proliferation. *Virology* 422:114–124.
- McNees AL, et al. 2009. SV40 lymphomagenesis in Syrian golden hamsters. *Virology* 384:114–124.
- Meroz D, et al. 2011. Putative amino acid determinants of the emergence of the 2009 influenza A (H1N1) virus in the human population. *Proc. Natl. Acad. Sci. U. S. A.* 108:13522–13527.
- Murata H, Peden K, and Lewis AM, Jr. 2008. Identification of a mutation in the SV40 capsid protein VP1 that influences plaque morphology, vacuolization, and receptor usage. *Virology* 370:343–351.
- Murata H, Teferedegne B, Sheng L, Lewis AM, Jr, Peden K. 2008. Identification of a neutralization epitope in the VP1 capsid protein of SV40. *Virology* 381:116–122.
- Naeger LK, et al. 1999. Bovine papillomavirus E2 protein activates a complex growth-inhibitory program in p53-negative HT-3 cervical carcinoma cells that includes repression of cyclin A and cdc25A phosphatase genes and accumulation of hypophosphorylated retinoblastoma protein. *Cell Growth Differ.* 10:413–422.
- Neu U, et al. 2010. Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif. *Cell Host Microbe* 8:309–319.
- Neu U, Stehle T, Atwood WJ. 2009. The Polyomaviridae: contributions of virus structure to our understanding of virus receptors and infectious entry. *Virology* 384:389–399.
- Neu U, Woellner K, Gauglitz G, Stehle T. 2008. Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc. Natl. Acad. Sci. U. S. A.* 105:5219–5224.
- Pelkmans L, Kartenbeck J, Helenius A. 2001. Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the ER. *Nat. Cell Biol.* 3:473–483.

38. Pho MT, Ashok A, Atwood WJ. 2000. JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. *J. Virol.* **74**: 2288–2292.
39. Qian M, Cai D, Verhey KJ, Tsai B. 2009. A lipid receptor sorts polyomavirus from the endolysosome to the endoplasmic reticulum to cause infection. *PLoS Pathog.* **5**:e1000465. doi:10.1371/journal.ppat.1000465.
40. Richardson JM, Milton MJ, Homans SW. 1995. Solution dynamics of the oligosaccharide moiety of ganglioside GM1: comparison of solution conformations with the bound state conformation in association with cholera toxin B-pentamer. *J. Mol. Recognit.* **8**:358–362.
41. Sapp M, Day PM. 2009. Structure, attachment and entry of polyoma- and papillomaviruses. *Virology* **384**:400–409.
42. Schon A, Freire E. 1989. Thermodynamics of intersubunit interactions in cholera toxin upon binding to the oligosaccharide portion of its cell surface receptor, ganglioside GM1. *Biochemistry* **28**:5019–5024.
43. Schrodinger LLC. 2010. The PyMOL molecular graphics system, version 1.3r1. Schrodinger LLC, New York, NY.
44. Stehle T, Gamblin SJ, Yan Y, Harrison SC. 1996. The structure of simian virus 40 refined at 3.1 Å resolution. *Structure* **4**:165–182.
45. Stehle T, Harrison SC. 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J.* **16**:5139–5148.
46. Sunyaev SR, Lugovskoy A, Simon K, Gorelik L. 2009. Adaptive mutations in the JC virus protein capsid are associated with progressive multifocal leukoencephalopathy (PML). *PLoS Genet.* **5**:e1000368. doi:10.1371/journal.pgen.1000368.
47. Sweet BH, Hilleman MR. 1960. The vacuolating virus, S.V. 40. *Proc. Soc. Exp. Biol. Med.* **105**:420–427.
48. Thomas MA, Lichtenstein DL, Krajcsi P, Wold WS. 2007. A real-time PCR method to rapidly titer adenovirus stocks. *Methods Mol. Med.* **130**: 185–192.
49. Tsai B, et al. 2003. Gangliosides are receptors for murine polyoma virus and SV40. *EMBO J.* **22**:4346–4355.
50. Turnbull WB, Precious BL, Homans SW. 2004. Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry. *J. Am. Chem. Soc.* **126**:1047–1054.
51. van Doremalen N, et al. 2011. A single amino acid in the HA of pH1N1 2009 influenza virus affects cell tropism in human airway epithelium, but not transmission in ferrets. *PLoS One* **6**:e25755. doi:10.1371/journal.pone.0025755.
52. Varki A. 2001. Loss of N-glycolylneuraminic acid in humans: mechanisms, consequences, and implications for hominid evolution. *Am. J. Phys. Anthropol. Suppl.* **33**:54–69.
53. Yates KE, Korb GA, Shtutman M, Roninson IB, DiMaio D. 2008. Repression of the SUMO-specific protease Senp1 induces p53-dependent premature senescence in normal human fibroblasts. *Aging Cell* **7**:609–621.
54. Yu RK, Bieberich E, Xia T, Zeng G. 2004. Regulation of ganglioside biosynthesis in the nervous system. *J. Lipid Res.* **45**:783–793.