Structural Analysis and Glycan Receptor Binding Specificities of Human Polyomaviruses

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Luisa Johanna Ströh

aus Tübingen

Tübingen

2015

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

Tag der mündlichen Qualifikation:	12.06.2015
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Thilo Stehle
2. Berichterstatter:	Prof. Dr. Thomas Iftner
3. Berichterstatter:	Prof. Dr. Martin Sapp

Table of Contents

Abbreviations

3'SL	α2,3-sialyllactose
3'SLN	α2,3-sialyllactosamine
5-HT	5-hydroxytryptamine, serotonin
5-HTR	Serotonin receptor
6'SL	α2,6-sialyllactose
AIDS	Acquired Immunodeficiency Syndrome
BKPyV	BK Polyomavirus
$CC_{1/2}$	Split-half correlation coefficient
ccw	Counterclockwise
cw	Clockwise
CFG	Consortium for Functional Glycomics
Da, kDa	Dalton, kilodalton
DNA	Desoxyribonucleic acid
dsDNA	Double-stranded DNA
E. coli	*Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
FBDD	Fragment-based drug discovery
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	*N*-acetyl-galactosamine
Glc	Glucose
GlcNAc	*N*-acetyl-glucosamine
GST	Glutathione-S-transferase
HBGA	Histo-blood group antigen
HIV	Human Immunodeficiency Virus
HPyV	Human Polyomavirus
His-tag	Hexahistidine-tag
HPyV6	Human Polyomavirus 6
HPyV7	Human Polyomavirus 7
HPyV9	Human Polyomavirus 9
HPyV10	Human Polyomavirus 10
HPyV12	Human Polyomavirus 12
HTS	High-throughput screening

ITC	Isothermal titration calorimetry
JCPyV	JC Polyomavirus
K_D	Dissociation constant
KIPyV	Karolinska Institut Polyomavirus
LSTc	Sialylneolacto-*N*-tetraose c, lactoseries tetrasaccharide c
LT	Large T-antigen
MCPyV	Merkel Cell Polyomavirus
MPyV	Murine Polyomavirus
MR	Molecular replacement
MW	Molecular weight
MWPyV	Malawi Polyomavirus
Neu5Ac	5-*N*-acetyl-neuraminic acid
Neu5Gc	5-*N*-glycolyl-neuraminic acid
NJPyV	New Jersey Polyomavirus
NMR	Nuclear magnetic resonance
PDB	Protein data bank
PML	Progressive Multifocal Leukoencephalopathy
PSV	Pseudovirus
PyV	Polyomavirus
r.m.s.d.	Root mean square deviation
RNA	Ribonucleic acid
SLS	Swiss Light Source
SSM	Secondary-structure matching
sT	Small T-antigen
STD	Saturation transfer difference
STLPyV	Saint Louis Polyomavirus
SV40	Simian Virus 40
TS	*Trichodysplasia spinulosa*
TSPyV	*Trichodysplasia spinulosa*-associated Polyomavirus
VLP	Virus-like particle
VP1	Virus Protein 1
VP2, VP3, VP4	Virus Protein 2
VP3	Virus Protein 3
WUPyV	Washington University Polyomavirus

The Svennerholm nomenclature was used for gangliosides.

Abstract

This thesis investigates the atomic structures and cell surface glycan receptor specificities of six human polyomaviruses. In healthy individuals, polyomaviruses establish usually asymptomatic chronic infections, but they may cause severe disease in immunocompromised patients. High-resolution X-ray structures of the major capsid protein VP1 alone and in complex with its specific glycan receptors allowed to analyze the molecular basis for underlying recognition events during early steps of the infection. The findings from crystallographic studies were corroborated with cell-based binding experiments using flow cytometry, interaction studies in solution by NMR spectroscopy, and infection studies in cell culture. The presented results highlight the enormous complexity of virus-glycan interactions and demonstrate that subtle differences in both the viral attachment protein and the cell-surface glycan receptor modulate binding specificities and affinities and thus, are key determinants for tissue and host tropism, viral infectivity, and pathogenesis.

JC Polyomavirus (JCPyV) causes the fatal demyelinating disease Progressive Multifocal Leukoencephalopathy (PML) in immunocompromised individuals. A brain isolate of JCPyV, a genotype 1 strain, requires α2,6-linked sialic acids on the LSTc glycan for attachment to host cells, whereas a kidney isolate, a genotype 3 strain, was reported to interact with gangliosides featuring α2,3- and α2,8-linked sialic acids. Comprehensive structural and functional analyses of these two representative strains and their glycan receptor specificities show that engagement of LSTc is a prerequisite for functional receptor engagement for all seven JCPyV genotypes while the weaker-binding gangliosides are not required for infection. Interestingly, the majority of JCPyV isolates from PML patients contain distinct mutations within or in proximity to the LSTc binding site on VP1. The presented results reveal that binding of these mutant viruses to glycans is abolished or severely compromised rendering them not infectious. Thus, these viruses likely utilize a so far unknown receptor for the infectious entry and/or play alternative roles in PML pathogenesis.

In order to explore potential strategies for the development of antiviral compounds against PML a fragment-based screening approach was carried out and subsequent X-ray structure analysis identified a novel compound binding site inside the hydrophobic cavity of the JCPyV VP1 pentamer. Further studies show that modifications to the five-fold pore of the VP1 pentamer result in a severe reduction of infectivity, suggesting that the pore is an important structural feature of polyomaviruses. Thus, targeting this pore may be proven to be an effective antiviral therapy.

Crystal structures of VP1 from *Trichodysplasia spinulosa*-associated Polyomavirus (TSPyV) in complex with three different glycans reveal a sialic acid binding site that is shifted by about

18 Å from the ‘classical’ sialic acid binding sites of JCPyV and other polyomaviruses. Functional and cell-based studies confirm the importance of these novel interactions with sialic acids and suggest that glycolipids play an important role during TSPyV infection. Surprisingly, this new sialic acid binding site is also conserved in VP1 of Human Polyomavirus 12 (HPyV12), whereas the human New Jersey Polyomavirus (NJPyV) employs a third location for the recognition of an Neu5Ac-α2,3-Gal-containing receptor. A structure-based phylogenetic analysis suggests that TSPyV, HPyV12, and NJPyV share their sialic acid binding site with closely related non-human polyomaviruses providing initial clues about determinants of host specificity and evolution of these viruses. In contrast, Human Polyomavirus 6 and 7 (HPyV6 and HPyV7, respectively) carry uniquely elongated loops that cover the bulk of the outer virion surfaces, and moreover, obstruct the groove that binds sialylated glycan receptors in related viruses. Consistently, cell attachment and NMR studies further suggest that sialylated glycans are not required for cell attachment of both, HPyV6 and HPyV7.

In conclusion, with its relatively high sequence homology and a conserved overall architecture, the growing *Polyomaviridae* family forms an excellent platform to analyze principles and molecular determinants of receptor specificity and antigenicity as well as critical factors for viral pathogenesis. A detailed understanding of the underlying molecular principles is important to establish a comprehensive toolbox, which can be used for new approaches for antiviral therapies and for the design of therapeutic gene vectors.

Zusammenfassung

Diese Dissertation untersucht die atomaren Strukturen von sechs menschlichen Polyomaviren und das spezifische Andocken dieser Viren an Zuckerstrukturen auf Zelloberflächen der Wirtszellen. Bei gesunden Menschen verursachen Polyomaviren meist chronische asymptomatische Infektionen, jedoch können einige dieser Viren bei immungeschwächten Patienten schwere Erkrankung auslösen. Mit Hilfe von Röntgenstrukturen der Kapsidproteine in Komplexen mit Zuckerverbindungen konnten die molekularen Grundlagen für das spezifische Andocken dieser Viren an Zellen identifiziert werden. Anschließend wurden die Erkenntnisse aus den kristallographischen Studien mit zellbasierten Bindungsexperimenten mittels Durchflusszytometrie, Interaktionsstudien mittels NMR-Spektroskopie in Lösung und Infektionsstudien in Zellkultur überprüft und weiter analysiert. Die erhaltenen Ergebnisse unterstreichen die enorme Komplexität von Interaktionen zwischen Viren und Zuckerstrukturen auf der Zelloberfläche und zeigen, dass feine Unterschiede in den viralen Proteine oder den gebundenen Zuckermolekülen eine wichtige Rolle für die Erkennungsprozess und die Bindungsaffinität spielen. Diese spezifischen und gut regulierten Wechselwirkungen zwischen Virus und Rezeptoren auf der Wirtszelle sind nicht nur für die initiale Erkennung wichtig sondern auch die viralen Infektiosität und Pathogenese.

JC Polyomavirus (JCPyV) kann bei immungeschwächten Personen die tödliche Gehirnerkrankung Progressive Multifokale Leukoenzephalopathie (PML) auslösen. Ein JCPyV Stamm aus dem Gehirn bindet an α2,6-verknüpfte Sialinsäure der linearen Zuckerstruktur LSTc auf der Wirtszelle, während Untersuchungen für einen anderen JCPyV Stamm, der aus der Niere isoliert wurde, zeigten, dass auch Gangliosiden mit α2,3- oder α2,8- verknüpfter Sialinsäure gebunden werden. Eine umfassende strukturelle und funktionelle Analyse der zwei repräsentative Virusstämme zeigt hier jedoch deutlich, dass die Bindung an LSTc eine Voraussetzung für die Infektion für alle sieben JCPyV Genotypen ist, während die schwächer gebundenen Ganglioside nicht für die Infektion erforderlich sind. Interessanterweise besitzen JCPyV Isolate von PML-Patienten oft eine kleine Anzahl von konservierten Mutationen innerhalb oder in unmittelbarer Nähe der LSTc-Bindungsstelle auf dem Kapsidprotein. Die hier vorgestellten Ergebnisse zeigen, dass die Bindung an Zuckerstrukturen durch die Veränderungen in VP1 verhindert oder stark beeinträchtigt ist und daher diesen Viren nicht mehr infektiös sind. Somit verwenden diese PML-assoziierten Viren wahrscheinlich ein bisher unbekanntes Molekül auf der Zelloberfläche für den infektiösen Eintritt und/oder sie tragen in einer anderen Weise zur PML-Pathogenese bei.

Ein Fragment-basiertes Screening wurde durchgeführt um mögliche Strategien für die Entwicklung von antiviralen Verbindungen gegen PML zu testen. Die Bindung von einem

Fragment im Innenraum des hydrophoben VP1 Pentamers wurde mittels Strukturanalyse bestätigt. Zusätzlich konnte in einer Untersuchung gezeigt werden, dass Änderungen der fünffachen Pore der VP1-Pentamere die Virusinfektion verringern. Diese Erkenntnisse deuten darauf hin, dass die Poren ein wichtiger struktureller Bestandteil des Viruskapsids ist. Somit stellt auch die Pore oder gar der hydrophobe Innenraum des VP1 Pentamers einen möglichen Angriffsort für die Entwicklung von wirksamen antiviralen Therapien dar.

Kristallstrukturen von VP1 des *Trichodysplasia spinulosa*-assoziierten Polyomavirus (TSPyV) im Komplex mit drei verschiedenen Zuckermolekülem zeigen eine Sialinsäure-Bindungsstelle, die um etwa 18 Å von der "klassischen" Sialinsäure-Bindungsstellen von JCPyV und anderer Polyomaviren entfernt ist. Funktionelle Untersuchungen bestätigen die Wichtigkeit dieser neuartigen Sialinsäure-Interaktion für die TSPyV Infektion und lassen zudem vermuten, dass besonders Glycolipide hierbei eine wichtige Rolle einnehmen. Diese neue Bindungsstelle ist auch im Kapsid des menschlichen Polyomavirus 12 (HPyV12) strukturell konserviert, während der menschliche New Jersey Polyomavirus (NJPyV) in einem dritten Bereich von VP1 mit einer spezifischer Zuckerstruktur wechselwirkt. Strukturbasierte phylogenetische Analysen zeigen, dass die Sialinsäure-Bindungsstellen von TSPyV, HPyV12 und NJPyV vermutlich auch in verwandten aber nicht-menschlichen Polyomaviren vorhanden sind und geben Hinweise über Determinanten der Wirtsspezifität und Evolution dieser Viren. Die humanen Polyomaviren 6 und 7 (HPyV6 und HPyV7) besitzen im Gegensatz dazu verlängerte VP1-Schleifenstrukturen, die den Großteil der Kapsidoberflächen abdecken und das Erkennen von Zuckermolekülen verhindert. Zusätzlich zeigen NMR-spektroskopische Untersuchungen, dass sialylierte Zucker nicht für das initiale Andocken von HPyV6 und HPyV7 an Zellen erforderlich sind.

Zusammenfassend lässt sich sagen, das die wachsende Familie der Polyomaviren mit einer hohen Sequenzhomologie und konservierten Gesamtarchitektur eine hervorragende Plattform bildet, um Einblicke in grundlegende Prinzipien und molekularen Determinanten für Rezeptorspezifität, viraler Antigenität und Pathogenese zu erhalten. Ein detailliertes Verständnis der molekularen Grundlagen ist wichtig für neue Ansätze in der Entwicklung von antiviralen Therapien und auch therapeutische Genvektoren.

Publications

Parts of this thesis have already been published.

Ströh, L.J., Gee, G.V. , Blaum, B.S., Dugan, A.S., Feltkamp, C.W., Atwood, W.J. and Stehle, T. (2015). *Trichodysplasia spinulosa*-associated polyomavirus employs a displaced binding site on VP1 to engage sialylated glycolipids. PLoS Pathogens *11*, e1005112.

Ströh, L.J. *, Maginnis, M.S. *, Blaum, B.S., Nelson, C.D., Neu, U., Gee, G.V., O'Hara, B.A., Motamedi, N., DiMaio, D., Atwood, W.J. and Stehle, T. (2015). Increased affinity of JC polyomavirus capsid for LSTc over other sialylated glycans is a major determinant of infectivity. Journal of Virology *89*, 6364-6375.
* Authors contributed equally to this work.

Nelson, C.D., **Ströh, L.J.**, Gee, G.V., O'Hara, B.A., Stehle, T., and Atwood, W.J. (2015). Modulation of a symmetry-related pore in the capsid of JC polyomavirus reduces infectivity and prevents exposure of the minor capsid proteins. Journal of Virology *89*, 3910-3921.

Ströh, L.J., Neu, U., Blaum, B.S., Buch, M.H.C., Garcea, R.L., and Stehle, T. (2014). Structure analysis of the major capsid proteins of human polyomaviruses 6 and 7 reveals an obstructed sialic acid binding site. Journal of Virology *88*, 10831-10839.

Ströh, L.J., and Stehle, T. (2014). Glycan engagement by viruses: receptor switches and specificity. Annual Review of Virology *1*, 285-306.

Neu, U., Allen, S.A., Blaum, B.S., Liu, Y., Frank, M., Palma, A.S., **Ströh, L.J.**, Feizi, T., Peters, T., Atwood, W.J., and Stehle, T. (2013). A structure-guided mutation in the major capsid protein retargets BK polyomavirus. PLoS Pathogens *9*, e1003688.

Maginnis, M.S. *, **Ströh, L.J.** *, Gee, G.V., O'Hara, B.A., Derdowski, A., Stehle, T., and Atwood, W.J. (2013). Progressive multifocal leukoencephalopathy-associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding. mBio *4*, e00247-00213.
* Authors contributed equally to this work.

Neu, U., Maginnis, M.S., Palma, A.S., **Ströh, L.J.**, Nelson, C.D., Feizi, T., Atwood, W.J., and 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.

Furthermore, I contributed to the following publication.

Forouzan, D., Ammelburg, M., Hobel, C.F., **Ströh, L.J.**, Sessler, N., Martin, J., and Lupas, A.N. (2012). The archaeal proteasome is regulated by a network of AAA ATPases. Journal of Biological Chemistry *287*, 39254-39262.

All relevant coordinates and structure factor amplitudes have been deposited in the RCSB Protein Data Bank.

1. Introduction

Viruses have been known as distinct biological entities since the end of the 19th century. They are the most abundant entities on our planet and infect all life forms, from bacteria and archaea to animals and plants. Although viruses themselves have no fossil records in the conventional sense, it seems that they have left their traces throughout the history of life. The global effort to understand and control these small agents responsible for many human diseases is rising and constantly being fostered by epidemic outbreaks through history until today, for example the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2002/03 in China, influenza pandemics such as the “swine flu” in 2009, and the Ebola virus outbreak in 2014. The development of new technologies, such as nucleic acid sequencing, but also discoveries and advances in closely related fields including molecular and cell biology, immunology, structural biology, and high-throughput methods such as fragment-based drug discovery (FBDD) for potentially active antiviral molecules have led towards a better understanding of viruses and the development of vaccines and antiviral therapies.
Prior to infection, viruses attach to specific receptors on the surface of their host cells in order to initiate cellular entry. Not only cellular attachment but also viral release and spread depend on precisely regulated interactions between viral proteins and their cognate receptors on the host cell surface. A detailed understanding of these interactions, which ultimately determine viral spread and tissue tropism, viral pathogenesis and host tropism, therefore provides a powerful platform for the development of antiviral strategies.

1.1 Polyomaviruses

Polyomaviruses are small non-enveloped viruses with a circular double stranded (ds) DNA genome. They were initially described in 1953 in mice by Ludwig Gross (Gross, 1953) but have been found in birds and several mammals, including humans, since then (DeCaprio and Garcea, 2013; Feltkamp et al., 2013). In addition, a fish-associated polyomavirus has been reported recently (Peretti et al., 2015). Their name is derived from Greek roots *poly-*, which means “many”, and *-oma*, which means “tumors”, due to the fact that the murine polyomavirus (MPyV) causes tumor formation when inoculated into mice. Between 1955 and 1963 lots of inactivated oral poliovirus vaccine that were contaminated with infectious amounts of another well-known member of the polyomavirus family, Simian Virus 40 (SV40), were administered to residents of the United States of America (Shah and Nathanson, 1976). SV40 infections are usually asymptomatic in rhesus macaques, but the virus has oncogenic potential in small animal models such as hamsters and transforming ability in cell culture. Thus, concerns about SV40-induced tumors in humans were raised but could not be proven (Shah, 2007). However,

more than 60 years of research focusing on molecular biology of the two model tumor viruses MPyV and SV40 resulted in fundamental discoveries about eukaryotic gene expression (Banerji et al., 1981; Benoist and Chambon, 1981; Fromm and Berg, 1983), DNA replication (Li and Kelly, 1984; Waga and Stillman, 1994) and oncogenic transfection (Lane and Crawford, 1979; Linzer and Levine, 1979).

The first two human polyomaviruses, JC polyomavirus (JCPyV) and BK polyomavirus (BKPyV), were identified in 1971 and are named after the initials of case patients (Gardner et al., 1971; Padgett et al., 1971). JCPyV is the causative agent of Progressive Multifocal Leukoencephalopathy (PML), which is a fatal brain disease in patients with Human Immunodeficiency Virus/ Acquired Immune Deficiency Syndrome (HIV/ AIDS) and more recently, in patients with multiple sclerosis (MS) who have been treated with immunomodulating drugs (Tan and Koralnik, 2010; Bellizzi et al., 2013; Diotti et al., 2013). Uncontrolled BKPyV infections contribute to polyomavirus-associated nephropathy (PVAN) and hemorrhagic cystitis in renal transplant recipients (Bennett et al., 2012; Rinaldo et al., 2013). JCPyV and BKPyV were the only known human polyomaviruses for almost 40 years, but in the last decade improved genetic screening and high-throughput sequencing techniques led to the identification of several previously unknown polyomaviruses including species frequently circulating within the human population (Figure 1 and Table 1). One of them, the Merkel Cell Polyomavirus (MCPyV), is associated with Merkel Cell Carcinoma (MCC), a rare but very aggressive skin tumor in immunosuppressed and elderly individuals. MCPyV was the first known polyomavirus causing cancer in humans (Feng et al., 2008).

The expansion of the *Polyomaviridae* family, including many human and non-human members, provoked its division into three genera, the Ortho-, Wuki-, and Avipolyomaviruses (Johne et al., 2011). Polyomavirus species must have less than 81% nucleotide sequence similarities to another according to the current guidelines of the Committee on Taxonomy of Viruses (ICTV) (Johne et al., 2011).

Seroprevalence studies have shown that polyomaviruses infect a large proportion of the human population worldwide (Table 1). The seroprevalence increases typically during childhood, suggesting that primary exposures occur early in life (Kean et al., 2009). Serological cross-reactivity appears between very closely related polyomaviruses such as SV40 and BKPyV (81% amino acid sequence identity of their major capsid protein VP1) but has been observed only to a minor extent between JCPyV and SV40 or BKPyV (≤ 79% VP1 amino acid sequence identity) (Viscidi and Clayman, 2006; Kean et al., 2009). However, the common human seroreactivity to African Green Monkey-derived Lymphotropic Polyomavirus (LPyV) found its explanation in the discovery of the Human Polyomavirus 9 (HPyV9) many years later (Sauvage et al., 2011; Scuda et al., 2011). In addition, serological studies could already predict the presence of a human polyomavirus species that is very similar to Chimpanzee

Polyomavirus (ChPyV) prior to the identification of New Jersey Polyomavirus (NJPyV) (Scuda et al., 2013; Mishra et al., 2014).

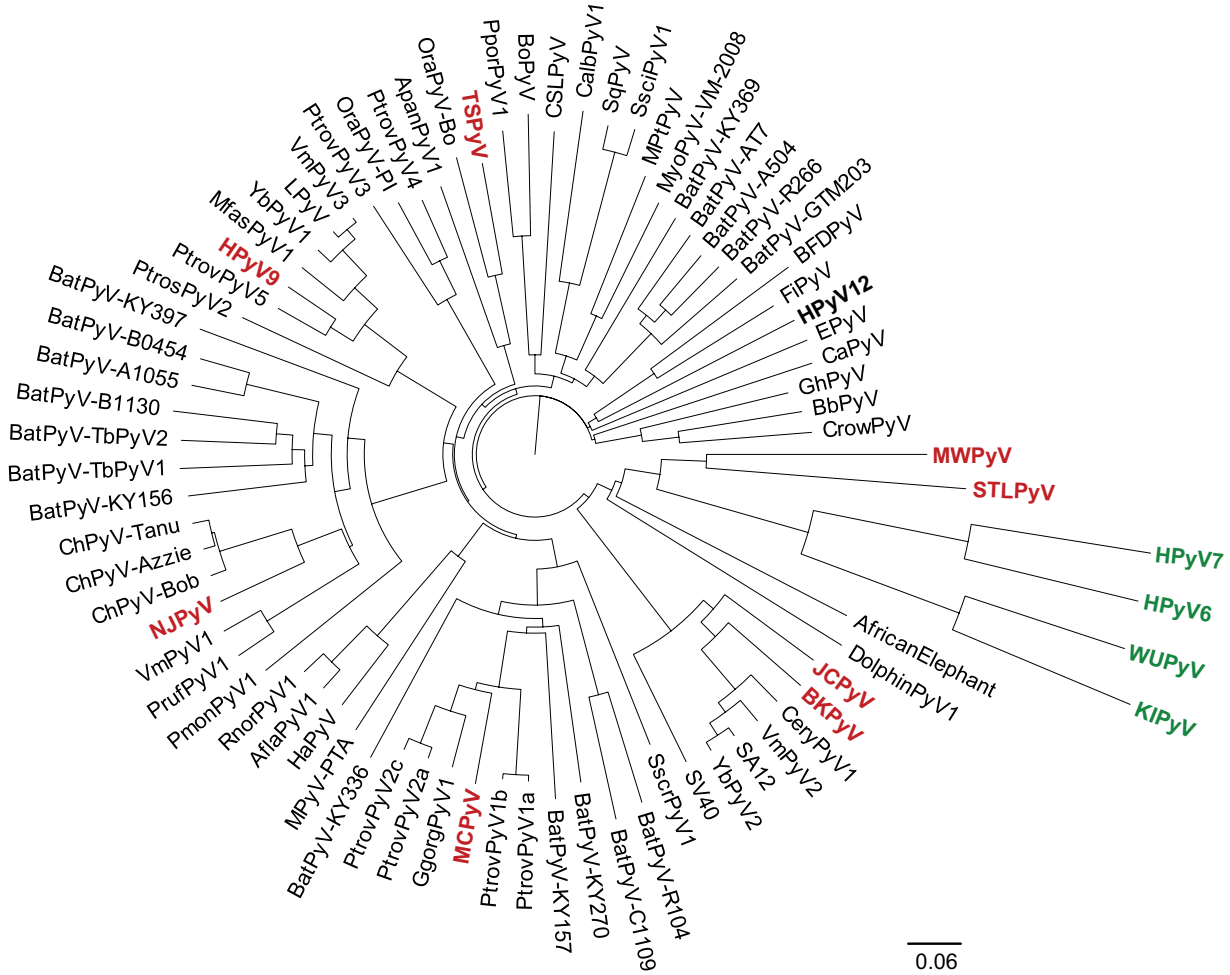


Figure 1: The Polyomavirus family. A neighboring-joining phylogenetic tree based on the amino acid sequence of the major viral capsid protein VP1 was calculated by Clustal Omega (Sievers et al., 2011) and visualized in a polar tree layout using Tree Figure Drawing ToolVersion 1.4.2, Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh. Full virus names and NCBI Genbank accession codes of sequences used for the analysis are given in the Appendix. Human polyomaviruses are highlighted in bold with Orthopolyomaviruses depicted in red and Wukipolyomaviruses in green. The origin of the Human Polyomavirus 12 (HPyV12) has not been confirmed yet.

The recent burst in discovery of new members has reinvigorated the interest in polyomaviruses and their potential contributions to human disease and cancer (Dalianis and Hirsch, 2013; De Gascun and Carr, 2013; Moens et al., 2014). Moreover, the coexistence of at least 13 different polyomavirus species in humans or even in the same organ raises questions about potential immunological interactions and feeds also speculations about synergetic oncogenic transformation mechanisms due to co-infections with polyoma- or even other oncoviruses (Rinaldo and Hirsch, 2013; Moens et al., 2014).

The work presented in this thesis focuses on JCPyV and the more recently discovered polyomaviruses, *Trichodysplasia spinulosa*-associated Polyomavirus (TSPyV), Human

Polyomaviruses 6 and 7 (HPyV6 and HPyV7, respectively), Human Polyomavirus 12 (HPyV12), and NJPyV. However, our knowledge gained from studies with the two model viruses SV40 and MPyV, BKPyV, and other human and non-human polyomaviruses over the last decades is critical for studies focusing on the newly discovered polyomaviruses. Consistently, general characteristics of polyomaviruses highlighted in the following chapters of the introduction are explained and supported by findings from other members of this virus family.

Table 1: Human Polyomaviruses. Source of isolation, associated diseases and seroprevalence; adapted from DeCaprio and Garcea, 2013. Viruses highlighted in grey were subjects of this thesis.

Year of Discovery	Full Name	Abbreviation	Reference	NCBI Reference	Source of isolation	Associated disease	Seroprevalence [%]
1971	BK Polyomavirus	BKPyV	Gardner et al., 1971	NC_001538.1	Urine	Polyomavirus-associated nephropathy (PVAN), hemorrhagic cystitis	82-92 (Knowles et al., 2003; Carter et al., 2009; Kean et al., 2009)
1971	JC Poylomavirus	JCPyV	Padgett et al., 1971	NC_001699.1	Urine, brain	Progressive multifocal leukoencephalopathy	35-80 (Knowles et al., 2003; Kean et al., 2009)
2007	Karolinska Institute Polyomavirus	KIPyV	Allander et al., 2007	NC_009238.1	Nasopharyngeal tissue	-	55-90 (Carter et al., 2009; Kean et al., 2009)
2007	Washington University Polyomavirus	WUPyV	Gaynor et al., 2007	NC_009539.1	Nasopharyngeal tissue	-	70-90 (Carter et al., 2009; Kean et al., 2009)
2008	Merkel Cell Polyomavirus	MCPyV	Feng et al., 2008	NC_010277.1	Lesion	Merkel cell carcinoma	25-87 (Carter et al., 2009; Kean et al., 2009; Nicol et al., 2013)
2010	Human Polyomavirus 6	HPyV6	Schowalter et al., 2010	NC_014406.1	Skin	-	70-85 (Schowalter et al., 2010; Nicol et al., 2013)
2010	Human Polyomavirus 7	HPyV7	Schowalter et al., 2010	NC_014407.1	Skin	HPyV7-associated pruritic rash and thymic epithelial tumors	35-60 (Schowalter et al., 2010; Nicol et al., 2013)
2010	*Trichodysplasia spinulosa*-associated Polyomavirus	TSPyV	van der Meijden et al., 2010	NC_014361.1	Lesion, skin	Trichodysplasia spinulosa	70-80 (van der Meijden et al., 2011; Nicol et al., 2013)
2011	Human Polyomavirus 9	HPyV9	Sauvage et al., 2011; Scuda et al., 2011	NC_015150.1	Skin, blood, urine	-	24-39 (Nicol et al., 2012; Nicol et al., 2013)
2012	Malawi Polyomavirus	MWPyV, MXPyV, HPyV10	Buck et al., 2012; Siebrasse et al., 2012; Yu et al., 2012	NC_018102.1	Stool, wart	-	42 (Nicol et al., 2014)
2012	Saint Louis Polyomavirus	STLPyV	Lim et al., 2013	NC_020106.1	Stool	-	68-70 (Lim et al., 2013)
2013	Human Polyomavirus 12	HPyV12	Korup et al., 2013	NC_020890.1	Liver	-	17-23 (Korup et al., 2013)
2014	New Jersey Polyomavirus	NJPyV	Mishra et al., 2014	NC_024118.1	Muscle, skin	-	-

The Polyomavirus Genome

The organization of the circular dsDNA genome is highly conserved among polyomaviruses and is divided into three genetic regions: (a) the early coding region, (b) the late coding region, and (c) the non-coding control region (NCCR). In Figure 2 the genome of JCPyV is shown as an example. With about 5.0 to 5.5 kilo base pairs in length, the genome of polyomaviruses is forming a chromatin-like structure called minichromosome (Griffith, 1975) and encodes at least five viral proteins - the two regulatory proteins large T antigen (LT) and small T antigen (sT) in the early region and three viral capsid proteins VP1, VP2, and VP3 in the late region (DeCaprio and Garcea, 2013). Additionally, alternatively spliced variants of T antigens have been characterized in the early genome region of SV40 (Zerrahn et al., 1993), MPyV, JCPyV (Trowbridge and Frisque, 1995; Frisque et al., 2003), BKPyV (Abend et al., 2009), MCPyV (Shuda et al., 2008), and Saint Louis Polyomavirus (STLPyV) (Lim et al., 2013). For MPyV, middle T antigen (MT) is the primary transforming protein and cooperates with sT (Asselin et al., 1984), but the role of a tiny T (tT) remains unknown (Riley et al., 1997). Three splicing derivatives of JCPyV T antigen, T'135, T'136 and T'165 seem to fine-tune the transformation potential of the virus (Bollag et al., 2006).

More recently, the overprinting gene ALTO (Alternate frame of the large T antigen) was identified for MCPyV (Carter et al., 2013). This novel protein with a potential role in cellular signaling processes is 248 amino acids long and is expressed during, but is not required for, MCPyV genome replication. The phylogenetic analysis reveals a single monophylogenic clade within the *Polyomaviridae* family named Almipolyomaviruses (ALTO or middle T containing polyomaviruses), which includes also TSPyV (Carter et al., 2013).

Some members of the family feature additional proteins in the late translated region. SV40 expresses VP4 about 24 hours after VP1, VP2, and VP3. VP4 acts as a viroporin and is not incorporated into the virion (Daniels et al., 2007; Raghava et al., 2011, 2013). The expression of agnoproteins has been detected for JCPyV, BKPyV, and SV40 (Jay et al., 1981; Rinaldo et al., 1998; Okada et al., 2002). These small proteins appear to be multifunctional with impact on various aspects of the viral life cycle from regulation of viral transcription and replication to inhibition of host DNA repair and to viroporin function (Suzuki et al., 2010; Unterstab et al., 2010; Sariyer et al., 2011; Saribas et al., 2012; Coric et al., 2014). Putative agnoproteins have been found in a few primate polyomavirus genomes but not in more recently described human polyomaviruses (Ehlers and Moens, 2014).

The NCCR encloses the origin of replication (ORI), TATA- and TATA-like sequences, binding sites for LT, and cellular transcription factors as well as other promoter and transcription enhancer elements. Replication proceeds in a bidirectional, temporally defined manner from the ORI. Genetic rearrangements can occur in the NCCR and seem to be important for JCPyV tropism and pathogenicity (Reid et al., 2011; Delbue et al., 2012; Ferenczy et al., 2012).

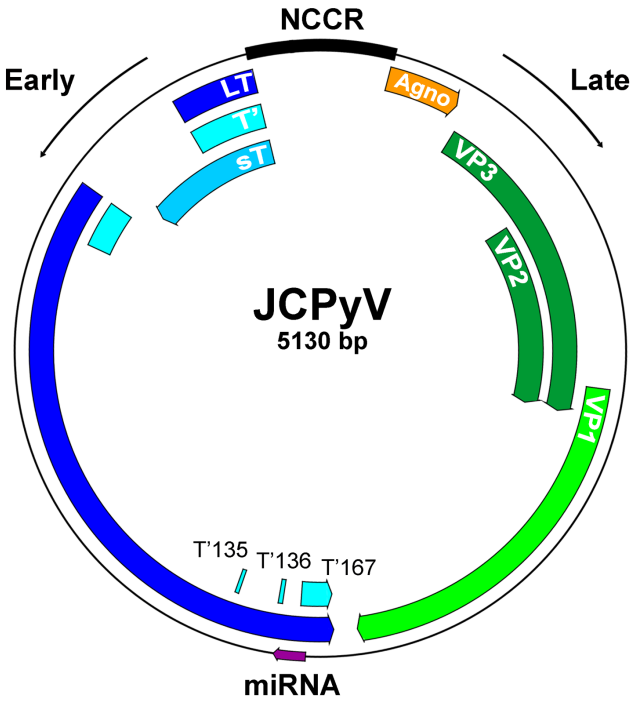


Figure 2: Organization of the circular dsDNA genome of JCPyV. The non-coding control region (NCCR) contains the early and late promoters, their transcription start sites and the origin of replication. The early region encodes LT and sT and alternatively spliced T' forms. The viral capsid proteins VP1, VP2, and VP3 are encoded in the late region. Agnoproteins have been confirmed for BKPyV and JCPyV, and a VP4 has been identified for SV40 but not for the recently discovered human polyomaviruses. Three alternative splicing forms of JCPyV T antigen, T'135, T'136 and T'165 are known. The position of a small non-coding regulatory microRNA (miRNA) is indicated.

Small non-coding regulatory microRNAs (miRNAs) have been reported for a subset of human and non-human members of the polyomavirus family, each of which expresses a single miRNA in the late region targeting early region mRNA in varying locations (Sullivan et al., 2005; Chen et al., 2008; Seo et al., 2008; Sullivan et al., 2009; Chen et al., 2014). The major role of the miRNA is likely the regulation of early gene expression in order to control viral replication in the host (Broekema and Imperiale, 2013). However, some evidence suggests that polyomavirus miRNAs may have also cellular targets, influence cellular miRNA expression, or help to evade the immune system (Bauman et al., 2011; Lee et al., 2011; Xie et al., 2014).

Epidemiology and Pathogenicity

JC Polyomavirus

JCPyV infects approximately 50 to 80% of the human population (Knowles et al., 2003; Kean et al., 2009). The mode of viral transmission is not yet definitely defined, but initial infections seems to be linked to stromal cells of the tonsils, followed by a lifelong persistent infection in the kidney, bone marrow and B-lymphocytes (Monaco et al., 1996; Monaco et al., 1998). JCPyV might enter the upper respiratory tract by close interpersonal contact and presumably spread by the hematogenous route (Ferenczy et al., 2012). Alternatively, transmission by semen and organ transplantation, or urine-to-oral and transplancental routes have been suggested (Bofill-Mas et al., 2001; Bellizzi et al., 2013). JCPyV sheds in the urine of healthy individuals (Yogo et al., 1990) and the detection of JCPyV in stool samples and in sewage and rivers worldwide raises also the possibility of transmission through ingestion of non-sterile water and by a fecal-oral route (Hamza et al., 2009; McQuaig et al., 2009; Ahmed et al., 2010). The infection is mostly asymptomatic in healthy individuals, but in immunosuppressed hosts JCPyV can spread to the central nervous system (CNS), where the virus infects astrocytes, oligodendroctyes, and occasionally neurons (Du Pasquier et al., 2003; Koralnik et al., 2005; Wuthrich et al., 2009b). B-lymphocytes have been proposed to be a major carrier for the virus, enabling JCPyV to disseminate from and/or towards the brain via the blood-brain barrier (BBB) (Tornatore et al., 1992; Chapagain and Nerurkar, 2010; Van Loy et al., 2015). In the CNS, glial cells and in particular oligodendrocytes are critical to the process of myelination. JCPyV infection of oligodendrocytes leads to their cytolytic destruction and causes the brain disease Progressive Multifocal Leukoencephalopathy (PML) (more recently reviewed in Tan and Koralnik, 2010; Bellizzi et al., 2013; Diotti et al., 2013). Symptoms for PML are muscle weakness, sensory deficit, cognitive dysfunction, aphasia, and loss of coordination. The diagnosis for PML includes detection of viral DNA and proteins in the cerebrospinal fluid (CSF) or by brain biopsy. Characteristic lesions in the PML-affected brain are usually visible by Magnetic Resonance Imaging (MRI) in the cortical white matter, cerebellum, peduncles, or the brain stem (Wattjes et al., 2013). HIV-positive patients represented 80% of PML cases corresponding to approximately 3-5% of HIV-1-positive individuals (Major, 2010). Therefore, PML was considered mostly as an AIDS-defining illness until more recently, when PML became a growing concern in other types of patients. These new PML cases are associated with immunomodulatory therapies using monoclonal antibodies for diseases such as MS, Crohn's disease, non-Hodgkin's lymphoma, systemic lupus erythematosus, rheumatoid arthritis, and psoriasis (Diotti et al., 2013). In particular, individuals with MS who are receiving a therapy with Tysabri™ (natalizumab) have a 1:500 chance of developing PML with the

incidence increasing to 1:100 in individuals that were seropositive for JCPyV prior to the immunosuppressive therapy and after treatment with natalizumab for 25 to 28 months (Bloomgren et al., 2012). MS is an autoimmune disease that is characterized by inflammations in the brain and spinal cord and worsens during relapses. Natalizumab is an humanized anti-VLA-4 ($\alpha4\beta1$ integrin) antibody that blocks migration of VLA-4^+ T- and B-lymphocytes across the BBB to the brain (Kawamoto et al., 2012). The treatment protects the brain of an MS patient from attacks by lymphocytes that bind to endothelial cells, but the lack of immune surveillance can result in increased spread of JCPyV to the brain (Diotti et al., 2013). As of December 2014, there had been 517 cases of natalizumab-induced PML and 119 deaths since Biogen Idec introduced the drug in 2004 (Biogen Idec, 2014). Moreover, also newly marketed therapies for MS and for the immune-mediated skin disease psoriasis have led to PML, including Novartis' drug Gilenya™ (fingolimod) and dimethyl fumarate-containing drugs Fumaderm™ and Tecfidera™ (Ermis et al., 2013,van Oosten, 2013 #209; Sweetser et al., 2013). However, a direct correlation between Gilenya™ and PML is disputed so far according to the company information (Novartis statement: Gilenya (fingolimod) safety information update; available online at www.novartis.com).

PML causes a rapid destruction of the white and grey matter in the CNS and usually results in death within one year. There is no specific antiviral therapy against JCPyV available indicating the immediate need to better understand the pathogenesis of the virus (Brew et al., 2010; Diotti et al., 2013). In some cases, restoration of the immune system can support viral clearance and resolution of PML, but the return of function leads frequently to an immune reconstitution inflammatory syndrome (IRIS) (Gray et al., 2005). Until an effective PML therapy is identified, an early diagnosis via clinical surveillance and restoration of the immune system, while monitoring for and managing IRIS, is currently the best attempt to maximize patient chance of survival (PML consortium, available online: pmlconsortium.org; Brew et al., 2010).

Genetic rearrangements of JCPyV, including duplication of enhancer elements in the NCCR that convert the virus to its neuropathogenic form, seem to be necessary for growth of JCPyV in the CNS (Ferenczy et al., 2012). The archetype strain (Cy) resides in the kidney and is usually detected in the urine of healthy individuals (Yogo et al., 1990). PML-type viruses are found in the CSF, brain tissue, and blood but not in the urine (Martin et al., 1985; Jensen and Major, 2001; Pietropaolo et al., 2003; Van Loy et al., 2015). The laboratory PML-prototype strain is the Mad-1 strain, which was originally isolated from the brain of a PML patient and contains a canonical 98-base pair tandem repeat in the NCCR (Frisque, 1983; Frisque et al., 1984). However, the high seropositivity rates for JCPyV within the human population and the low rate of incidences for PML suggest that viral, cellular, and individual host factors play important roles in PML pathogenesis.

In 2005 and 2009 two other distinct CNS diseases were linked to the pathogenicity of JCPyV, JCPyV granule cell neuronopathy (GCN) and JCPyV encephalopathy (JCVE) (Koralnik et al., 2005; Wuthrich et al., 2009b). Both clinical conditions have been observed in PML patients and independently (Wuthrich et al., 2009a; Schippling et al., 2013). In infected granule cell neurons in the cerebellum of GNC patients JCPyV variants with deletions in the C-terminus of the VP1 gene, resulting in frame shifts and alterations of up to 13 amino acids, have been found (Dang and Koralnik, 2006; Agnihotri et al., 2014). Nevertheless, these GCN1 variants have also been detected in CSF and peripheral blood mononuclear cell (PBMC) samples from PML patients. These finding indicate that these deletions may originate outside of the CNS suggesting that these virus variants frequently co-exist as a minor species in the blood of PML patients (Dang and Koralnik, 2006; Dang et al., 2012a). JCVE is an infection of the cortical pyramidal neurons and has been associated with JCPyV variants with deletions in the agnoprotein gene leading to a truncated protein (Dang et al., 2012b). So far mechanisms underlying the neuronal tropism of the JCVE- and GCN-associated JCPyV variants remain unresolved.

Skin-tropic Polyomaviruses: Merkel Cell Polyomavirus, *Trichodysplasia spinulosa*-associated Polyomavirus and Human Polyomaviruses 6 and 7

Our skin is the largest organ of our body, is an important physical barrier against external, potential aggressive agents, and it harbors a diverse microbiota comprising symbiotic bacteria, fungi and viruses. Four members of the polyomavirus family, MCPyV, TSPyV, HPyV6, and HPyV7, have been found to be associated with the human skin, whereas at least MCPyV, TSPyV, and HPyV7 are following the “opportunistic” strategy of polyomaviruses (Feng et al., 2008; Schowalter et al., 2010; van der Meijden et al., 2010; Ho et al., 2014). Infections with these viruses are usually latent and persistent, but a symptomatic and severe reactivation can occur during immunosuppression of the host.

The MCPyV genome was found to be integrated into the host genome of MCCs, which are rare but aggressive and highly lethal, neuroendocrine tumors of the skin in elderly and immunocompromised individuals (Feng et al., 2008; Houben et al., 2010; Hughes et al., 2014). In the initial study, viral DNA sequences were identified in eight of ten and integrated in five of eight MCCs, linking MCPyV to the pathogenesis of MCC (Feng et al., 2008). Others studies have also confirmed the association between MCC and MCPyV (Kassem et al., 2008; Shuda et al., 2008), but no preferential integration sites in the genome have been identified so far (Feng et al., 2008; Sastre-Garau et al., 2009; Laude et al., 2010). The disruption of the viral genome as a result of integration leads to viral sequences with mutations and deletions in the NCCR and in genes coding for LT and VP1, suggesting a strong selection pressure to eliminate viral replication within MCCs (Feng et al., 2008; Sastre-Garau et al., 2009; Laude et al., 2010).

TSPyV was discovered in 2010 in facial skin samples from a patient with *Trichodysplasia spinulosa* (TS), a skin disease that is characterized by the development of follicular papules and keratin spines (van der Meijden et al., 2010). Electron microscopic studies have confirmed the present of icosahedral viral particles in affected hair follicles, which predominantly occur in the face and less frequently on the trunk and limbs of the patients (Osswald et al., 2007; Matthews et al., 2011). Seroprevalence values of about 70% within the adult human population indicate that TSPyV infections occur during childhood and suggest the scenario of a persistent asymptomatic infection and reactivation in immunocompromised patients (van der Meijden et al., 2011; Nicol et al., 2013; van der Meijden et al., 2013). However, high viral loads were only detected in TS patients (van der Meijden et al., 2010; Kazem et al., 2012), whereas samples from the skin, plucked eyebrows, serum/plasma, and urine of immunocompetent and -compromised individuals have been tested mostly negative for TSPyV DNA (Scuda et al., 2011; van der Meijden et al., 2011; Kazem et al., 2012). Thus, a persistent infection at undetectable levels, persistency in an so far undiscovered latent extracutaneous reservoir, or primary infections during an immunosuppressive therapy or after organ transplantation are possible (Kazem et al., 2013). The detection of TSPyV DNA in tonsillar samples from healthy individuals indicates that the virus may establish a persistent infection lymphoid tissue (Sadeghi et al., 2014). Viral shedding and spreading from this site may then be crucial for transmission and reactivation during immunosuppression. The pathogenic mechanism of TSPyV includes uncontrolled hyperproliferation of inner root sheath (IRS) cells but remains otherwise largely unknown (Kazem et al., 2012; Kazem et al., 2014).

HPyV6 and HPyV7 are commonly shed from the skin of healthy individuals together with MCPyV (Schowalter et al., 2010). A persistent infection with both viruses seems very common, resulting in seropositivity rates of 35 to 90% by adulthood (Schowalter et al., 2010; Nicol et al., 2013). Pruritic, brown plaques on the trunk and extremities of transplant patients characterized by distinctive epidermal hyperplasia with HPyV7 capsid-laden keratinocytes indicates that HPyV7 is pathogenic in immunosuppressed individuals (Ho et al., 2014). More recently, the association of HPyV7 with thymic epithelial tumors has been demonstrated (Rennspiess et al., 2015) and thus, also for HPyV6 an involvement in cutaneous diseases or tumors has to be considered (Duncavage and Pfeifer, 2011; Schrama et al., 2012; Scola et al., 2012). Due to their close sequence relationship with the two respiratory Washington University and Karolinska Institute Polyomaviruses (WUPyV and KIPyV, respectively) HPyV6 and HPyV7 have been classified as Wukipolyomaviruses (Johne et al., 2011).

Human Polyomavirus 12

HPyV12 has been described in a study focusing on the identification of new human polyomaviruses in the gastrointestinal tract (esophagus, stomach, colon, rectum, liver, gall

bladder), spleen, and lymph nodes (Korup et al., 2013). In this study, HPyV12 was detected in liver tissue by PCR, but its human origin is not yet confirmed (personal communication B. Ehlers, Division of Viral Infections, Robert Koch Institute, Germany). However, initial seroprevalence studies performed with a VP1 pentamer-based enzyme-linked immunosorbent assay (ELISA) reveal a prevalence of 15 to 33% in healthy adults (Korup et al., 2013). A cross-reactivity of antibodies against other known human polyomaviruses is rather unlikely because none of their VP1 possesses more than 60% amino acid sequence identity to HPyV12 VP1. So far, cross-reactivity has been only a problem with VP1 sequence identities above 75% (Kean et al., 2009). Nevertheless, cross-reactivity against so far unknown human polyomaviruses cannot be ruled out. In the phylogenetic analysis based on LT, HPyV12 was described as an earliest offshoot of a large and diverse clade comprising polyomaviruses from apes, bats, monkeys, rodent, and humans (MCPyV and TSPyV) (Korup et al., 2013), but the analysis based on the VP1 amino acid sequence clusters HPyV12 with Avipolyomaviruses (Figure 1).

New Jersey Polyomavirus

NJPyV was identified via a muscle biopsy in an organ transplant recipient with systemic vasculitis, myositis, and retinal blindness (Mishra et al., 2014). NJPyV is closely related to ChPyV and prior to its discovery its existence had been predicted in a serological study (Scuda et al., 2013). Origin and duration of the infection with NJPyV remain unknown, but the patient did not recall exposure to non-human primates (Mishra et al., 2014). Histopathology analysis suggests a tropism for endothelial cells and thus, NJPyV may have contributed to muscle and ocular damage in the patient. Further studies are needed to clarify factors and incidences for human infections.

Structure and Organization of the Polyomavirus Capsid

The polyomavirus capsid with a diameter of approximately 45 to 50 nm comprises 72 pentamers of the major capsid protein VP1, which are centered on the vertices of a T=7*dextro* (*d*) icosahedral lattice (Liddington et al., 1991; Stehle et al., 1994; Yan et al., 1996). Each VP1 pentamer associates usually with a single molecule of either minor capsid protein VP2 or VP3, which is located inside the capsid (Barouch and Harrison, 1994). The tertiary structure of VP1 can be divided into an N-terminal arm, a jelly-roll β-sheet sandwich core structure, and a long C-terminal extension (Liddington et al., 1991). X-ray structures of polyomavirus VP1 pentamers feature often extended and structurally variable surface loops that emanate from the conserved jelly-roll core formed by β-strands B, I, D, G2 and C, H, E, F (Figure 3). These surface loops are named BC-, DE-, HI-, and EF-loops after the strands connected by them (Liddington et al., 1991; Stehle and Harrison, 1997; Neu et al., 2008; Neu et al., 2010; Neu et al., 2011; Neu et al., 2012; Neu et al., 2013a; Neu et al., 2013b; Khan et al., 2014). The BC-loop is subdivided into two parts, which emanate from the top of VP1 into different directions (BC1- and BC2-loops).

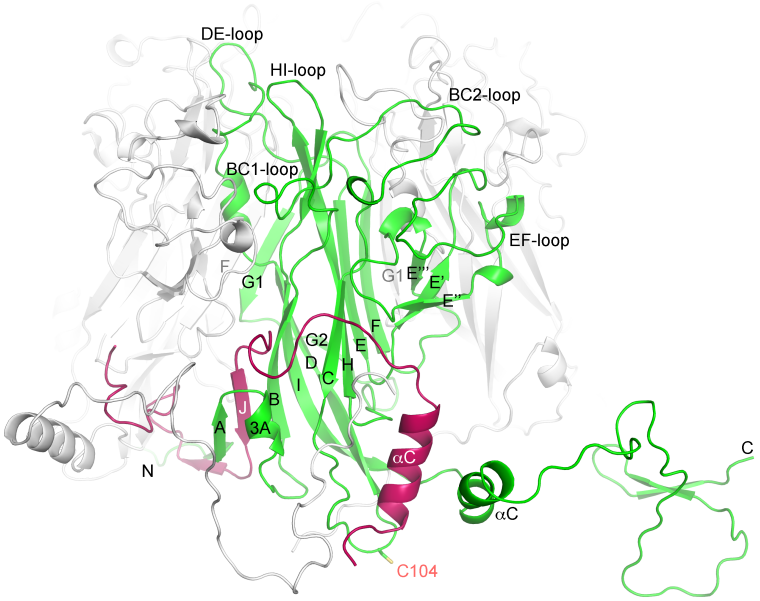


Figure 3: Structure of SV40 VP1. One VP1 monomer of a VP1 pentamer is colored in green and neighboring VP1 monomers in grey (pdb: 1SVA). The conserved jelly-roll fold is built up by β-strands B, I, D, G2 and C, H, E, F. Surface loops are named BC-, DE-, HI-, and EF-loops. The extensive BC-loop is subdivided into BC1- and BC2-loops, which are facing into different directions. The incoming C-terminal arm of a neighboring pentamer in the context of the virion is shown in red. A cysteine residue in the CD-loop (C104 in SV40) is highly conserved among polyomaviruses.

Likewise, the G-strand is divided into the G1- and G2-strands, so that the N-terminal part (G1) of one monomer contributes to the CHEF-sheet of the clockwise neighboring monomer (C, H, E, F, G1) and the G2 to the BIDG2-sheet. In contrast to the other surface loops, the EF-loop

emanates from the bottom VP1 and forms a small three-stranded antiparallel β-sheet (E''', E', E'') at the side of the pentamer (Stehle et al., 1996). Each VP1 monomer forms extensive contacts with neighboring monomers in the pentamer, contributing to the large interface between two VP1 monomers that compromises about 2,200 to 2,800 $Å^2$ (Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute; http://www.ebi.ac.uk/pdbe/prot_int/pistart.html; Krissinel and Henrick, 2007).

The overall structure of the VP1 β-sheet core is essentially identical in the assembled capsid, but significant differences occur in the C-terminal arm (Liddington et al., 1991; Stehle et al., 1996). This arm emerges from the base of each VP1 monomer, and interacts with a VP1 of the adjacent pentamer, thus tying together the pentamers with allowance for the required variability in geometry of the icosahedral capsid geometry (Figure 4).

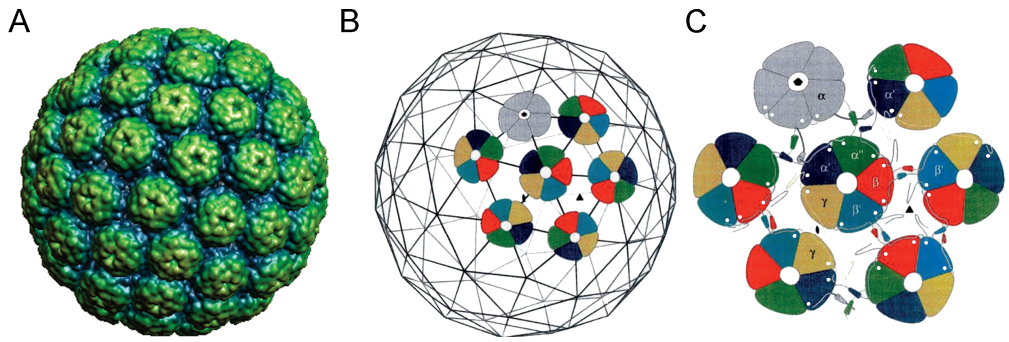


Figure 4: Structure of the polyomavirus virion. (A) The complete SV40 virion (pdb: 1SVA) in a surface representation created with the VIPERdb tool (Carrillo-Tripp et al., 2009). (B) Architecture of the virion shell showing the arrangement of the pentamers on the T=7*d* icosahedral lattice with 5-fold axes through the vertices of the icosahedron, 3-fold axes at the centers of each triangle, and 2-fold axes through the middle edges of each icosahedron. 5-fold and 6-fold coordinated pentamers are depicted in grey and colors, respectively. (C) Interaction between pentamers. C-terminal α-helical segments are shown as cylindrical tubes. Ca^{2+} binding sites are indicated with white spheres. Panels B and C were adapted from Stehle et al., 1996.

In general, icosahedral particles follow three symmetry requirements: 5-fold axes through the vertices of the icosahedron, 3-fold axes at the centers of each triangle, and 2-fold axes through the middle edges of each icosahedron (Figure 4 B). Twelve pentamers lie on 5-fold rotation axes and the remaining 60 pentamers are surrounded by six other pentamers and are not located on any icosahedral symmetry axis. The details of the polyomavirus capsid organization remained a puzzle until structural analysis of SV40 with a sufficient resolution of 3.8 Å could answer the question with a concept of quasi-equivalence at the basis of capsid construction (Liddington et al., 1991). Both, 5- and 6-coordinated pentamers, donate and accept five C-terminal arms, but three different kinds of interpentamer interactions are realized through alternative conformations and arrangements of the long C-terminal arm. A pseudo 3-fold axis (monomers α, α' and α''), a pseudo 2-fold axis (monomers β and β'), and a strict 2-fold

icosahedral axis (γ monomers) are found (Figure 4 B and C). The C-terminus possesses an α-helical segment (αC) with the exceptions of the γ-monomers. Contacts between monomers β and β' involves two helices, between α, α', and α'' monomers three helices, and γ-monomers are mediated by flexible non-helical structures. The C-terminal end can form either a short antiparallel terminal β-ribbon composed of strands named K and L (KL β-ribbon) or is flexible.
However, in all cases, the β-strand A within the N-terminal arm of one pentamer clamps the invading C-arm of a neighboring pentamer at the J-strand resulting in a six-stranded β-sheet (AJBIDG2) (Figure 3). The N-terminus of VP1 binds DNA and contains a nuclear localization signal (NLS) (Moreland et al., 1991; Li et al., 2001).
A cysteine residue in the tip of the CD-loop is conserved among MPyV and SV40 (Stehle et al., 1996; Stehle and Harrison, 1996) as well as other polyomaviruses (Figure 3). In contrast, the CD-loop itself adopts various conformations. In SV40, cysteine residue C104 from the subunit α (VP1 from a 5-fold coordinated pentamer) approaches VP1 monomers α' and α'' from 6-fold coordinated pentamers and can form a disulfide bond with one of the equivalent cysteines of another loop, and also the CD-loops of the β and β' monomers are bridged by disulfide bonding. In addition, cysteine residues C9 in the N-terminal arms may be involved in inter-pentamer disulfide bonds, but these were not visible in the X-ray structure (Stehle et al., 1996; Jao et al., 1999). Pentamers of the MPyV capsid are not cross-linked by disulfide bonds in the X-ray structure, but the equivalent cysteine residue in the CD-loop, C114, forms a disulfide bond with C19 in the N-terminal arm of the counterclockwise VP1 monomer, holding the N-terminal clamp in place (Stehle and Harrison, 1996). Additionally, the assembled SV40 capsid seems to be stabilized by two calcium ions (Ca^{2+}) bridging the incoming C-terminus with core residues of the accepting pentamer (Liddington et al., 1991; Stehle et al., 1996). Thus, Ca^{2+} serves potentially as a trigger at the pentamer interface, and Ca^{2+} addition or depletion can switch on capsid assembly or disassembly, respectively (Salunke et al., 1986; Kawano et al., 2009). The Ca^{2+}-binding sites and cysteine residues are highly conserved among polyomaviruses, suggesting that their capsids are likely stabilized, assembled and disassembled *in vivo* in a similar manner (Stehle et al., 1996; Stehle and Harrison, 1996; Schelhaas et al., 2007; Walczak and Tsai, 2011). Consistently, SV40, MPyV and also other polyomavirus capsids only disassemble *in vitro* by treatment with reducing and Ca^{2+}-chelating agents (Brady et al., 1977; Kosukegawa et al., 1996; Ishizu et al., 2001).
VP1 possesses usually about 350 to 400 amino acids. Only Vervet Monkey Polyomavirus (VMPyV) VP1 and others such as MCPyV, ChPyV, and Bat Polyomavirus (BatPyV) VP1 are about 500 amino acids in length due to an extended C-terminus (Yamaguchi et al., 2014). Deletion of this extended C-terminus seems not to affect the overall morphology of VLPs, but the elongated C-terminus is important for efficient VMPyV VLP assembly (Yamaguchi et al., 2014). In contrast, VP1 of the Avian Polyomavirus (APyV), also known as Budgerigar fledgling

disease Polyomavirus (BFDPyV), has a uniquely truncated C-terminus. Amino acid residues that are forming the KL β-ribbon in SV40 and MPyV are missing (Stehle et al., 1994; Stehle et al., 1996; Shen et al., 2011). The terminal KL β-ribbon may help to lock the capsid in a stable conformation and size variations of the APyV capsid could be explained by a missing β-ribbon (Shen et al., 2011).

The Minor Capsid Proteins VP2 and VP3

VP1 is able to self-assemble into virus-like particles (VLPs) in the absence of the minor capsid proteins VP2 and VP3 with similar structural properties (Kosukegawa et al., 1996; Gillock et al., 1997). In full virions, VP1 pentamers are usually associated with a copy of either VP2 or VP3, although presence and roles of both proteins seem to differ among virus species (Gasparovic et al., 2006; Kawano et al., 2006; Schowalter and Buck, 2013). VP2 and VP3 share a common C-terminus but VP2 features additional amino acids at its N-terminus. VP2 and VP3 comprise about 320-360 and 200-240 amino acid residues, respectively, in MPyV and SV40 as well as JCPyV and BKPyV. MPyV VP1 and VP3 interact with their C-termini with VP1 within the central pore, so that in one VP1 pentamer binds either one molecule of VP2 or VP3 (Barouch and Harrison, 1994; Chen et al., 1998). The X-ray structure of MPyV VP1 in complex with a C-terminal VP2 fragment (residues 214 to 318) shows that mainly an hydrophobic α-helical segment of VP2 is involved in interactions, but also salt bridges and hydrogen bonds at two positions at the inside of the VP1 pentamer are tying both proteins together (Chen et al., 1998) (Figure 5). However, only residues 269 to 296 of VP2 were visible in the electron density map and are included in the complex structure (Chen et al., 1998) (Figure 5 C and D). VP2 and VP3 were found to feature a NLS at their C-termini (Wychowski et al., 1987). The NLS of VP2 and VP3 are likely placed in close distance to the N-terminal NLS of VP1 in the assembled virion (Chen et al., 1998).

During viral infection minor capsid proteins are exposed in the ER. JCPyV mutants lacking VP2 or VP3 are non-infectious suggesting that their exposure is likely a critical factor for viral entry (Gasparovic et al., 2006). Following the exposure step VP2 and VP3 of MPyV and SV40 bind to and/or insert into cellular membranes (Magnuson et al., 2005; Daniels et al., 2006; Geiger et al., 2011). The N-terminus of VP2 is myristylated (Streuli and Griffin, 1987), which is important for infection efficiency likely due its interaction with membranes (Krauzewicz et al., 1990; Sahli et al., 1993).

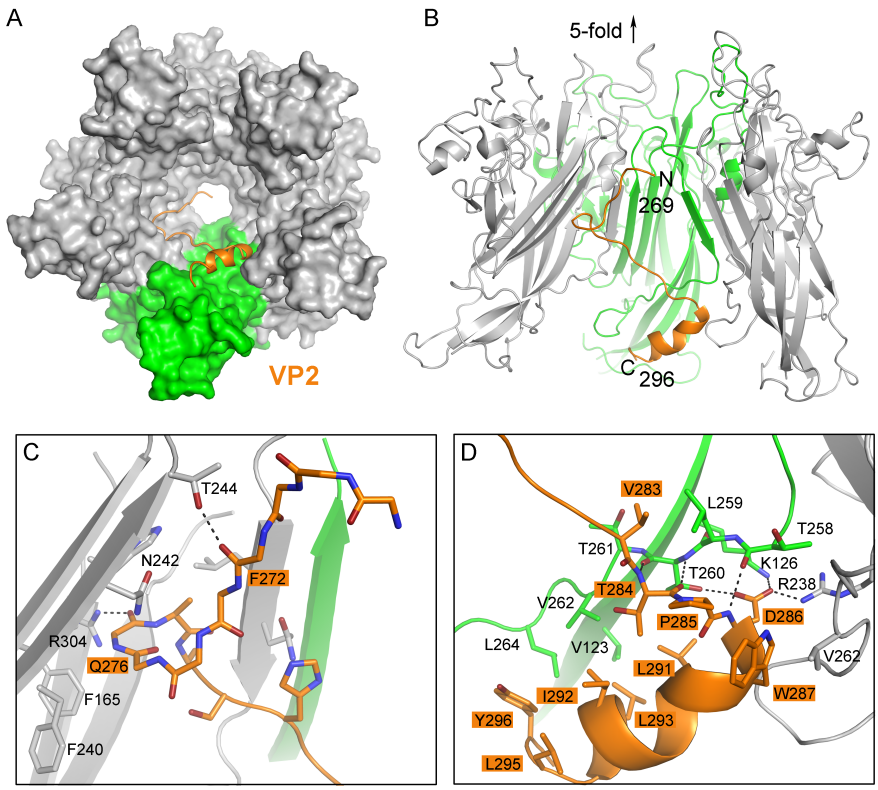


Figure 5: Interaction between VP1 and VP2. (A) - (D) X-ray structure of MPyV VP1 in complex with a C-terminal fragment of VP2 (residues 214-318), which is sufficient for complex formation of VP1-VP2 (Chen et al., 1998). Only VP2 residues 269-296 were defined by electron density and are included in the deposited complex structure (pdb: 1CN3). (A) The VP1-VP2 complex seen from the inside of the virion. The VP1 pentamer is shown in surface representation with one VP1 monomer highlighted in green. VP2 (orange) is shown in cartoon representation. (B) Side view of three VP1 monomers in complex with VP2. (C) Close-up view of VP1-VP2 interactions at the N-terminal region of VP2. (D) Interaction between VP1 and VP2 in the C-terminal region of the VP2 fragment. Hydrogen bonding and salt bridges are shown in dashed lines.

Recombinant VP2 and VP3 produced in *Escherichia coli* (*E. coli*) require a N-terminal glutathione-S-transferase (GST)-fusion protein for solubility (Giorda et al., 2013), co-expression of VP1 (Chen et al., 1998), or are found in inclusion bodies, but refolding of MPyV VP2 and VP3 *in vitro* for biophysical and functional characterizations is possible (Burkert et al., 2014). Both proteins are monomeric and are highly α-helical in solution, but further conformational changes upon interaction with membranes may be required to allowed the insertion of the helices into membranes (Burkert et al., 2014). Studies report membrane insertion via three potential transmembrane domains for SV40 VP2 and VP3 and viroporin activity (Giorda et al., 2013). Such transmembrane helices may not be fully exposed in the soluble form of the protein (Burkert et al., 2014).

Overall, the larger N-terminal part of VP2 seems to be more flexible and not as tightly folded as the common C-terminal region of VP2 and VP3, which engages VP1. Consequently, the N-terminus may emerge more easily from the inside of the virion during infection and

disassembly. However, despite its critical impact on the infection, the mechanism of VP2 exposure in the ER remains unclear. In the VP1-VP2 complex structure, the N-terminal part of the co-crystallized VP2 fragment adopts a hairpin-like conformation and loops back towards the inside of the pentamer (Figure 5). Due the pentameric assembly of VP1, there is a gap along the 5-fold axis, which would be in general sufficient to allow an unfolded polypeptide to emerge either as an elongated loop structure or after structural rearrangement, so that the N-terminus of VP2 emerges first from the inside. Alternatively, VP2 could be exposed through other mechanisms that require capsid expansion or dissociation.

Entry Pathways and Life Cycle

The polyomavirus life cycle starts with the attachment of the capsid to the target cell, proceeds with endocytosis, intracellular trafficking via early, maturating, and late endosomes to the ER lumen for initial uncoating and to the nucleus for genome replication, and finally ends with the assembly and release of progeny virions (Figure 6). Being non-enveloped viruses, polyomaviruses lack a surrounding lipid bilayer and must penetrate biological membranes. This critical step during the infection is fundamentally different from enveloped viruses, which normally gain access to host cells by membrane fusion. Studies with the two model systems SV40 and MPyV but also with JCPyV and BKPyV have revealed common as well as virus-specific cellular mechanisms and interactions during entry and trafficking. Nonetheless, it remains to be investigated in the future how the current knowledge can be transferred to cellular entry mechanisms of the more recently discovered human polyomaviruses.

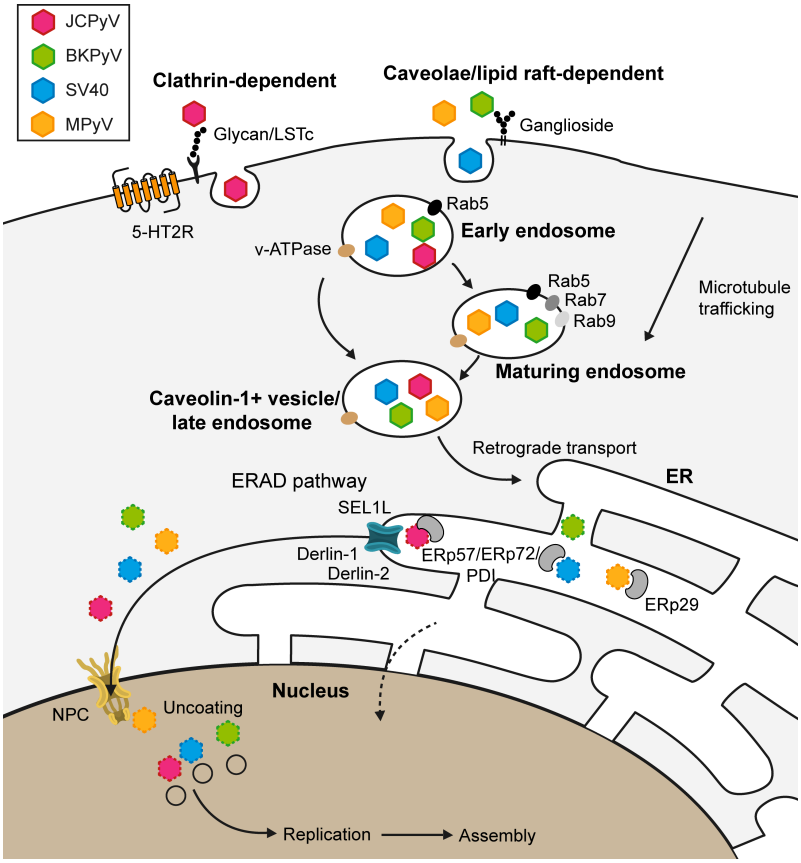


Figure 6: Simplified model of the infectious entry pathways utilized by polyomaviruses. JCPyV enters cells by clathrin-dependent endocytosis. MPyV, SV40, and BKPyV utilize a caveolae/lipid raft-dependent endocytosis pathway. Polyomaviruses traffic to the ER for partial uncoating and then to the nucleus for replication. ERAD, ER-associated degradation; NPC, nuclear pore complex

SV40, Murine Polyomavirus and BK Polyomavirus

SV40, MPyV, and BKPyV bind to glycolipid receptors known as gangliosides (Tsai et al., 2003; Low et al., 2006) and undergo endocytosis in a caveolae- and lipid raft-dependent manner (Gilbert and Benjamin, 2000; Eash et al., 2004; Gilbert and Benjamin, 2004). The multivalent binding of the viral capsid to ganglioside receptors that are located on the extracellular leaflet of the plasma membrane is leading to membrane curvature and viruses initiate hereby their internalization (Ewers et al., 2010). The induced curvature of the membrane seems to be a critical step for viral entry, as even artificial particles coated with GD1a ganglioside antibodies are delivered from the plasma membrane to the ER in a similar manner as MPyV (Qian et al., 2009). Gangliosides have been found in ER fractions, indicating that viruses stay attached to their receptors all the way until reaching the ER (Tsai et al., 2003). Moreover, since the affinity of an individual binding site on the capsid for its receptor is low (Neu et al., 2008) it is assumed that the multivalent binding event of viruses to several ganglioside molecules is necessary for the attachment. How concentration and diffusivity of the ganglioside receptor control the multivalent binding has been investigated for SV40 in more detail. Binding of at least four to five ganglioside molecules in a short time appears to be sufficient for effective viral attachment and initiation of the infection (Szklarczyk et al., 2013).

In the subsequent steps of the infection polyomaviruses rely on microtubules and they are transferred first to maturating endosomes, late endosomes, and probably endolysosomes before they travel to the ER (Gilbert and Benjamin, 2004; Eash and Atwood, 2005; Qian et al., 2009; Engel et al., 2011; Zila et al., 2014). Earlier models suggested that SV40 and MPyV bypass the classical endosomal pathway and accumulate instead in so-called “caveosomes,” which were defined as caveolin-1 (Cav-1)-positive, pH-neutral, endocytic organelles (Pelkmans et al., 2001). Cav-1 belongs to a family of integral membrane proteins that are key players in caveolae (Parton and del Pozo, 2013). However, more recent studies have demonstrated that caveosomes are not distinct cell organelles but instead modified late endosomes or endolysosomes in which Cav-1 is found after overexpression or interference with caveolae assembly (Hayer et al., 2010; Engel et al., 2011). After internalization, polyomaviruses colocalize with several endosomal Rab proteins, which belong to a family of small GTPases and regulate membrane trafficking through effector protein recruitment (Querbes et al., 2006; Qian et al., 2009; Engel et al., 2011). Rab5 is involved in formation, transport, and fusion of vesicles to early endosomes, whereas Rab7 and Rab9 are markers of late endosomes. Studies have detected MPyV and SV40 in Rab5-, as well as Rab7- and Rab9-positive endosomes. In addition, both viruses were also found to be associated with the lysosomal-associated membrane protein 1 (LAMP1), which is a marker of the endolysosomes (Qian et al., 2009; Engel et al., 2011). The vacuolar ATPase (v-ATPase) is responsible for the

acidification of endosomes and lysosomes and seems to be required for pH-dependent conformational changes in the polyomavirus capsid (Qian et al., 2009; Engel et al., 2011). The subsequent retrograde transport from the late endosomes to the ER is a critical and also unique step during the polyomavirus infection since most extracellular ligands do not reach the ER after endocytosis (Querbes et al., 2006; Nelson et al., 2012).

Upon arriving in the ER lumen chaperones and protein disulfide isomerase (PDI) family members such as ERp29, ERp57, ERp72, and the canonical PDI are responsible for the disruption of disulfide bonds to allow conformational changes of the capsid leading to partial exposure of VP2 and VP3 (Magnuson et al., 2005; Gilbert et al., 2006; Schelhaas et al., 2007; Nelson et al., 2012). The precise combination of the utilized PDI proteins differs between polyomviruses possibly due to differences in the disulfide bond network of the capsid (Gilbert et al., 2006; Walczak and Tsai, 2011). It has been postulated that the exposed hydrophobic and myristylated N-terminal part of VP2 helps anchoring the virus to the ER membrane, while the Hsp70 chaperone BiP together with the J-domain of the chaperone DNAJB11 may prevent aggregation of the now partial unfolded and hydrophobic viral particle (Geiger et al., 2011; Goodwin et al., 2011). Studies with SV40 suggests that this partial uncoated virus particle then employs components of the ER-associated degradation (ERAD) machinery for cytosolic entry, although the dependence on the ERAD pathway for ER egress seems to vary between polyomavirus species and furthermore between different cell types (Bennett et al., 2013). ERAD pathway components such as Derlin-1 (Schelhaas et al., 2007), Derlin-2 (Lilley et al., 2006), SEL1 (Schelhaas et al., 2007), RMA1 (Geiger et al., 2011), and BAP29/BAP31 (Geiger et al., 2011) have been proposed to be mediators for the transport across the ER membrane. In the cytosol, ER membrane J proteins may stimulate interactions between the cytosolic chaperon Hsp70 and the virus and could help with viral release into the cytosol (Goodwin et al., 2011; Walczak et al., 2014).

Next, polyomaviruses enter the nucleus likely via the nuclear pore complex (NPC) (Qu et al., 2004). The translocation into the nucleus depends on the recognition of the NLS motif of VP1 and/or of minor capsids proteins by alpha and beta importins (Nakanishi et al., 2002; Qu et al., 2004; Nakanishi et al., 2007). However, polyomaviruses may also be able to enter the nucleoplasm directly from the ER by crossing the inner nuclear membrane by deformation (Butin-Israeli et al., 2011).

Later in the virus replication cycle, capsid proteins are synthesized in the cytoplasm and are transported into the nucleus for assembly at nuclear domain 10 (ND10) bodies, which are also called promyelocytic leukemia (PML) protein nuclear bodies (Ishov and Maul, 1996; Shishido-Hara et al., 2004). Viral DNA replication has been localized to these discrete functional heterogeneous nuclear scaffolds (Shishido-Hara et al., 2004), suggesting that nuclear bodies provide the architecture for “polyomavirus assembly factories” (Erickson et al., 2012). Virus

assembly occurs here in tubular structures with “budding” events at the end, but the PML protein itself seems not to be critical for this process (Shishido-Hara et al., 2004; Erickson et al., 2012). VP1 pentamers are thought to polymerize into tubes around the viral chromatin or the chromatin might later traverse into until a terminal event buds off icosahedral particles at the end. How the VP1 subunits specifically identify and bind the viral genome remains unknown (Oppenheim et al., 1992; Dalyot-Herman et al., 1996; Carbone et al., 2004; Erickson et al., 2012). After assembly, agnoproteins or remaining VP2 and VP3 may support the release of the virus progeny by disruption of the host cell membrane (Suzuki et al., 2010).

Pentameric bacterial toxins such as cholera and shiga toxin share certain characteristics of cellular entry with polyomaviruses (Ewers et al., 2010; Nelson et al., 2013). They attach to gangliosides, use the cholesterol-dependent entry mechanism, and traffic to the ER (Fujinaga et al., 2003; Cho et al., 2012). Although multivalent receptor binding is not essential for the toxicity of cholera toxin, higher receptor binding avidity and/or the induced remodeling of the membranes through lipid clustering increase its effect (Jobling et al., 2012). However, in contrast to ER-targeting bacterial toxins polyomavirus particles are usually not detected in the trans-Golgi network or cisternae of the Golgi complex (Engel et al., 2011; Cho et al., 2012).

JC Polyomavirus

JCPyV attaches to cells via engagement of its glycan receptor motif lactoseries tetrasaccharide c (LSTc) and then enters cells by clathrin-dependent endocytosis through a mechanism involving serotonin 5-hydroxytryptamine-type 2 receptors (5-HT2Rs) and the epidermal growth factor receptor pathway substrate clone 15 (Pho et al., 2000; Elphick et al., 2004; Querbes et al., 2004; Neu et al., 2010; Assetta et al., 2013). 5-HT2Rs are transmembrane-spanning G protein-coupled receptors that are abundantly expressed on multiple cell types, including brain and kidney cells, correlating with the major JCPyV infection sites (Bonhaus et al., 1995). Especially 5-HT2R type 2A (5-HT2AR) is highly expressed in the cerebral cortex (Bonhaus et al., 1995), a site in which JCPyV and demyelinating lesions are usually detected in PML patients (Du Pasquier et al., 2003; Wuthrich and Koralnik, 2012). Overexpression of 5-HT2Rs enhances JCPyV infection in poorly permissive cell lines and 5-HT2AR and 5-HT2CR antibodies block infection (Elphick et al., 2004; Maginnis et al., 2010; Assetta et al., 2013), however, JCPyV still infects human brain microvascular endothelial cells, which have low levels of 5-HT2AR (Chapagain et al., 2007). The drug chlorpromazine, an inhibitor of clathrin-dependent endocytosis and antagonist of 5-HTRs, reduces JCPyV infection in glial cells and the kidney cell line HEK293, which indicates, that viral entry in the brain as well as kidney is mediated by clathrin-dependent uptake (Pho et al., 2000; Assetta et al., 2013). Probing the glycosylation sites on the extracellular N-terminus of 5-HT2AR by

mutagenesis suggests that the sialylated glycan attachment receptor LSTc for JCPyV is not linked to 5-HT2R (Maginnis et al., 2010).

Internalized JCPyV is found together with 5-HT2AR in early endosomes already 5 min post-infection (Elphick et al., 2004). The virus enters then the endocytic system where it colocalizes 15 min post-infection with Rab5 and 2 hours post-infection with Cav-1 (Querbes et al., 2006). Unlike other polyomaviruses, JCPyV does not colocalize with Rab7, which is important for the endosomal maturation process (Querbes et al., 2006). Thus it remains to be clarified whether JCPyV actually enters maturing and late endosomes like other polyomaviruses, and if the association with Cav-1 occurs in distinct independent organelles or in Cav-1-positive late endosomes (Querbes et al., 2006). The fact that JCPyV does not employ a ganglioside receptor for attachment and entry may impact subsequent virus accumulation and trafficking (Qian et al., 2009; Ströh et al., 2015). Nevertheless, intracellular pathways of JCPyV and other polyomaviruses seem to intersect in large parts. For example, JCPyV and SV40 infection is inhibited by the drug Brefeldin A, indicating that productive trafficking of both viruses to the ER is Arf1-GTPase-dependent and shared despite differences during the early steps of the infection (Querbes et al., 2006).

Similarly to other polyomaviruses, the cellular uptake of JCPyV and intracellular transport to the ER is rather slow. JCPyV colocalizes with the ER protein calreticulin 6 to 16 hours post-infection (Nelson et al., 2012). In the ER the canonical PDI and ERp57 are necessary for isomerization of interpentamer disulfide bonds of the JCPyV capsid and the partial exposure of VP2 is a critical step for the following ER egress (Nelson et al., 2012; Nelson et al., 2013).

1.2 Virus Glycan Receptors

Many viruses, including several human pathogens, interact with glycans on the cell surface of target cells during the initial steps of the infection. Viral glycan receptors such as sialic acid-containing carbohydrates or glycosaminoglycans (GAGs) are often negatively charged, but also neutral glycans such as histo-blood group antigens (HBGAs) function as receptors or attachment factors for viruses (Figure 7).

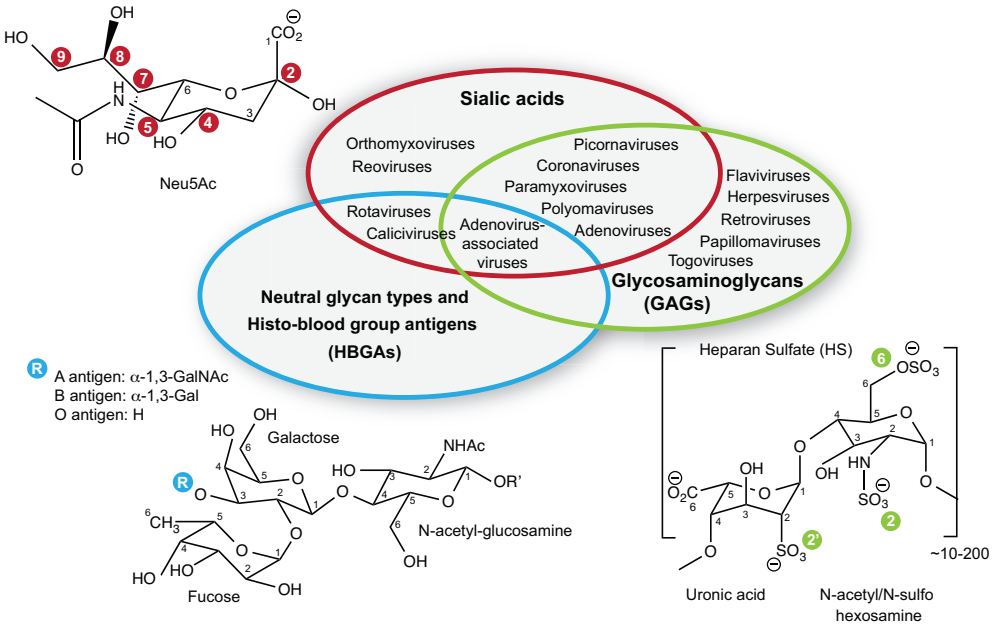
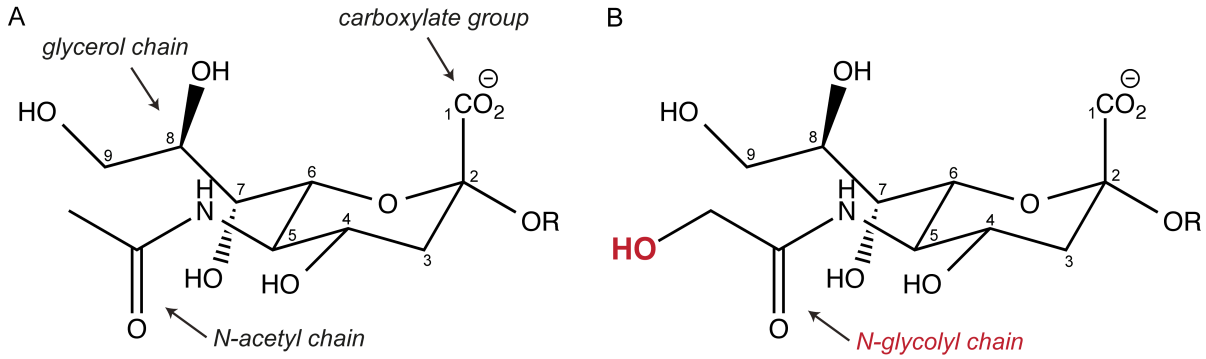


Figure 7: Glycans as viral receptors. Host- and cell-specific variants of sialic acids, neutral oligosaccharides (for example histo-blood group antigens (HBGAs)), and glycosaminoglycans (GAGs) are used as virus receptors or attachment factors. Virus families with members that are known to interact with one or more glycan types are indicated in the respective overlapping ellipses of the Venn diagram. The most common human sialic acid Neu5Ac is shown. Modifications occur at several carbon positions (see also Figure 8). The type 2 core antigen (blue) is part of the ABO blood group determinants. R′ indicates where the core is attached to a glycoprotein. Disaccharide units consisting of N-acetyl- or N-sulfohexosamine and either an uronic acid (glucuronic or iduronic acid) or galactose form long GAG chains. Heparan sulfate (HS) is shown as an example. HS usually vary by N-sulfation or N-acetylation at carbon position C2 of the hexosamine and sulfation at C2 of the uronic acid and at C6 of the hexosamine. Modified from Ströh and Stehle, 2014.

Sialic acids, which are 9-carbon, α-acidic keto sugars, are ubiquitously expressed at cell membranes of higher vertebrates, where they are located at terminal ends of N-glycans, O-glycans and lipids (Varki and Varki, 2007; Varki and Schauer, 2009). In humans, the α-5-*N*-acetyl neuraminic acid (Neu5Ac), with a carboxylate group at C2, a N-acetyl group at the C5 position and a glycerol chain at C6, is the most abundant variant, but various sialic acid substitutions (acetylation, methylation, and sulfatation) at C4, C5, C7, C8 and C9 are known (Angata and Varki, 2002) (Figure 8). α-5-N-glycolyl neuraminic acid (Neu5Gc) is not produced

by humans due to a mutation in the gene coding for the enzyme cytidine monophosphate-*N*-acetyl neuraminic acid hydroxylase (CMAH), which hydroxylates Neu5Ac (Chou et al., 1998; Irie et al., 1998). Thus, the functional group at C5 plays an important role in human evolution, immunology and other sialic acid-mediated biological processes (Varki, 2001; Chou et al., 2002; Varki, 2007; Padler-Karavani et al., 2008). Furthermore, the absence of Neu5Gc affects susceptibility to various pathogens, and many human-specific pathogens target Neu5Ac on cells (Byres et al., 2008; Deng et al., 2014). However, Neu5Gc can be absorbed from diary sources and is found in certain cancers (Tangvoranuntakul et al., 2003; Samraj et al., 2014).



5α-*N*-Acetylneuraminic acid (Neu5Ac) 5α-*N*-Glycolylneuraminic acid (Neu5Gc)

Figure 8: Structures of (A) Neu5Ac and (B) Neu5Gc. Neu5Ac, with an N-acetyl group at the C5 position, is the most abundant sialic acid variant in humans. Neu5Gc is produced in monkeys and other vertebrates but not in humans. Characteristic functional groups of sialic acids are the carboxylate group, glycerol and N-acetyl or N-glycolyl chain, respectively, but also various substitutions (acetylation, methylation, and sulfation) at C4, C5, C7, C8, and C9 of the 9-carbon backbone are known (Angata and Varki, 2002).

In glycoconjugates such as N- or O-linked proteins and glycolipids such as gangliosides, sialic acids are usually attached via α2,3- or α2,6-linkages to either a galactose (Gal) or to an *N*-acetylgalactosamine (GalNAc), or via α2,8- or α2,9-linkages to one another (Yu et al., 2011a; Sato and Kitajima, 2013). Gangliosides are divided into four groups, or “series”, termed asialo-, a-, b- and c-series based on the number and on complexity of their sialic acid branching patterns (Schnaar et al., 2009). They occur ubiquitously in most tissues with varying concentrations of the different forms according to cell types. In the adult human brain, GM1, GD1a, GD1b, and GT1b are very abundant, comprising up to 97% of the present gangliosides, whereas for example the structurally simplest member of gangliosides GM3 is predominant in liver and kidney (Nilsson and Svennerholm, 1982; Svennerholm et al., 1989; Prokazova et al., 2009; Yu et al., 2009; Sturgill et al., 2012). Oriented towards the extracellular matrix and anchored through their ceramide lipids into the outer leaflet of the plasma membrane, gangliosides are involved in various signaling processes, as well as cell-cell and cell-matrix

interactions (Lopez and Schnaar, 2009). However, they are usually not uniformly distributed in the membrane but clustered in lipid rafts (Lingwood and Simons, 2010).

Occurrence and distribution of glycan receptors and pseudo receptors on target cells as well as binding specificity and binding affinity determine tissue and host tropism and ultimately pathogenicity of viruses. Hereby influenza virus A is probably the best-studied and known example to demonstrate the importance of a particular glycosidic linkage (recently reviewed in Imai and Kawaoka, 2013; de Graaf and Fouchier, 2014). Avian strains engage α2,3-linked sialic acids on intestine tissue in birds and usually do not infect humans, while human strains recognize α2,6-linked sialic acids of the human respiratory tract. However, a receptor switch from α2,3- to α2,6-linked sialic acids in an avian or porcine reservoir could eventually lead to transmission to humans causing widespread infection in humans and viral pandemics.

Polyomavirus Sialic Acid Glycan Receptors

Glycans featuring terminal sialic acids function as cell surface attachment receptors for several polyomaviruses (Table 2). SV40 uses the GM1 ganglioside, whereas MPyV attaches to GD1a and GT1b gangliosides and BKPyV engages GD1b and GT1b gangliosides as receptors (Tsai et al., 2003; Low et al., 2006). MCPyV VP1 pentamers bind the GT1b ganglioside *in vitro* (Erickson et al., 2009) and for JCPyV, LPyV, and HPyV9 several sialylated glycan receptor motifs have been identified although location and presentation on cells remains unclear (Komagome et al., 2002; Neu et al., 2010; Gorelik et al., 2011; Neu et al., 2013b; Khan et al., 2014). Crystal structures of polymavirus VP1 in complex with specific sialylated glycan receptors motifs have been solved and reveal that surface loops connecting β-strands mediate the interactions with the terminal sialic acid residues in all cases (Neu et al., 2008; Neu et al., 2010; Neu et al., 2012; Neu et al., 2013a; Neu et al., 2013b; Khan et al., 2014). The protruding functional groups of the sialic acid are commonly employed for specific engagement by VP1 in the same general area of the protein (Figure 9). A small number of additional contacts outside the binding pocket achieve glycan specificity in the case of closely related SV40, BKPyV, and JCPyV, which are engaging the terminal Neu5Ac in a highly conserved manner (Neu et al., 2013a) (Figure 10). Alternatively, sialic acids are recognized in a different and unique orientation by applying specific binding strategies as observed for MCPyV, MPyC, LPyV, and HPyV9 (reviewed in Ströh and Stehle, 2014) (Figure 9).

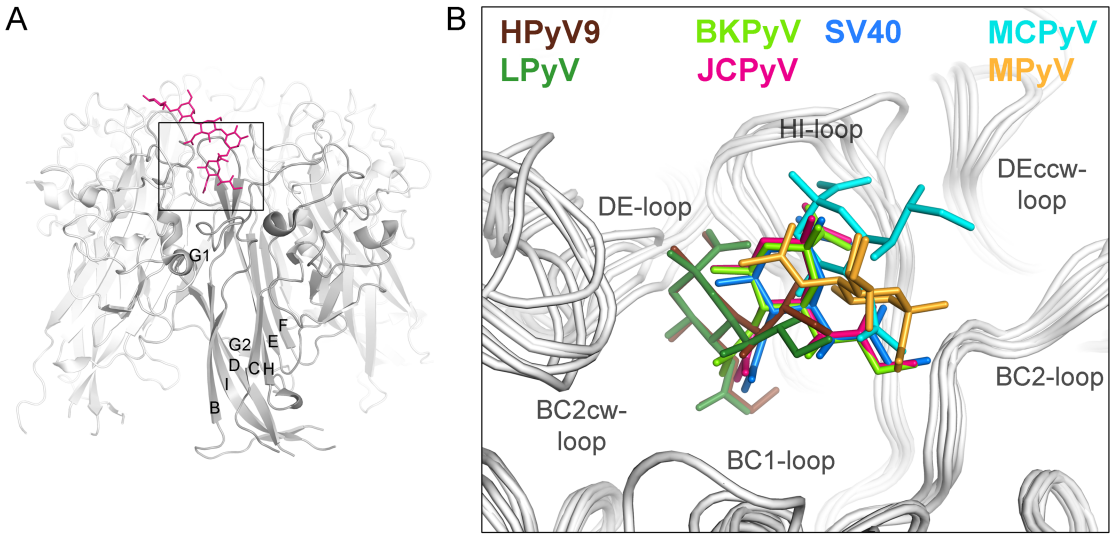


Figure 9: A highly plastic sialic acid binding site on polyomavirus VP1. The structural superposition of polyomavirus VP1-glycan complex structures reveals an overall conserved location of the sialic acid binding site, which is build up by variable surface loops connecting the β-strands of the jelly-roll core. (A) The complex structure of the JCPyV VP1 pentamer with the LSTc glycan (pdb: 3NXD) is shown exemplary to highlight the overall location of the binding site. (B) Close-up view of the binding site. Terminal sialic acid residues are shown in stick representations. SV40 VP1-GM1 glycan (pdb: 3BWR), BKPyV VP1-GD3 glycan (pdb: 4MJ0), JCPyV VP1-LSTc, MCPyV VP1-GD1a glycan (pdb: 4FMJ), LPyV VP1-α2,3 Neu5Ac-sialyllactose (pdb: 4MBY), MPyV VP1-Neu5Ac-α2,3-Gal-β1,3-[α2,6-Neu5Ac]-GlcNAc-β1,3-Gal-β1,4-Glc (pdb: 1VPS), HPyV9 VP1-α2,3 Neu5Gc-sialyllactose (pdb: 4POT). Modified from Ströh and Stehle, 2014.

The pentasaccharide LSTc was identified as receptor motif for JCPyV by glycan array screening (Neu et al., 2010). Subsequent structural studies revealed that JCPyV VP1 from the prototype Mad-1 strain engages LSTc in a unique L-shaped conformation (Figure 10 A). Mutations of VP1 residues contacting LSTc lead to a severe defect in viral growth in glial cells (Neu et al., 2010). The LSTc glycan features a terminal α2,6-linked Neu5Ac, whereas previous studies had suggested that JCPyV can use both α2,3- and α2,6-linked sialic acid receptors for infection (Dugan et al., 2008). Additionally, virus-like particles (VLPs) of a urine isolate were found to have an altered capacity to engage sialylated receptors, suggesting a possible link to differences in viral tropism and spread in JCPyV pathogenesis (Gorelik et al., 2011). The urine isolate belongs to the genotype 3 and is therefore referred to as WT3, whereas the Mad-1 strain is a genotype 1 strain. WT3 VLPs engage asialo-GM1, GD1a, GD1b, GD2, and GT1b (Gorelik et al., 2011) as well as GM1 and GM2 (Leonid Gorelik, Consortium for Functional Glycomics (CFG); available online at www.functionalglycomics.org according to CFG policy). However, Mad-1 VLPs have also been found to bind to several gangliosides, including GM3, GD2, GD3, GD1a, GD1b, GT1b, and GQ1b, but not GM1 or GM2, by a virus-overlay-blotting-assay (Komagome et al., 2002). These findings suggest that JCPyV subtypes may either have

altered binding specificities or affinities for α2,3- and α2,6-linked Neu5Ac glycans, which may have then different roles and function during infection and for JCPyV pathogenicity.

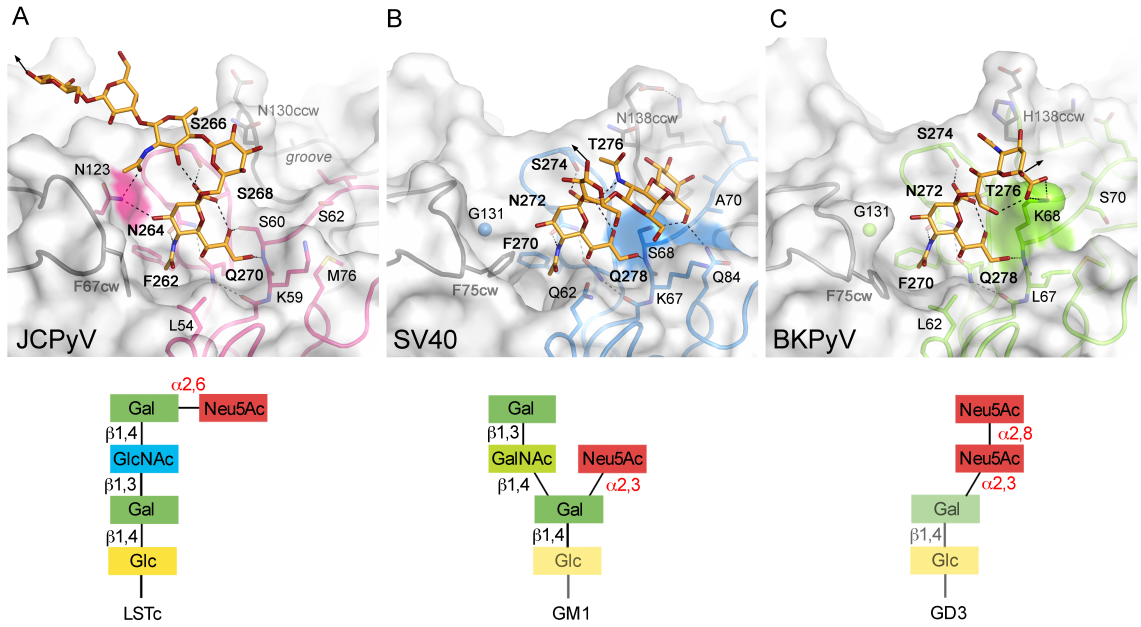


Figure 10: Glycan binding site of JCPyV, SV40, and BKPyV. (A) JCPyV VP1-LSTc (pdb: 3NXD), (B) SV40 VP1-GM1 glycan (pdb: 3BWR), and (C) BKPyV VP1-GD3 glycan (pdb: 4MJ0) complex structures are shown with VP1 in surface and cartoon representations. Glycans and VP1 residues in and adjacent to binding sites are represented with sticks. Hydrogen bonds and salt bridges are indicated with black dashed lines. Residues highlighted in bold are involved in conserved interactions with the terminal Neu5Ac in all three structures. Schematic representations of glycans used in in crystallographic studies are shown below the structural figures. Glycan residues that are contacted by VP1 are highlighted in comparison to residues not involved in protein glycan interactions. Structural figures and glycan representations were adapted from Neu et al., 2013a.

In simians, the natural hosts of the closely related SV40, the GM1 ganglioside contains often a terminal Neu5Gc instead of NeuAc as seen in humans. Indeed, glycan array screening has shown that the GM1 ganglioside featuring a terminal Neu5Gc is a better receptor than the human counterpart (Campanero-Rhodes et al., 2007). The crystal structure of a SV40 VP1-Neu5Ac-GM1 glycan complex provides the explanation (Neu et al., 2008). The N-acetyl group of Neu5Ac points into a large cavity that is built up by polar and hydrophobic residues and is much larger than in VP1 of the related human JCPyV and BKPyV (Figure 10 C). It could accept the hydroxymethyl group of the Neu5Gc and the predicted conformation would even allow additional hydrogen bonds that result in increased affinity for Neu5Gc-GM1 compared to Neu5Ac-GM1 (Stehle and Khan, 2014). A single point mutation in BKPyV VP1 (K68S) switches the glycan receptor binding specificity from α2,8-α2,3-diNeu5Ac containing b-series gangliosides to the SV40 GM1 ganglioside receptor (Neu et al., 2013a). However, BKPyV

K68S retains the binding specificity for the human Neu5Ac-modified GM1 ganglioside due to its shallow cavity on VP1 compared to SV40 (Neu et al., 2013a).

In contrast to the receptor binding specificity of SV40 and BKPyV corresponding to the most prevent sialic acid variant in their host, glycan array screening and structural analysis of HPyV9 VP1 displayed a preference for α2,3-Neu5Gc-linked linear receptors whereas the closely related monkey virus LPyV seems to engage an α2,3-Neu5Ac-linked motif (Neu et al., 2013b; Khan et al., 2014). Interactions with an additional hydroxymethyl group are not favored in the binding pocket of LPyV VP1 that is otherwise largely conserved compared to HPyV9. Thus, the detailed understanding of the recognition of Neu5Ac variants by viruses is only beginning to emerge, but also bacterial toxins are known to engage Neu5Gc-linked glycans on human gut endothelium cells, which present Neu5Gc absorbed from red meat or milk products (Byres et al., 2008).

Functional cell-based and structural studies demonstrate that several b-series gangliosides support BKPyV infection albeit with differences in affinities (Neu et al., 2013a). However, the increased affinity toward more complex b-series gangliosides may be only one determinant for receptors usage. Theoretically, lower affinity could be balanced by greater receptor abundance on cells, so that weaker binding gangliosides could have still implications for BKPyV tropism and pathogenesis (Neu et al., 2013a). BKPyV establishes a persistent infection in the human kidney, and here cells possess in particularly more simple gangliosides such as GM3 and GD3 (Shayman and Radin, 1991). The relative abundance of gangliosides differs not only between brain and kidney, expression of gangliosides varies also between the different tissues of the adult human kidney (Holthofer et al., 1994; Yu et al., 2009). Consistently, developmental or drug-induced changes in b-series ganglioside distribution may play a role in latency and reactivation of BKPyV (Neu et al., 2013a).

The impact of precisely regulated glycan-binding affinities on viral pathogenicity has been demonstrated for MPyV (Freund et al., 1991; Bauer et al., 1995; Bauer et al., 1999). Sialylated glycolipids and -proteins, two major constituents of the plasma membrane, execute opposing functions during MPyV infection. The engagement of “decoy” or pseudo receptors such as glycoproteins restricts the infection of MPyV (Qian and Tsai, 2010). A mutation in the VP1 sialic acid binding pocket reduces the affinity for sialylated pseudo receptors on cells and decreased cell binding results in increasing viral dissemination and pathogenicity.

Table 2: Polyomavirus receptors and receptor candidates. Adapted and modified from O'Hara et al., 2014b. 3'SL, α2,3-sialyllactose; DSLNT, Disialyllacto-N-tetraose; 3'SLN, α2,3-sialyllactosamine.

Virus	Structural studies	Glycan receptors or glycan receptor motifs; identified via (a) glycan array screening, (b) floating assay, (c) ELISA, (d) cell supplementation assay, (e) virus overlay assay	Co-receptors
MPyV	3'SL (Stehle et al., 1994) DSLNT (Stehle and Harrison, 1996)	(b) GD1a, GT1b (Tsai et al., 2003) (d) GD1a (Tsai et al., 2003)	α4β1 integrin (Caruso et al., 2003)
SV40	GM1 (Neu et al., 2008)	(a) GM1 (Neu et al., 2008) (b) GM1 (Tsai et al., 2003) (d) GM1 (Campanero-Rhodes et al., 2007)	MHC class 1 (Atwood and Norkin, 1989) Axl (Drayman et al., 2013) α2β1 integrin (Stergiou et al., 2013)
BKPyV	GD3 (Neu et al., 2013a)	(a) GD1b (Neu et al., 2013a) (b) GT1b, GD1b (Low et al., 2006) (d) GD3, GD2, GD1b, GT1b (Neu et al., 2013a)	-
JCPyV	LSTc (Neu et al., 2010)	(a) LSTc (Neu et al., 2010), GM2, GM1 (Consortium of Functional Glycomics; www.functionalglycomics.org) (c) GD1a, GD2, GD1b, GT1b (Gorelik et al., 2011) (e) GM3, GD2, GD3, GD1a, GD1b, GT1b, GQ1b (Komagome et al., 2002)	5-HT2R (Elphick et al., 2004; Assetta et al., 2013)
MCPyV	Neu5Ac-α2,3-Gal (Neu et al., 2012)	(b) GT1b (Erickson et al., 2009)	Glycosaminoglycans (Schowalter et al., 2011)
LPyV	3'SL, 3'SLN (Neu et al., 2013b)	(a) 3'SL, 3'SLN (Neu et al., 2013b)	-
HPyV9	Neu5Ac-3'SL, Neu5Gc-3'SL (Khan et al., 2014)	(a) Neu5Ac-3'SL, Neu5Gc-3'SL (Khan et al., 2014)	-

Studying Virus-Glycan Interactions

There has been a large increase in studies on virus-glycan interaction over the past years due to recent technological advances such as reverse genetic systems for more and more viruses, glycan binding arrays, X-ray crystallography, and biophysical methods to study low affinity interactions (recently reviewed in Ströh and Stehle, 2014). However, virus-glycan interaction studies have still many experimental restrictions. Glycans cannot be cloned and sequenced in contrast to nucleic acids and proteins, the usage of established molecular biology tools is often limited, the chemical synthesis of complex glycans is difficult, or the available material is very expensive. Additionally, information about their three-dimensional structures and their distribution on cells is rare. Nevertheless, rapid and high-throughput approaches using glycan arrays have led to the identification of glycan receptors or glycan receptor motifs for several viruses such as influenza virus (Childs et al., 2009), polyoma- (Neu et al., 2008; Neu et al., 2010), adeno- (Nilsson et al., 2011), reo- (Reiss et al., 2012), and picornaviruses (Imamura et al., 2014). Critical differences in glycan binding preferences of certain strains have also been determined (Stevens et al., 2006). On the array, either amine-terminating glycans are immobilized by covalent linkages to N-hydroxysuccinimide (NHS)-activated glass (Blixt et al., 2004) or lipid-inked glycans are printed on nitrocellulose-coated glass, known as neoglycolipid (NGL)-based arrays (Palma et al., 2014). Whole virus, VLPs, or the viral attachment protein alone are incubated on the array and binding is usually quantified via fluorescence-based detection systems.

In order to test for functionality of the interaction in cell-based infectivity assays virus-glycan binding can be blocked by lectins or the virus can be incubated with free glycans prior to infection (Neu et al., 2010; Reiss et al., 2012). The enzymatic removal of glycans using neuraminidases for sialic acids or heparinases and chondrotinases for glycosaminoglycans, respectively, is also possible (Schowalter et al., 2011). Treatment of cells with drugs that specifically inhibit N- and O-linked glycosylation or the synthesis of gangliosides can help to understand where the identified glycan is linked to *in vivo* (Taube et al., 2012). Moreover, small interfering RNAs (siRNAs) and cell lines deficient for certain glycans are useful tools to investigate the functional relevance of virus-glycan interactions for attachment and infection (O'Donnell et al., 2010; Schowalter et al., 2011).

Ligand-based nuclear magnetic resonance (NMR) spectroscopy techniques such as saturation transfer difference (STD) NMR that were developed for monitoring weak protein-ligand interactions (K_d in the mM- to µM-range) are suitable to determine viral binding epitopes on glycans (Reiss et al., 2012; Neu et al., 2013a). The STD NMR experiment is based on the intermolecular Nuclear Overhauser Effect (NOE) between the protein and the ligand in the case of complex formation with fast off-rates. Resonances belonging to protons of the ligand in

distances of up to 5 Å to the protein during the binding event are observed in the STD NMR spectrum after selective excitation of the protein proton resonances (Mayer and Meyer, 1999). Protein resonances are selectively saturated in a spectral region of the ^{1}H NMR spectrum that contains resonances of the protein but not the ligand (*on resonance spectrum*). The *on resonance* spectrum is then subtracted from a spectrum recorded without selective saturation (*off resonance* spectrum), usually the aromatic side chain or methyl region. In the resulting difference spectrum only proton resonances of the ligand that received saturation transfer from the protein via spin diffusion through the intermolecular NOEs are visible. A control experiment without the protein serves to verify that no direct excitation of the ligand takes place (Mayer and Meyer, 1999).

The list of biophysical methods to determine binding affinities or kinetics of virus-glycan interaction are often restricted by the availability and cost of the material for the experiment. In general, routinely used biophysical studies require due to the low binding affinity especially high amounts of expensive or only rarely available glycans. Nevertheless, biophysical methods including isothermal titration calorimetry (ITC) (Neu et al., 2008), microscale thermophoresis (MST) (Xiong et al., 2013), surface plasmon resonance (SPR) (Nilsson et al., 2011), or mass spectrometry (MS) (Han et al., 2015) have been used to study binding affinities and to determine equilibrium dissociation constants for virus-glycan interactions. ITC measures the heat that is released when a ligand binds to a protein and from this enthalpy and entropy of the binding event can be calculated. During the MST experiment the motion of molecules along microscopic temperature gradients are measured and thereby changes in the hydration shell, charge, or size are detected (Jerabek-Willemsen et al., 2011). One major advantage of SPR techniques is the possibility to investigate binding and dissociation events separately. SPR was used to measure association (k_a) and dissociation (k_d) rates for the interaction between influenza virus A strains and α2,3- and α2,6-linked Neu5Ac using an immobilized lipid bilayer containing gangliosides on the surface of the sensor chip (Hidari et al., 2007). Moreover, the binding affinity of Adenovirus 37 (Ad37) towards its glycan receptor motif GD1a has been determined by SPR (Nilsson et al., 2011). Hereby, the Ad37 fiber knob (ligand) was immobilized on the sensor ship surface and the GD1a glycan (analyte) was injected in varying concentrations and passed over the surface. A relatively high-affinity binding site (K_d of about 20 µM, with $K_d = k_d / k_a$) was detected at low glycan concentrations and additionally, a lower-affinity binding event (K_d of about 270 µM) was observed at higher glycan concentrations (Nilsson et al., 2011). Crystal structure analysis has shown that Ad37 binds the branched GD1a glycan by engaging both terminal Neu5Ac residues of its two branches. Although the proposed two-site model for binding based on the SPR measurements requires further evaluations, this study shows exemplary how a bivalent binding event (Ad37-GD1a) can increase the overall binding affinity of the virus-glycan interaction compared to a monovalent

interaction (Nilsson et al., 2011). A much lower binding affinity (K_d of about 5 mM) had been determined for the monovalent Ad37-3'SL interaction in earlier studies (Burmeister et al., 2004).

1.3 Antiviral Strategies

In addition to immunization strategies, effective antiviral agents are important for managing the emergence and re-emergence of viral infections. The fact that viruses are obligatory intracellular pathogens and rely on host cell function for replication facilitates the definition of suitable targets for antiviral therapy (Muller and Krausslich, 2009; Lou et al., 2014). However, a better understanding about structure and function of viral proteins and molecular mechanisms of virus-host interactions provides a useful platform for a rational development and discovery of new antiviral agents. Initial antiviral drugs were directed at virus-encoded enzymes, but more recently compounds inhibiting viral entry or release have been developed, and also strategies targeting host factors and the innate immune response system as well as gene silencing are pursued (Blake et al., 2012; Lou et al., 2014). Nevertheless, due to the rapid development of antiviral resistance in some cases, the use of drugs can be rather limited (Bartholomeusz and Locarnini, 2006; Okomo-Adhiambo et al., 2015).

Small molecules as attachment inhibitors have been developed for several enveloped as well as for non-enveloped viruses. For example in the case of HIV, small molecule compounds that bind in a highly conserved pocket within the HIV-1 envelope protein gp120 and interfere with the recognition of the CD4 receptor on the cell surface are tested in clinical trials (Henrich and Kuritzkes, 2013). Other examples for small compound attachment inhibitors are oseltamivir (Tamiflu™) and zanamivir (Relenza™), which were the first drugs approved for their usage in anti-flu therapy (Ison, 2011). The neuraminidase (NA) is a surface glycoprotein of influenza viruses that cleaves cell surface sialic acids, which are engaged by influenza HA and allows thereby the release of progeny virions. Oseltamivir and zanamivir bind both to the influenza NA. In particular structure-based drug design strategies including *in silico* docking- and pharmacophore-based virtual screening in combination with X-ray structure analysis have been applied to develop NA inhibitors. However, emergence and spread of viral resistance limit their use and have become an important concern particularly for influenza virus A (H1N1) and oseltamivir (Okomo-Adhiambo et al., 2015). The development and clinical approval of two novel inhibitors, laninamivir and peramivir, could at least partially overcome the oseltamivir resistance (Yamashita, 2010) or can be used when drug delivery by a route other than intravenous is not feasible (Mancuso et al., 2010).

Crystal structure analysis has also inspired the design and development of a multivalent inhibitor against Ad37, a virus that causes the severe ocular infection Epidemic

keratoconjunctivitis (EKC) (Nilsson et al., 2011; Spjut et al., 2011). Ad37 engages its glycan receptor GD1a with its trimeric fiber knob protein (Nilsson et al., 2011). A trivalent inhibitor featuring sialic acid-derived head groups is more potent than the monovalent sialic acid and also the divalent GD1a glycan (Spjut et al., 2011). So far, several carbohydrate-based or glycomimetic drugs have reached the market, but the development of therapies is often hindered by challenges such as poor absorption of the molecules and/or their rapid elimination (Ernst and Magnani, 2009). However, in the case of Ad37 an administration of the inhibitor directly to the eye, the site of infection, may circumvent many of the pharmacokinetic hurdles (Spjut et al., 2011).

Targeting interaction with sialylated receptors during viral attachment is also one rational antiviral approach for human polyomaviruses such as JCPyV and BKPyV. Although Neu5Ac binding sites of two neighboring VP1 monomers are further apart in the context of the pentamer or the capsid compared to the trimeric Ad37 fiber knob, compounds mimicking sialic acids are still promising antiviral candidates. Unique surface features on VP1 are hereby likely critical to improve binding affinities and kinetics compared to the VP1-glycan interaction. For example in the case of JCPyV, an elongated groove is connected with the Neu5Ac binding pocket and could be explored to improve binding characteristics of sialic-acid derived compounds (Neu et al., 2010). In contrast to more complex viruses, polyomaviruses possess only the capsid protein VP1 that mediates interaction with cell surface receptors and can be used as a target for attachment inhibition. However, interactions between VP1 and minor capsid proteins VP2 and VP3 or interactions between neighboring VP1 pentamers that are important for the general architecture of the virion are also rational targets (Chen et al., 1998; Nakanishi et al., 2006). Furthermore, alternative strategies, including inhibition of LT (Goodwin et al., 2009; Seguin et al., 2012; Randhawa et al., 2014) and blocking of retrograde trafficking to the ER (Nelson et al., 2013; Carney et al., 2014) have been tested. Antiviral effects of different small compounds have been observed and explored against BKPyV and JCPyV infection (Jiang et al., 2010; Rinaldo et al., 2010; O'Hara et al., 2014). For MCPyV, an attempt has been made to identify potential siRNA molecules for silencing of T antigen coding mRNA (Hoque et al., 2012). The specific knockdown of a gene function of interest makes the siRNA technology not only a powerful tool for studying gene function, but it may also be used in antiviral therapy. Promising antiviral effects of siRNAs have been shown for example against hepatitis B and C virus (HBV, HCV) (Chen et al., 2008).

In addition to small antiviral molecules, the development of effective human neutralization monoclonal antibodies against the JCPyV capsid is currently pursued in order to prevent, manage, or treat one or more conditions associated with a JCPyV infection (Simon et al., 2013; Burioni and Clementi, 2014). Neutralizing antibodies bind idealistically to a plurality of

JCPyV genotypes compromising mutations associated with PML or other clinical disorders caused by JCPyV.

Fragment-based Drug Discovery

Fragment-based drug discovery (FBDD) has emerged in the past decade as a popular alternative to high-throughput screening (HTS) for the identification of lead compounds during drug discovery (Davies and Tickle, 2012). The approach aims to identify very small molecules out of a smaller collection instead of searching huge collections containing larger molecules, which are potentially more active from the beginning. The chemical diversity can be represented in a FBDD library of only several thousand fragments, and collections with about 1,000 molecules represent the chemical diversity of libraries containing tens of millions of larger and more drug-like compounds (Hajduk et al., 2011). In the FBDD approach, small molecules are usually expanded or linked to increase specificity as well as biological potency. Fragments in FBDD compound libraries have often a higher degree of binding promiscuity and are in line with the “rule-of-three” (Congreve et al., 2003), which was named in analogy to the earlier described Lipinski’s “rule-of-five” (Lipinski et al., 2001). The rule-of-three demands a molecular weight of less than 300 Da, no more than three hydrogen bonding donors and three hydrogen bonding acceptors, and a computed partition coefficient (clogP) of less than three (Congreve et al., 2003). The partition coefficient (P) is a quantitative descriptor of lipophilicity and is defined as the ratio of the concentrations of a neutral compound in organic (for example octanol) and aqueous phases of a two-compartment system under equilibrium conditions. It can be measured experientially and is mostly used in the logarithmic form. However, several methods have been developed in order to calculate log P values based on the molecular structure (Mannhold et al., 2009). Other criteria include a polar surface area (PSA) < 60 $Å^2$, and a number of rotatable bonds of less or equal than three. Fragments according to these rules are thought to possess a high quality of binding, although they exhibit high dissociation constants (K_d ranging from µM to mM) (Davies and Tickle, 2012). Nevertheless, FBDD is facing serious problems such as the reactivity of molecules and aggregation in aqueous solution at the high concentrations that are needed for most assays to detect low-affinity binders (Erlanson, 2012). These aggregates can non-specifically inhibit proteins, thereby interfering with biochemical assays (Rishton, 2003). Another critical challenge is the poor solubility of many organic molecules at concentrations that are needed. Despite these drawbacks, FBDD has gained credibility in many drug discovery projects, especially together with improvements in complimentary methods including X-ray crystallography, biophysical methods such SPR, and computational approaches to establish structure-activity relationships (SAR) (Davies and Tickle, 2012; Murray et al., 2012).

2. Objectives

2.1 Determinants of JC Polyomavirus Cellular Attachment and Infection and the Development of Specific Antagonist for PML

At the beginning of this study it was established that JCPyV engages sialic acid receptors on host cells in order to initiate infection. Studies with the Mad-1 prototype strain of JCPyV showed that the virus binds the LSTc pentasaccharide during cell attachment and/or entry, but interactions with several other glycans, including gangliosides, were reported in other studies. It was unknown how differences that occur in VP1 proteins from different virus strains, genotypes as well as PML isolates influence glycan receptor specificity, the infectious entry pathway, and ultimately pathogenicity. Therefore, experiments were conducted in order to answer the following questions:

- What are the structural determinants of the JCPyV receptor binding specificity?
- What is the effect of VP1 mutations frequently found in JCPyV samples isolated from the brain of PML patients on virus receptor binding, glycan receptor specificity, and virus infectivity?
- Are there strain-specific or genotype-specific differences in glycan receptor binding that affect JCPyV cell tropism, the entry pathway, and pathogenicity?
- Are gangliosides important for JCPyV infection?
- How do binding affinities influence JCPyV VP1-glycan interactions? What are the requirements of the glycan receptors to promote JCPyV infection?
- What is the role of the pore at the 5-fold axis of the VP1 pentamer during JCPyV infection?
- How can the structural information be used for the development of antiviral therapies and treatment of PML?

2.2 Structural and Functional Analysis of Receptor Specificities of the Recently Discovered Human Polyomaviruses

With relatively high sequence homology and conserved overall architecture across the three existing genera, the polyomavirus family forms an attractive platform for analyzing determinants of antigenicity, cell entry, cell tropism, and host range, as well as other factors that contribute to viral pathogenesis. By comparing the receptor specificity of several newly discovered human polyomaviruses with better-studied human viruses JCPyV, BKPyV, and MCPyV, this part of the thesis aimed at understanding how structural changes of the virion enable the virus to adapt to new host cells, spread to different tissues, or to infect a new host species since all of these events are dependent on precisely regulated virus-receptor interactions. In detail, the experiments presented here were designed to address the following questions:

- What kind of cellular receptors do the more recently discovered human polyomaviruses HPyV6, HPyV7, TSPyV, HPyV12, and NJPyV engage?
- Which are the structural determinants of viruses' receptor binding specificities as well as antigenicity?
- Are the structural determinants derived by X-ray structure analysis functional in the context of cells?
- Where are glycan receptors located on the cell?
- How does the virus receptor specificity explain features of cell and host tropism and contribute to viral pathogenicity?
- How do close phylogenetic relationships of human and non-human polyomaviruses affect glycan receptor specificity, host tropism, and viral adaption?
- Is it possible to use structure-based phylogenetic analysis to predict polyomavirus-glycan receptor interactions in a reliable manner?

3. Results and Discussion

3.1 Progressive Multifocal Leukoencephalopathy-Associated Mutations in the JC Polyomavirus Capsid Disrupt Lactoseries Tetrasaccharide c Binding

Maginnis, M.S. *, **Ströh, L.J.** *, Gee, G.V., O'Hara, B.A., Derdowski, A., Stehle, T., and Atwood, W.J. (2013). Progressive multifocal leukoencephalopathy-associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding. mBio *4*, e00247-00213.
* Authors contributed equally to this work.

Studies with the Mad-1 prototype strain of JCPyV have shown that the LSTc glycan and 5-HT2Rs are important for attachment and endocytosis, respectively. However, the majority of viral isolates from PML patients contain distinct VP1 mutations that were hypothesized to arise from positive selection within the patient. The most common mutations are L54F and S268F, which account for approximately 50% of PML-associated mutations, while mutations S266F and S268Y are less frequent but occur still very often. All four mutations are localized within the LSTc binding pocket of VP1. These distinct single point mutations have never been detected in the urine of healthy individuals, nor are they present in the urine of PML patients, from whom viruses with PML-associated mutations were isolated from their CSF and blood. These findings suggest that specific mutations can arise within the JCPyV-infected individual. We sought to answer the question whether these viruses represent the pathogenic form of the virus that infects glial cells in the CNS by utilizing a structural-functional approach.
First we introduced the mutations into the JCPyV infectious genomic clone with a Mad-1 VP1 and transfected viral DNA into the permissive glial cell line SVGA. The results showed clearly that only wild type JCPyV was able to propagate. Similarly, the attachment of the mutants to SVGA cells was abolished as monitored by flow cytometry. Next, we assessed the structural effects and solved high-resolution X-ray structures of L54F, S268F, and S268Y VP1. Crystal soaking experiments with LSTc revealed that binding was abolished or severely compromised in all three cases. Then we used the JCPyV pseudovirus system to investigate the infectivity of the mutant viruses in a panel of glial and kidney cell lines. Pseudoviruses are produced by transfecting plasmids coding for viral capsid proteins VP1, VP2, and VP3 into HEK293FT cells together with a reporter plasmid that expresses the green fluorescent protein and a secreted form of luciferase (*Gaussia* luciferase). The pseudovirus assay provides a sensitive method of measuring the infectivity of mutant viruses that do not propagate using traditional cultural methods. However, also here pseudoviruses with PML-associated mutations were not infectious.

All three crystallized mutants and likely also S226F block binding to sialylated glycan motifs, independent of their location on gangliosides or glycoproteins and their glycosidic linkages. Viruses with PML-associated mutations are rendered not infectious in glial and kidney cells, suggesting that these distinct mutations may play alternative roles in PML pathogenesis. One scenario could be that PML-associated mutations arise as a strategy of immune escape. Mutations in the viral capsid, making these mutant viruses capable of evading the host immune responses, may result in increased spread to the CNS. It seems also plausible that the loss of sialic acid binding may be necessary for increased spread to the brain and subsequent infection of glial cells in a sialic acid-independent manner. This hypothesis is similar to the scenario demonstrated for a mutation in the VP1 sialic acid-binding pocket of MPyV, which leads to decreased cell binding and increases viral dissemination and thus higher pathogenicity of the virus. It is also possible that an unknown host cell factor, e.g. a non-sialic acid receptor on glial cells in the CNS, which is necessary for infection by JCPyV with PML-associated mutations, is not expressed abundantly on the tested cell lines. Noticeable, JCPyV VP1 pentamers feature an elongated groove on the surface of the pentameric capsid protein that is unique among all polyomavirus VP1 proteins crystallized to date. This groove could be involved in the engagement of other receptor structures on host cells, and such an interaction might be favored if binding to LSTc is significantly reduced or blocked. In conclusion, this study illustrates that specific engagement of cell surface glycan receptors is an important determinant of tissue tropism and possesses a critical role for JCPyV pathogenesis.

3.2 Increased Affinity of JC Polyomavirus Capsid for LSTc Over Other Sialylated Glycans is a Major Determinant of Infectivity

Ströh, L.J. *, Maginnis, M.S. *, Blaum, B.S., Nelson, C.D., Neu, U., Gee, G.V., O'Hara, B.A., Motamedi, N., DiMaio, D., Atwood, W.J. and Stehle, T. (2015). Increased affinity of JC polyomavirus capsid for LSTc over other sialylated glycans is a major determinant of infectivity. Journal of Virology *89*, 6364-6375.

* Authors contributed equally to this work.

The prototype Mad-1 strain of JCPyV uses the α2,6-linked LSTc glycan motif for initial attachment to cells, but in contrast, the JCPyV Type 3 strain “WT3”, a kidney isolate, has been reported to interact with several gangliosides although their role in JCPyV infection was not clear. In order to define whether there are strain-specific differences in glycan receptor binding specificity with possible impact on cell tropism, the viral entry pathway and eventually pathogenicity, the JCPyV brain isolate Mad-1 was compared to the WT3 strain in a structural and functional study. WT3 and Mad-1 differ in eight amino acid positions in their mayor capsid

proteins VP1 and thus, both viruses enclose the majority of strain-specific VP1 differences of the seven JCPyV genotypes. The WT3 VP1 amino acid changes were introduced into the Mad-1 capsid of infectious virus, pseudovirus, and recombinant VP1 pentamers.

High-resolution X-ray glycan complex structures of Mad-1 and WT3 VP1 pentamers reveal that in addition to the LSTc glycan motif both viruses are capable of engaging gangliosides glycans GM1, GM2, GD1a and GD1b, which feature α2,3- and α2,8-linked sialic acids. However, based on the observed interaction within in the complex structures, we hypothesized that binding affinities may modulate differences in glycan specificities between the two JCPyV strains. Here it was not possible to perform reliable affinity measurements using SPR and other techniques such as ITC have experimental and financial drawbacks when working with low affinity interactions with dissociation constants in the mM-range.

In order to obtain an estimate for binding affinities we here applied a crystallographic approach. Therefore, highly isomorphous crystals of Mad-1 and WT3 VP1 were soaked in parallel for 30 min in crystal harvesting solutions supplemented with the glycans LSTc, GM1, GM2, GD1a or GD1b in a concentration-dependent manner. Data sets were collected at beamlines X06DA and X06SA at Swiss Light Source (SLS) (Villigen, Switzerland). The data set of the native WT3 VP1 crystal was taken as a reference data set for data processing. Structure factor amplitudes and structure factor sigmas from all data sets were combined into one single file including one set of free R flags from the reference data set. Then all data sets were given the unit cell parameters from the reference data set and were also scaled to the reference data set using a low- and a high-resolution cutoff (12.0-2.4 Å). Refinement was carried out again with Refmac5 using the data set from the crystals with the respective highest glycan concentration to obtain previously solved Mad-1 and WT3 VP1 complex structures with either the LSTc, GM1, GM2, GD1a or GD1b glycan in the referenced unit cell. Simulated annealing F_{obs}-F_{calc} omit electron density maps were calculated 5 Å around the glycans and VP1 marker residues (Y80, Y160, W200, F262) for all scaled data sets. Next, masks 1 Å around the respective terminal Neu5Ac residues of all five glycans and around VP1 marker residues from each VP1 chain in the asymmetric unit were generated. Electron densities of the Neu5Ac or the marker residues were then integrated for all data sets by summation of the values of the grid points within the respective mask. The resulting electron density-derived binding curves approach saturation when the ligand occupancy in the crystal reaches a value of 1. The data from the electron density integration shows that both JCPyV strains engage the LSTc glycan with the highest affinity, but possess similar lower binding affinities for GD1a, GD1b and GM2. Assuming a simple equilibrium of the VP1-glycan interaction, dissociation constants $K_{d\ crystal}$ of about 0.5 mM and 6 mM for the strongest interaction with LSTc and the weakest interaction with GD1a glycan, respectively, could be determined using least squares

fit to the experimental data points. Of the ganglioside glycans tested GM1 binds best to both JCPyV VP1 pentamers, although the binding curves differ slightly between both strains.
The GM1 and GD1b gangliosides are functional receptors of the most closely related SV40 and BKPyV, respectively. STD NMR experiments reveal that interactions with the GM1 glycan in solution are strikingly similar for Mad-1, WT3, and SV40 VP1, but the exogenous addition of gangliosides to cells did not alter the infectivity of viruses with Mad-1 or WT3 capsids in permissive or non-permissive cells. In order to further test whether ganglioside expression contributes to JCPyV infection at all, the enzyme responsible for the ganglioside biosynthesis, the GM3 synthase, was silenced by siRNA. The finding that silencing does not affect JCPyV infection demonstrates that gangliosides are not required for attachment and infection.
Based on our functional and structural data it is unlikely that other genotype-specific VP1 variations alter the glycan specificity of JCPyV. While JCPyV VP1 can in principle engage multiple sialic acid motifs, including functional receptors of the most closely related viruses, the higher binding affinity of JCPyV for LSTc seems to be prerequisite for functional receptor engagement. It is tempting to speculate that the increased affinity for LSTc has co-evolved with a switch in the viral entry pathway towards the clathrin-mediated pathway, whereas BKPyV and SV40 employ gangliosides receptors and a cholesterol-dependent endocytosis.
Thus, our findings reconcile divergent findings from previous studies by showing that differences in glycan binding affinities and the location of the glycan on the cell, rather than possible binding events alone, determine JCPyV infection.

3.3 Modulation of a Pore in the Capsid of JC Polyomavirus Reduces Infectivity and Prevents Exposure of the Minor Capsid Proteins

Nelson, C.D., **Ströh, L.J.**, Gee, G.V., O'Hara, B.A., Stehle, T., and Atwood, W.J. (2015). Modulation of a symmetry-related pore in the capsid of JC polyomavirus reduces infectivity and prevents exposure of the minor capsid proteins. Journal of Virology *89*, 3910-3921.

The icosahedral viral capsid of JCPyV is composed primarily of the major capsid protein VP1. Due to the pentameric arrangement of VP1 monomers, a pore-like structure arises at the five-fold axis of symmetry. The sequence of lining residues and architecture of this pore is rather conserved among polyomaviruses, with accessible gap diameters from 7.6 Å (WUPyV) to 8.6 Å (MPyV) at the respective bottleneck. Thus, VP1 has either limited structural plasticity in this region or the pore could be important during viral infection or assembly of the virion.
In order to investigate the role of the five-fold gap several amino residues mutations lining the five-fold gap were introduced into VP1 pentamers, pseudoviruses, and infectious viruses. Functional studies showed that exposure of VP2 upon arrival to the ER, a step that is critical

for infection, was impaired in mutated pseudoviruses, whereas viral assembly, packaging of minor capsid proteins, binding to cells, and transport to the host cell ER were not affected. X-ray crystal structures of pentamers from three representative mutants demonstrate that only subtle changes of the five-fold gap diameter, but an increased hydrophobicity or minor changes of the diameter or the molecular surface might alter, and perhaps strengthen, the contacts with VP2. As known, VP2 egress does not appear to be accompanied by complete disassembly of the viral capsid and thus, an exposure through inter- and intra-pentameric gaps as well as the 5-fold gap seems possible. It is also possible that these substitutions alter a recognition site for an as of yet unidentified ER-resident chaperone or inhibit productive association with ER-resident chaperones. Another possibility is that these pore mutants are successfully transported to the ER but are deficient in transport to critical regions of the ER necessary for VP2 exposure and do not associate with chaperones such as BIP and DNAJ and retrotranslocation.

This study demonstrates that the five-fold pore is not only an artifact due to the five-fold symmetry of the pentamer but is rather a very important structural feature of JCPyV and likely other polyomaviruses. Modifications to this structure result in a severe reduction in infectivity suggesting that small molecules targeting this part of VP1 may be a promising approach for the development of antiviral strategies.

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

Ströh, L.J., Neu, U., Blaum, B.S., Buch, M.H.C., Garcea, R.L., and Stehle, T. (2014). Structure analysis of the major capsid proteins of human polyomaviruses 6 and 7 reveals an obstructed sialic acid binding site. Journal of Virology *88*, 10831-10839.

Serological studies within the human population suggest that a persistent infection with HPyV6 and HPyV7 is very common, but routes of infection, cell tropism and entry pathways are unknown. In immunosuppressed individuals HPyV7 can cause distinctive epidermal hyperplasia, but although the pathogenicity of HPyV6 remains to be proven, an involvement in cutaneous diseases or tumors has also to be considered. In order to investigate critical determinantes for antigenicity and cell-surface receptor specificities that are linked to attachment, tropism and ultimately viral pathogenicity, high-resolution X-ray structures of HPyV6 and HPyV7 VP1 pentamers were determined. Our structural analysis revealed uniquely elongated loops that cover the bulk of the outer virion surfaces, and moreover, obstruct the groove that binds sialylated glycan receptors in the case of related polyomaviruses. In addition, NMR spectroscopy and single-cell binding studies indicate that sialylated glycans are likely not

required for initial attachment of both viruses to human cells. Overall findings elucidate distinct antigenic properties of HPyV6 and HPyV7 capsids and moreover highlight the need for structure-based comparisons to better define and predict phylogenetic relationships among the raising number of human polyomaviruses.

3.5 *Trichodysplasia Spinulosa*-associated Polyomavirus Employs a Displaced Binding Site on VP1 to Engage Sialic Acid Containing Glycolipids

Ströh, L.J., Gee, G.V. , Blaum, B.S., Dugan, A.S., Feltkamp, C.W., Atwood, W.J. and Stehle, T. (2015). *Trichodysplasia spinulosa*-associated polyomavirus employs a displaced binding site on VP1 to engage sialylated glycolipids. PLoS Pathogens *11*, e1005112.

The skin-tropic TSPyV is the fourth human member of the polyomavirus family for which clear evidence of a direct involvement in a clinical disease was shown. The pathogenic mechanism includes hyperproliferation of IRS cells, but molecular determinants underlying the infection and the associated disease remain unknown. Using cell binding and pseudovirus infection studies, we found that sialylated glycans are required for TSPyV attachment and infection of cultured human cells. To gain further knowledge about glycan specificity of the virus, X-ray structures of TSPyV VP1 pentamers in complex with the branched GM1 glycan and the linear compounds 3'SL and 6'SL were solved. TSPyV VP1 features a unique glycan binding site, which is shifted by about 18 Å from sialic acid binding sites of other structurally studied polyomaviruses. The Neu5Ac ring of all three sialylated compounds is recognized in this binding site, which is built up by the exposed BC2-loop of VP1. Structure-based site-directed mutagenesis reduces binding of VP1 pentamers to cells and pseudovirus infectivity, highlighting the importance of this novel sialic acid binding site for TSPyV infection. In addition, treatments of cells with inhibitors for glycosphingolipid synthesis, O- or N-glycosylation prior to cell binding and pseudovirus infection studies suggest that especially glycolipids may play an important role during attachment and entry.

According to VP1 amino acid sequence alignments, TSPyV is most closely related to the Bornean Orangutan Polyomavirus (OraPyV-Bo). Interestingly, the core of the Neu5Ac binding site is shared with the most closely related simian virus and may be employed for receptor binding by other known or still unknown members of the growing polyomavirus family. Our correlation of structural and functional data provides first insights into molecular recognition events during TSPyV infection and, furthermore, demonstrates clearly that glycan receptor recognition of (polyoma-) viruses is fine-tuned to a greater extent than initially predicted, providing clues about determinants of host adaption and viral evolution.

3.6 Glycan Engagement by Viruses: Receptor Switches and Specificity

Ströh, L.J., and Stehle, T. (2014). Glycan engagement by viruses: receptor switches and specificity. Annual Review of Virology *1*, 285-306.

Many viruses, including several human pathogens, employ cell surface glycans as attachment receptors or attachment factors. Hereby, the engagement of glycans facilitates not only viral attachment to cells and cellular entry, which are the first steps during a viral infection, but consequently also determines host range, tissue tropism, pathogenicity, and transmissibility of viruses. Based on representative crystal structures of viral attachment proteins in complex with glycans this review focuses on current knowledge about virus-glycan interactions including mainly the recognition of sialic acids and non-charged oligosaccharide such as HBGAs. General principles and determinants of specificity utilized by different glycan-binding viruses were highlighted, and in particular, we focused on the potential of these interactions for switching receptor specificities within or even between glycan classes. Although numerous viruses have been found to interact with a third class of cell surface glycans, the polyanionic GAGs, high-resolution complex structures of viral receptor-binding proteins with motifs or analogs, which were built unambiguously into electron density, are very rare due to various reasons such as sample heterogeneity. Thus, an understanding on the molecular level about specificities and conformational changes of virus-GAG interactions is still very limited.
In conclusion, recent advances in glycan microarray screening coupled with high-resolution structural studies and reverse genetics systems provide a powerful platform for the identification and characterization of viral glycan receptors. However, further technical improvements are needed for the analysis and quantification of cell surface glycans densities and developmentally regulated host- and tissue-specific glycan distributions. In addition, a correlation of affinity measurements with structural studies seems to be necessary for a more rigorous evaluation of glycan specificity and receptor switching potentials in order to establish a comprehensive toolbox, which can then be used for the design of therapeutic gene vector and new antiviral drugs.

3.7 Ongoing Research

Human Polyomavirus 12 and New Jersey Polyomavirus

Current research focuses on the structural and functional characterization of VP1 from the two human polyomaviruses HPyV12 and NJPyV.

Cloning, Expression and Protein Purification

A synthetic *E. coli* codon-optimized gene (Mr.Gene GmbH) coding for amino acids 44 to 316 of HPyV12 VP1 (NCBI accession code: YP_007684355) was amplified by PCR and cloned into the pET15b vector (Novagen) via *Nde*I and *Bam*HI restriction enzymes in frame with an N-terminal hexahistidine-tag (His-tag) and a thrombin cleavage site. The pET15b-derived expression plasmid for NJPyV VP1 (amino acids 36 to 323; NCBI accession code: YP_009030020) was obtained via the InFusion HD cloning strategy according to the manufacturer's protocol (Clontech). The vector was linearized via *Xho*I.
Recombinant expression in *E. coli* Rosetta 2 (DE3) and *E. coli* BL21 (DE3) for N- and C-terminal truncated NJPyV VP1 and HPyV12 VP1, respectively, was induced by addition of 0.4 mM IPTG and carried out at 20°C for about 18 h. The truncated VP1 cannot assemble into a capsid but forms pentamers in solution (Stehle and Harrison, 1997). Proteins were first purified by nickel affinity chromatography and then by size exclusion chromatography on a Superdex-200 column. For crystallization, the His-tag was cleaved off in solution with thrombin (GE Healthcare) prior to size exclusion chromatography leaving the non-native amino acid sequence GSHM at the N-terminus of both VP1 proteins. After size exclusion chromatography, HPyV12 VP1 pentamers were kept in 20 mM HEPES pH 7.5, 150 mM NaCl, 20 mM Dithiothreitol (DTT). Dimers of NJPyV VP1 pentamers were stored in an equivalent buffer without DTT.

Crystallization, Data Collection, and Structure Determination

HPyV12 VP1 was concentrated to 7 mg/ml and crystallized at 20°C by hanging drop vapor diffusion against a reservoir solution containing 4% (v/v) Tacsimate pH 7.0 and 16% (w/v) polyethylene glycol (PEG) 3,350. These crystals were not suitable for crystal soaking experiments because all five binding sites of VP1 were blocked in the crystal by a symmetry-related protomer. Additionally, crystals of HPyV12 VP1 were grown by hanging drop vapor diffusion using a protein concentration of 6 mg/ml and a reservoir solution containing 0.2 M ammonium acetate, 0.1 M Bis Tris pH 7.0, and 45% (v/v) 2-Methyl-2,4-pentanediol (MPD). These crystals were soaked for 24 h in the reservoir solution supplemented with 20 mM α2,3-sialyllactosamine (3'SLN) in order to solve the complex structure (Table 3).

Crystals of NJPyV VP1 obtained at 4°C using the sitting drop vapor diffusion method, a reservoir solution containing 0.1 M succinic acid pH 7.0 and 15% (w/v) PEG 3,350, and a protein solution with a concentration of 7 mg/ml yielded the native structure of the protein (Table 3). Crystals grown in 0.1 M MES pH 6.5, 12% (w/v) PEG 20,000 at 20°C were derivatized by incubation for 10 min in a reservoir solution supplemented with 50 mM 3'SL.
For the crystallization of HPyV12 VP1 by the hanging drop technique, drops were set up with 1 µl protein solution and 1 µl reservoir solution. 0.4 µl of the protein and 0.4 µl of the reservoir solution were used for the crystallization of NJPyV VP1 using the sitting drop vapor diffusion method. All crystals were transferred for 2 sec into the respective reservoir or soaking solution, supplemented with 30% (v/v) glycerol prior to freezing in liquid nitrogen. Data sets were collected at beamline X06DA of the SLS and at beamline 14.1 of the Berliner Elektronenspeicherring Gesellschaft für Synchrotronstrahlung (BESSY), respectively. Diffraction data sets were processed with XDS (Kabsch, 2010). HPyV12 VP1 and NJPyV VP1 structures were solved by Molecular replacement (MR) using Phaser MR (McCoy et al., 2007) included in the CCP4 program suite (Winn et al., 2011). The native MCPyV VP1 structure (pdb: 4FMG), modified by CHAINSAW (Stein, 2008), was taken as MR model for HPyV12 and NJPyV VP1. Rigid body and simulated annealing refinement was carried out with Phenix (Adams et al., 2010). Subsequently, alternating rounds of model building in Coot (Emsley et al., 2010) and refinement with Refmac5 (Murshudov et al., 1997), including 5-fold non-crystallographic symmetry (NCS)-restraints, the translation-libration-screw (TLS) method (Painter and Merritt, 2006), CCP4 library restraints, and user-defined restraints for the α2,3-glycosidic bond were performed.

Table 3: Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

	HPyV12 VP1	**HPyV12 VP1-3'SLN**	**NJPyV VP1**	**NJPyV VP1-3'SL**
Beamline	SLS X06DA	SLS X06DA	BESSY 14.1	BESSY 14.1
Data collection				
Space group	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_1$
a, b, c [Å]	61.19, 136.31, 85.33	83.44, 141.91, 251.49	86.48, 151.76, 130.91	86.38, 151.08, 130.62
β [°]	109.57	90	106.85	106.56
Resolution [Å]	50-1.50 (1.54-1.50)	50-1.90 (1.95-1.90)	50-2.3 (2.36-2.30)	50-1.80 (1.85-1.80)
Unique reflections	198,796 (13,662)	234,039 (16,726)	143,046 (10,546)	295,092 (21,596)
Total reflections	768,174 (51,438)	1,655,154 (102,457)	828,661 (60,401)	1,695,619 (115,347)
R_{meas} [%]	8.0 (174.5)	11.9 (115.7)	25.1 (104.5)	15.4 (124.0)
I/σI	15.4 (2.2)	14.3 (2.0)	6.7 (1.6)	9.6 (1.4)
$CC_{1/2}$ [%]	99.9 (47.4)	99.9 (75.7)	98.4 (63.8)	99.6 (56.6)
Completeness [%]	94.7 (87.0)	99.7 (97.2)	99.9 (99.9)	99.8 (98.9)
Wilson B-factors [$Å^2$]	23.7	31.1	29.3	26.1
Refinement				
R_{work} / R_{free} [%]	16.1 / 18.5	21.7 / 24.8	21.2 / 23.1	18.1 / 20.1
No. of atoms				
protein	10,621	19,902	21,307	21,519
water	1,424	1,077	477	1,813
glycan	-	276	-	419
B-factor [$Å^2$]				
protein	15.6	31.2	29.2	23.9
water	27.8	28.7	21.3	28.5
glycan	-	36.7	-	30.1
R.m.s.d.				
Bond length [Å]	0.008	0.007	0.009	0.011
Bond angles [°]	1.337	1.331	1.524	1.549

Overall Structures of HPyV12 and NJPyV VP1 Pentamers

The structures of the unassembled VP1 pentamer from HPyV12 and NJPyV feature the typical conserved jelly-roll core of the polyomavirus VP1. In the crystal of the HPyV12 VP1-3'SLN complex structure and in both NJPyV VP1 structures two pentamers stack against one another with their bottom surfaces. The conserved cysteine residues in the CD-loops, C119 of HPyV12 VP1 and C113 of NJPyV VP1, respectively, are engaged in interpentamer disulfide bonds in each chain. From the structure of the SV40 virion it is already known that the CD-loops adopt different conformations, some of which are engaging in interpentamer disulfide bonds (Liddington et al., 1991; Stehle et al., 1996). In other structures of unassembled polyomavirus VP1 pentamers, DE-loops are also characterized by relatively high temperature factors (B-factors) or could only be built in defined electron density when fixed by crystal contacts.
HPyV12 shares a VP1 amino acid sequence identity of 53-58% for the full-length protein with human HPyV9, JCPyV, BKPyV, TSPyV, and NJPyV as well as with other structurally investigated viruses SV40, MPyV, and LPyV. Only Wukipolyomaviruses, MWPyV, and STLPyV possess a lower amino acid sequence identity for VP1 with HPyV12 (23-46%). Nevertheless, Cα atoms of the HPyV12 VP1 monomer superimpose with the lowest root mean square deviation (r.m.s.d.) value of 0.7 Å onto JCPyV VP1, but other human VP1 are also structurally conserved in a similar magnitude (r.m.s.d. values of 0.8 Å for BKPyV VP1 to 1.2 Å for NJPyV VP1). Based on the amino acid sequence of the full length VP1, NJPyV is most closely related with TSPyV and HPyV9 (59% sequence identity). The superimposition with VP1 structures from other Orthopolyomaviruses reveals r.m.s.d. values between 1 Å for MCPyV and 1.3 Å for MPyV with a value of about 1.2 Å for TSPyV and HPyV9.

HPyV12 VP1 Binds Specifically Terminal Sialic Acids

In the complex structure of the HPyV12 VP1 pentamer with 3'SLN, the Neu5Ac moiety of the linear trisaccharide is engaged in a binding site mainly formed by the BC2-loop on top of one VP1 monomer (Figure 11). Well-defined electron density was seen for Neu5Ac in all ten VP1 binding sites in the asymmetric unit, whereas in some sites the Gal could also be unambiguously built into the electron density (Figure 11 B). However, only specific interactions of VP1 with the terminal Neu5Ac were observed in the complex structure. The Gal does not interact with VP1, but it is likely that conformational restraints of the glycosidic bond stabilize the Gal ring in a preferred orientation. The N-acetyl group of Neu5Ac inserts into a shallow cavity (Figure 11 C), where it is recognized by a hydrogen bond between residue T85 and the -NH group. Additionally, it depends on non-polar interactions between the methyl group and residues P95 and H144 of the ccw monomer at one site and on the bottom of the cavity, respectively. The oxygen of the N-acetyl group interacts by hydrogen bonding with the

backbone amine of residue R94. The glycerol chain of Neu5Ac forms hydrogen bonds to residue R149 from the ccw DE-loop and with the backbone amine of residue V86. On the other side of the BC2-loop, the C4 hydroxyl group of Neu5Ac interacts with the protein backbone. Interestingly, the carboxylate group of Neu5Ac is not involved in any direct interactions with VP1. Cα atoms of the BC2-loop (residues 83-98) from the liganded HPyV12 VP1 structure superpose with the respective residues of the native structure with an r.m.s.d. of 0.3 Å, indicating that no major conformational changes occur upon ligand binding. Neu5Ac docks into a preformed binding pocket.

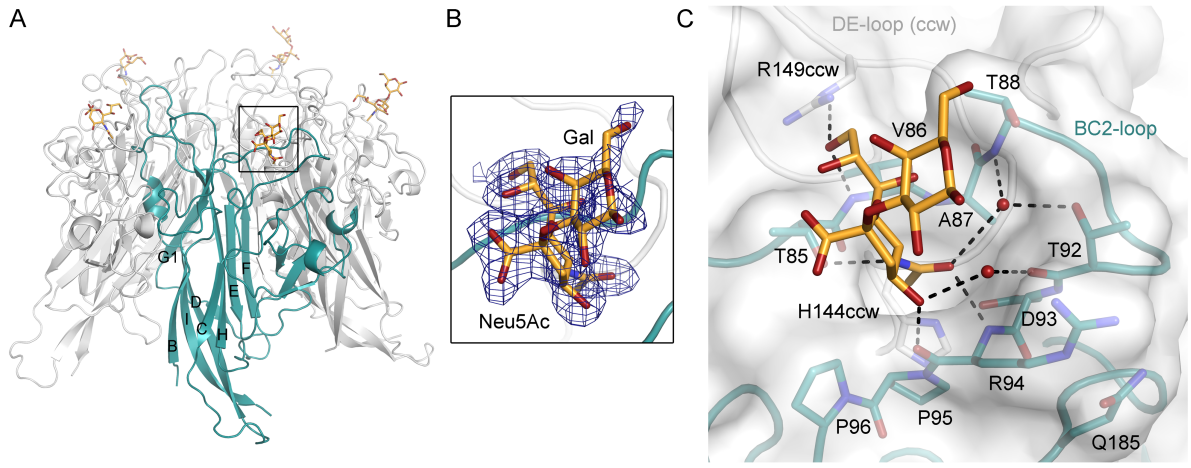


Figure 11: Structure of HPyV12 VP1 in complex with 3'SLN. (A) Structure of the HPyV12 VP1 pentamer in complex with 3'SLN. One VP1 monomer is highlighted in deep teal, the others are colored gray. The carbohydrate ligand is shown in stick representation, with oxygens in red and nitrogens in blue. (B) The simulated annealed difference electron density map for the terminal Neu5Ac-Gal motif of 3'SLN bound to HPyV12 VP1 is contoured at a σ level of 2.5 and displayed with a radius of 2 Å around the ligand. (C) Binding surface for 3'SLN. VP1 side chains in the BC2-loop binding site are shown in sticks. Interactions are indicated with black dashed lines. Water-mediated interactions are also shown with water molecules depicted as spheres.

Next, we sought to validate the VP1-Neu5Ac interaction in solution and on the surface of cells. STD NMR measurements were performed with HPyV12 VP1 pentamers, 3'SL, and 6'SL as described in 3.2. The spectra reveal that HPyV12 VP1 interacts with both glycans in solution but no additional glycan rings besides the respective terminal Neu5Ac residue were involved in binding (Figure 12 A). These findings are consistent with the crystal structure analysis. In addition, attachment of HPyV12 VP1 pentamers to cultured cells depends on cell surface sialic acids as assessed by flow cytometry according to the protocol described in 3.5. Cells were mock-treated or incubated with *Clostridium perfringens* neuraminidase prior to the binding experiments (Figure 12 B). The *Clostridium perfringens* neuraminidase cleaves α2,3- and α2,6-linked sialic acids.

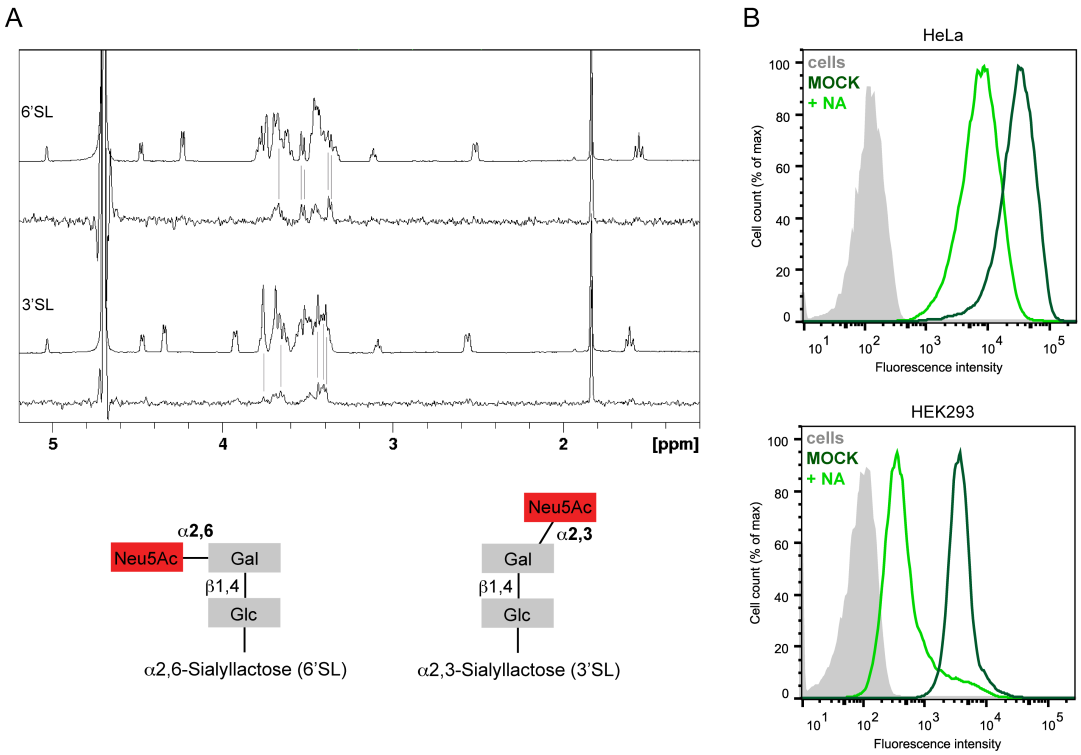


Figure 12: STD NMR spectroscopy and cell binding analysis of HPyV12 VP1 pentamers. (A) From top to bottom: ^{1}H NMR reference spectrum of 50 µM HPyV12 VP1 with 1 mM of 6'SL, STD NMR difference spectrum recorded with the same sample, ^{1}H NMR reference spectrum of 50 µM HPyV12 VP1 with 1 mM of 3'SL, STD NMR difference spectrum with the same sample. STD NMR experiments were performed by Bärbel Blaum, Tübingen. (B) HeLa and HEK293 cells were mock (PBS) or pretreated with 0.2 U/ml *Clostridium perfringens* neuraminidase (+ NA) for 30 min, washed and then incubated with Alexa Fluor 488 conjugated HPyV12 VP1 pentamers according to the protocol described in chapter 3.5. VP1 pentamer binding to cells was then analyzed by flow cytometry. Histograms represent the fluorescence intensity of Alexa Fluor 488 for 15,000 gated events in each case. Cells alone are colored gray. Three independent experiments were performed and results of a typical experiment are presented.

NJPyV VP1 Binds Specifically to 3'SL

In contrast to the broader binding specificity observed for HPyV12, the glycan complex structure of NJPyV VP1 suggests that the virus engages α2,3-linked sialic acids in a more specific manner. NJPyV VP1 interacts with the α2,3-Neu5Ac-Gal motif of 3'SL by polar and non-polar interactions in a binding site that is located between HI- and BC-loops of one VP1 monomer (Figure 13). The DE-loop, which is remarkably elongated in NJPyV VP1 compared to other structurally characterized polyomavirus VP1 pentamers, is not involved in any interactions with the glycan. Instead, side chains of residues K296 and N299 from the HI-loop and of BC2-loop residue H80 engage the carboxylate group of Neu5Ac. Residue E77 interacts with the C4 hydroxyl group of Neu5Ac and the N-acetyl group is involved in a water-mediated hydrogen bond with the backbone of residue V78. Additionally, hydrophobic interactions between the methyl group of the N-acetyl group, residue L90, and the hydrophobic part of the

E77 side chain are observed, but the glycerol chain of Neu5Ac is pointing away and is not engaged by VP1. The Gal ring interacts with residue D298 via a hydrogen bond formed with by the C4 hydroxyl group and via hydrophobic interactions of its C6. Furthermore, there is a weak hydrogen bond with a distance of about 3.7 Å between K296 and O4 of the Gal. Cα atoms from structures of native and liganded NJPyV VP1 superpose with a r.m.s.d value of 0.2 Å and no major rearrangements upon glycan binding were observed.

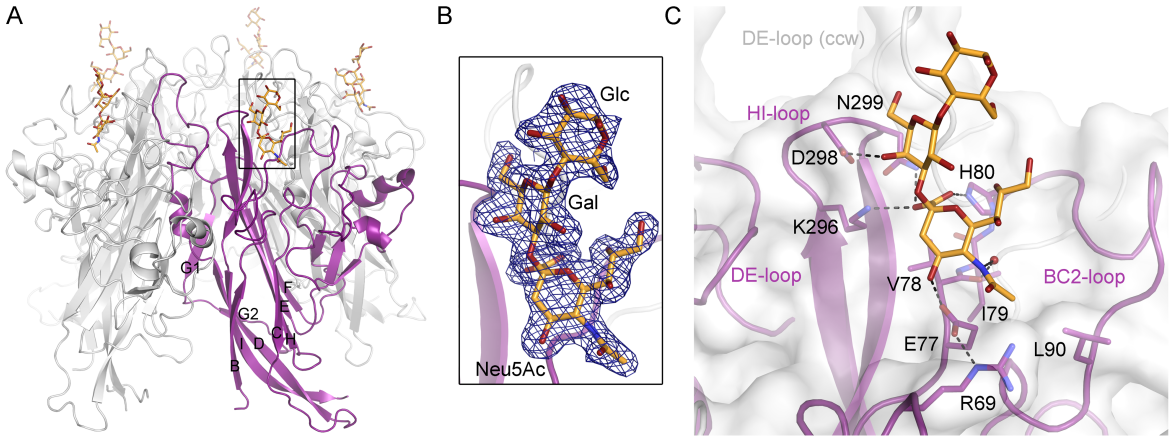


Figure 13: Complex structure of NJPyV VP1 with 3'SL. (A) NJPyV VP1 pentamer structure in complex with the 3'SL glycan. One VP1 monomer is highlighted in purple, the others are colored in gray. The glycan is shown in stick representation with carbons in orange, oxygens in red, and nitrogens in blue. (B) The simulated annealed difference electron density map for the engaged ligand 3'SL is contoured at a σ level of 2.5 and is shown with a radius of 2 Å around the ligand. (C) Interactions with 3'SL. Side chains of VP1 in the binding site are shown in sticks. Direct and water-mediated interactions are indicated with dashed lines. A water molecule in the binding site is shown (red sphere).

Next, STD NMR spectroscopy was used to rationalize the interaction seen in the crystal structure (Figure 14). The α2,3-sialic acid residue of 3'SL is bound by NJPyV VP1 pentamers in solution, but interactions with 6'SL and 2-O-methyl sialic acid were not detected. For 3'SL only resonances belonging to the sialic acid were observed and peaks for protons H1 and H3 of the Gal are not visible in the STD NMR difference spectrum. However, although Gal engages in fewer contacts compared to Neu5Ac in the crystal structure and contributes only about 30% of the total VP1 glycan interface area (Krissinel and Henrick, 2007), its interactions with VP1 seems to be critical for the glycan recognition event. A terminal sialic acid alone represented by 2-O-methyl sialic acid is not sufficient for the interaction with VP1 (Figure 14 D and E).

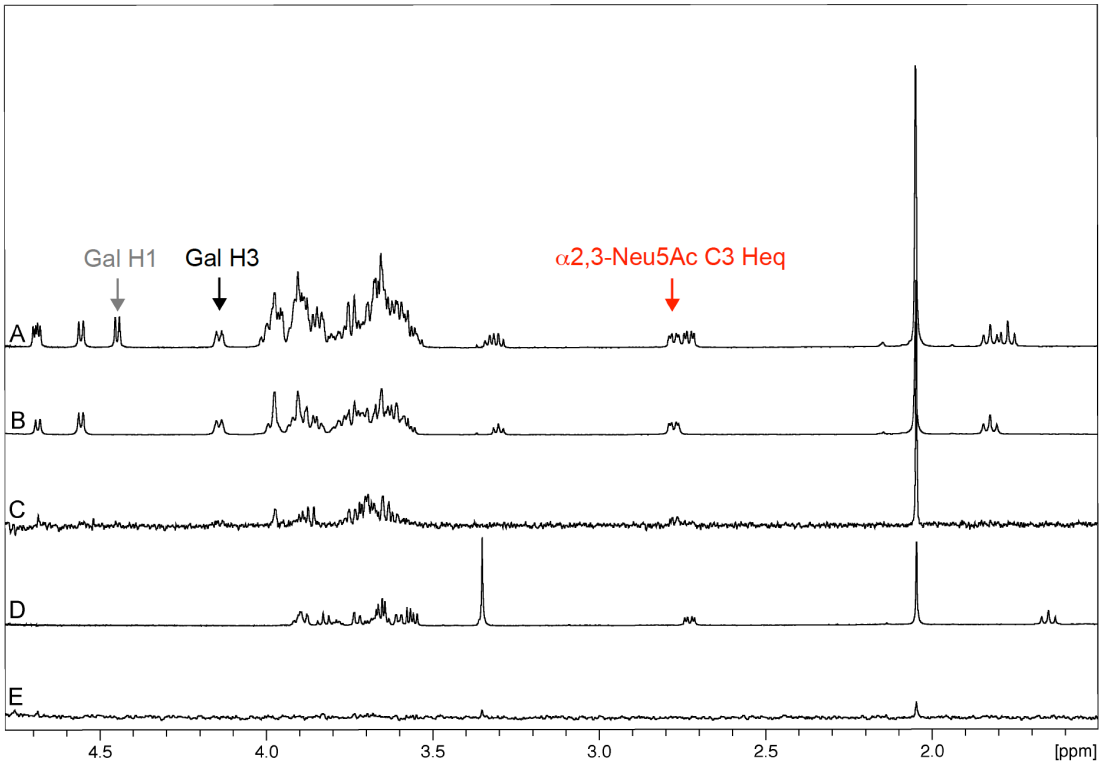


Figure 14: STD NMR spectroscopy of NJPyV VP1 with 3'SL, 6'SL, and 2-O-methyl sialic acid. (A) ^{1}H NMR reference spectrum of 50 µM NJPyV VP1 with 1 mM of 6'SL and 1 mM of 3'SL. (B) ^{1}H NMR reference spectrum of 1 mM of 3'SL. (C) STD NMR difference spectrum of the same sample as in A. (D) ^{1}H NMR reference spectrum of 50 µM HPyV12 VP1 with 1 mM 2-O-methyl sialic acid. (E) STD NMR difference spectrum of the same sample as in D. The red arrow marks the equatorial H3 resonance of α2,3-Neu5Ac, which is slightly shifted with respect to the same resonance of α2,6-Neu5Ac and serves to distinguish which of the two compounds receives magnetization transfer from the protein in spectrum C. Only peaks for the 3'SL proton are visible in the STD NMR difference spectrum in C. H1 (grey arrow) and H3 (black arrow) of the Gal ring are not visible in C.

Comparison of Glycan Receptor Binding Sites from HPyV12 and NJPyV with Those from Other Polyomaviruses

The superimposition of VP1-glycan complex structures from HPyV12 and NJPyV VP1 with liganded VP1 structures from other polyomaviruses highlights that HPyV12 engages the terminal Neu5Ac highly similar to TSPyV (Figure 15).

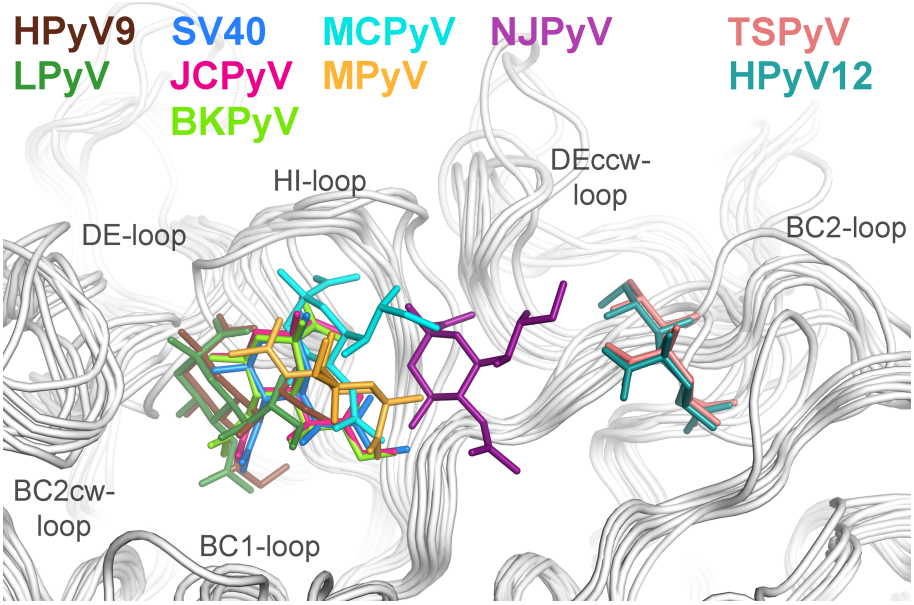


Figure 15: Glycan binding sites of HPyV12 and NJPyV VP1 in comparison with other polyomavirus sialic acid binding sites. The structural superposition of polyomavirus VP1-glycan complex structures highlights the variability of Neu5Ac engagement on VP1. HPyV12 VP1-3'SLN, NJPyV VP1-3'SL, SV40 VP1-GM1 glycan (pdb: 3BWR), BKPyV VP1-GD3 glycan (pdb: 4MJ0), JCPyV VP1-LSTc (pdb: 3NXD), MCPyV VP1-GD1a glycan (pdb: 4FMJ), LPyV VP1-3'SL (pdb: 4MBY), MPyV VP1-Neu5Ac-α2,3-Gal-β1,3-[α2,6-Neu5Ac]-GlcNAc-β1,3-Gal-β1,4-Glc (pdb: 1VPS), HPyV9 VP1-3' Neu5Gc-SL (pdb: 4POT), TSPyV VP1-GM1 glycan (pdb: 4U60).

A more detailed comparison reveals that residues of the binding site are highly conserved between VP1 from HPyV12 and TSPyV (Figure 16). Thus, the BC2-loop folds in a very similar manner on top of both VP1 and allows conserved contacts with Neu5Ac.

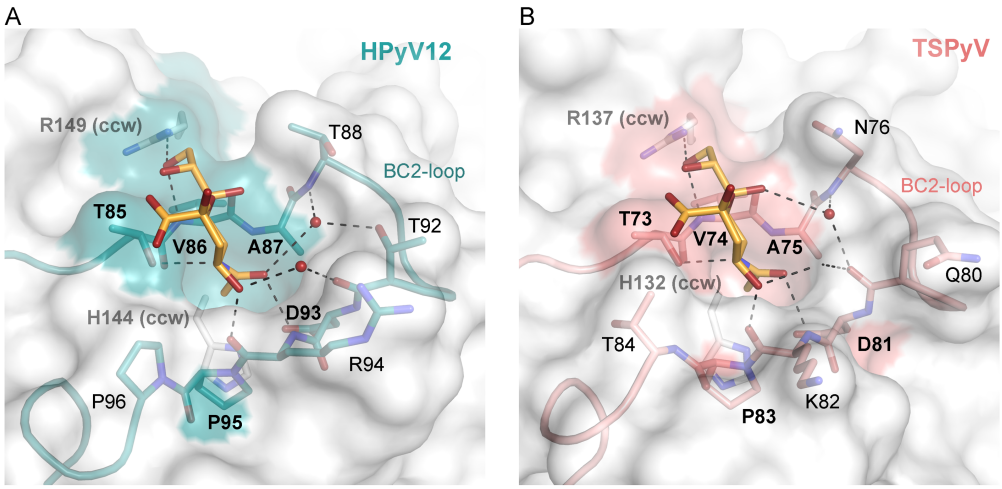


Figure 16: Glycan binding sites of HPyV12 and TSPyV. Neu5Ac binding sites of (A) HPyV12 and (B) TSPyV VP1 (pdb: 4U60) are shown in the same orientation. VP1 complex structures were aligned using Cα atoms and the secondary-structure matching (SSM) tool in Coot (Emsley et al., 2010). One VP1 monomer is highlighted in deep teal and salmon for HPyV12 and TSPyV, respectively. Glycan residues are show in orange sticks with carbons in orange, oxygens in red and nitrogens in blue. Glycan residues are colored grey if they are not engaged by VP1. Direct and water-mediated hydrogen bonds between VP1 side chain residues and the terminal Neu5Ac are depicted as black dashed line. Conserved residues are colored on the VP1 surface and are labeled in bold.

In contrast, NJPyV employs a Neu5Ac binding site between the BC- and HI-loops of one VP1 monomer. This binding site is related to MCPyV, in means of its location and the orientation of the respective Neu5Ac engaged (Figure 15), although the amino acid sequence identity in the binding region is not very high. VP1 amino acid differences between MCPyV and NJPyV result in a shift of the binding site towards the BC2-loop (Figure 17). However, the N-acetyl chain and the carboxylate group of Neu5Ac are engaged in a very similar manner on NJPyV and MCPyV VP1, although specific interactions are realized by different amino acids. A common feature of the Neu5Ac binding sites of SV40, JCPyV, BKPyV, HPyV9, and LPyV is a hydrophobic pocket, which is formed by residues of HI- and BC-loops from one VP1 monomer and BC2-loop residues of the cw VP1 monomer (Neu et al., 2008; Neu et al., 2010; Neu et al., 2013a; Neu et al., 2013b). The location of the hydrophobic pocket is not conserved on NJPyV VP1, but instead BC2-loop residues E77 and L90 form a hydrophobic environment for the methyl group of Neu5Ac on NJPyV VP1 in a different location.

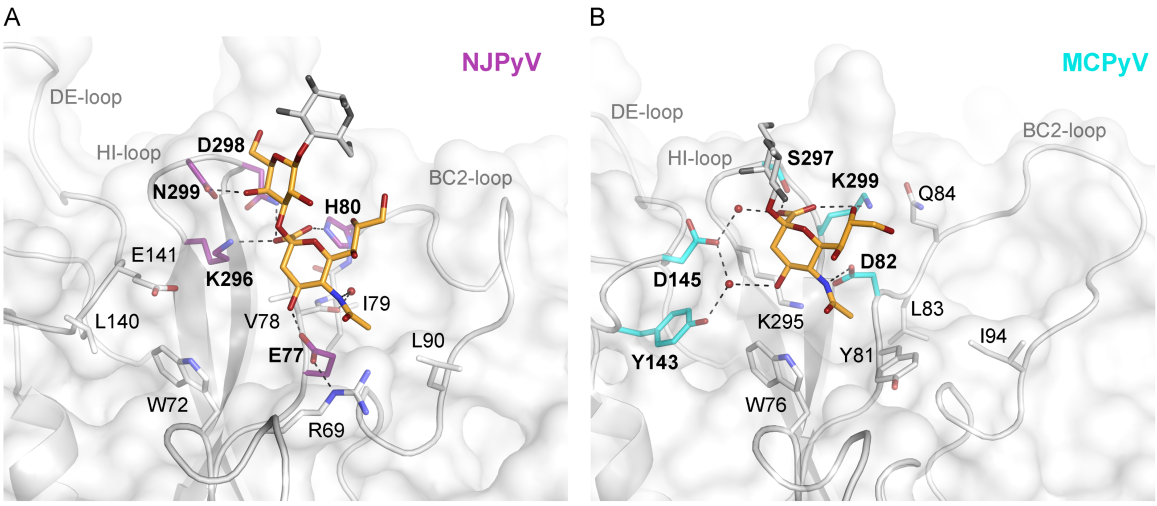


Figure 17: Comparison of glycan binding sites on NJPyV and MCPyV VP1. VP1 glycan binding sites from (A) NJPyV and (B) MCPyV (pdb: 4FMJ) are shown in the same orientation. Glycan residues not involved in intermolecular interactions are colored in grey. Direct and water-mediated hydrogen bonds between VP1 side chain residues and the terminal Neu5Ac are depicted as black dashed line. Side chains of residues interacting with the glycan via hydrogen bonding or charged interactions are colored in purple and cyan for NJPyV and MCPyV, respectively.

The phylogenetic analysis of the polyomavirus family indicates that many human polyomaviruses are closely related to one or two monkey polyomaviruses. These relationships are also present on the amino acid level of VP1 (Figure 1). NJPyV shares a VP1 amino acid sequence identity of about 84% with ChPyV isolates (ChPyV-Azzi, ChPyV-Bob, and ChPyV-Tanu) and of about 73% with two other primate polyomaviruses, Vervet monkey Polyomavirus 1 (VmPyV1) and the Piliocolobus rufomitratus Polyomavirus 1 (PrufPyV1). In the phylogenetic tree based on the VP1 amino acid sequence HPyV12 clusters with Avipolyomaviruses, including Finch Polyomavirus (FiPyV) and Budgerigar fledgling disease Polyomavirus (BFDPyV), but also Crow Polyomavirus (CrowPyV), Butcherbird Polyomavirus (BbPyV), Goose hemorrhagic polyomavirus (GhPyV) and Canary Polyomavirus (CaPyV) have amino acid sequence identities of about 60% with HPyV12 (Figure 1). In order to rationalize the impact of VP1 differences on receptor specificity of closely related viruses from different species, we mapped amino acid differences on the surfaces of HPyV12 and NJPyV VP1 pentamer structures, respectively (Figure 18). Differences occur predominantly within the exposed surface loops on top of VP1. Interestingly, the elongated DE-loop, a prominent feature of NJPyV VP1, is unique for the human virus, whereas VP1 residues involved in the recognition of the Neu5Ac-Gal motif are highly conserved in ChPyV, VmPyV1, and PrufPyV1 (Figure 18 B). The BC2-loop Neu5Ac binding region of HPyV12 VP1 may also be partially conserved in avian viruses FiPyV and BFDPyV, whereas most of the top surface differs (Figure 18 C and D). A broader VP1 sequence alignment reveals that key residues of the

HPyV12 BC2-loop Neu5Ac binding region, in contrast to the rest of the VP1 top surfaces, are conserved across Avipolyomavirus species, EPyV, and TSPyV but not in other Orthopolyomaviruses (Figure 18 E). Nevertheless, predictions of the surface loop architecture can be rather error-prone, hampering modeling and mapping of glycan binding sites and the prediction of glycan receptor specificities based on sequence information.

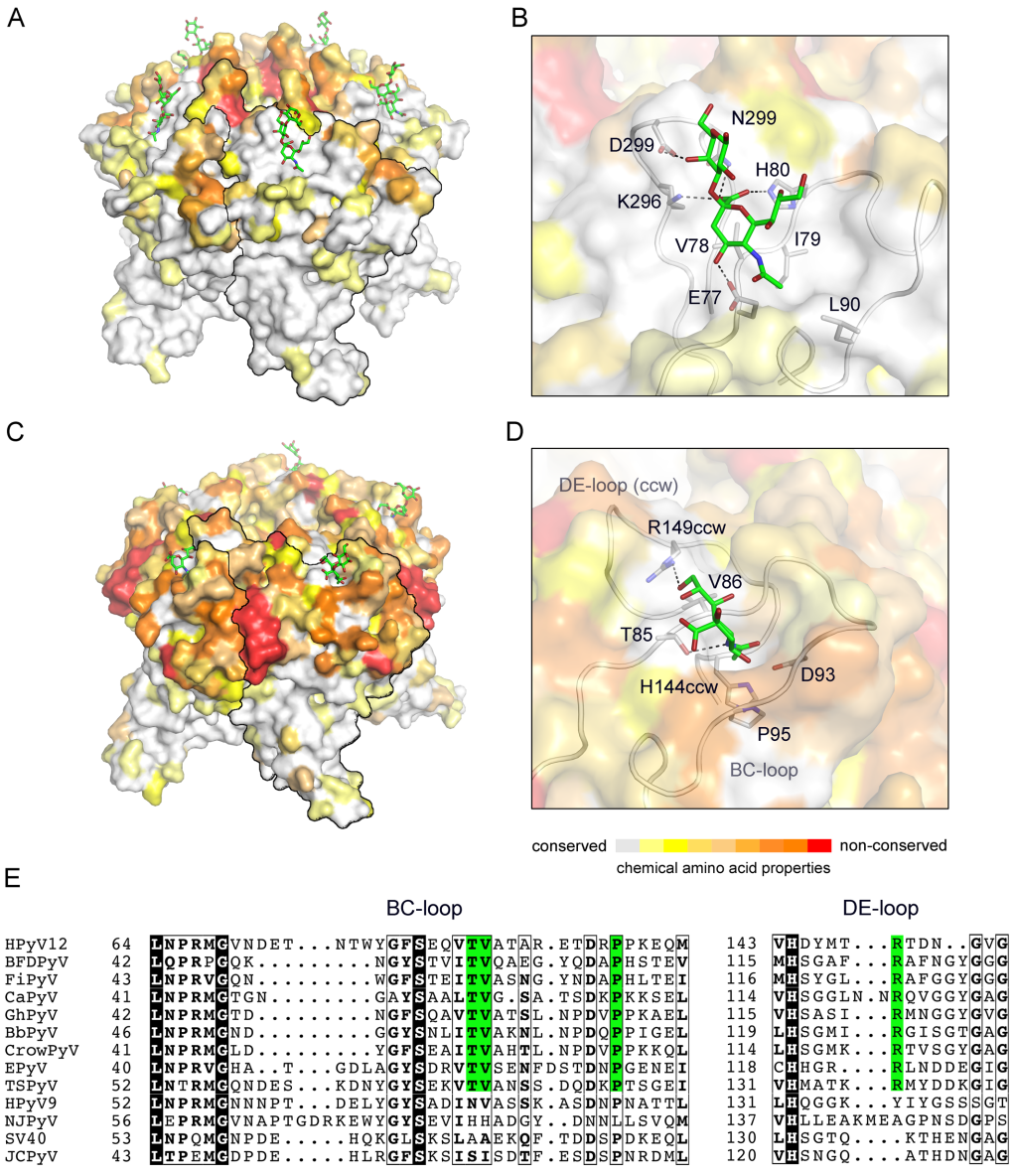


Figure 18: Features of the NJPyV and HPyV12 VP1 Neu5Ac binding sites are conserved in closely related polyomaviruses from different species. (A), (B) VP1 differences between NJPyV, ChPyV-Azzi (GFR692336), ChPyV-Bob (FR692334), ChPyV-Tanu (FR692335), VmPyV1 (NC_019844), and PrufPyV1 (JX159984) are mapped onto the NJPyV VP1-3'SL complex structure and shown in overall and close-up views. (C), (D) Based on the HPyV12 VP1-3'SLN complex structure, VP1 differences between HPyV12 and two representative Avipolyomaviruses FiPyV (NC_007923) and BFDPyV (AY672646) are shown. Amino acid differences are coloured according to their chemical conservations in panel A-D. Amino acid sequence alignments were done with Clustal Omega (Sievers et al., 2011) and JalView (Waterhouse et al., 2009) to assign values for the conservation of their

chemical properties from 10 (conserved; grey) to 0 (non-conserved; red). Values of 2 to 0 are coloured in red. In the close-up views conserved residues of the glycan binding site are highlighted in sticks representations. (E) Structure-based sequence alignment of BC2- and DE-loop regions from HPyV12, Avipolyomaviruses, and representative Orthopolyomaviruses. Canary Polyomavirus (CaPyV; NC_017085), Goose hemorrhagic Polyomavirus (GhPyV; HQ681903), Butcherbird Polyomavirus (BbPyV; NC_023008), Crow Polyomavirus (CrowPyV; NC_007922), and Equine Polyomavirus (EPyV; NC_017982). Conserved key residues of the Neu5Ac binding site are highlighted in green. NCBI gene bank numbers are given in parenthesis.

Fragment-based Screening with JCPyV VP1

Fragment-based screening by ^{19}F and *transverse relaxation time* (T2)-filtered (Carr-Purcell-Meilboom-Gill, CPMG) NMR spectroscopy experiments were performed in collaboration with Christoph Rademacher (Max Planck Institute of Colloids and Interfaces, Potsdam) to identify small compounds that bind to JCPyV VP1 pentamers. Both NMR methods are 1D ligand-based approaches that are often used in FBDD. For NMR studies ^{19}F has excellent properties that are comparable to those of ^{1}H (Chen et al., 2013). The ^{19}F nucleus has a 1/2 spin that exists in 100% natural abundance, but also the intrinsic sensitivity of the ^{19}F nucleus and its chemical shift to changes in the local chemical environment are useful for FBDD (Jordan et al., 2012; Chen et al., 2013). A library consisting of 300 low molecular weight fluorinated compounds was used for JCPyV VP1. All compounds in the library contain one mono fluoro group or a trifluorometyl group. 1D NMR spectra were analyzed for changes in peak intensity and chemical shift. In the case of a binding event, a shift of the characteristic peak for the compound occurs in the spectrum with the protein compared to the spectrum of the sample without protein (Figure 19). The screening resulted in 40 double positive hits obtained by both NMR spectroscopy methods.

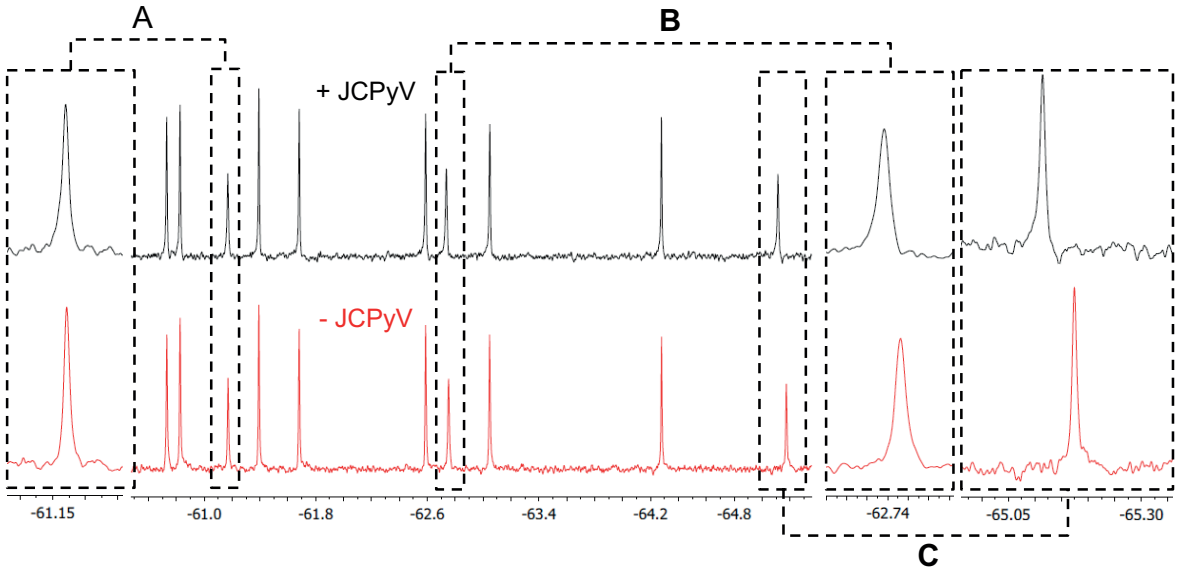


Figure 19: Fragment-based screening by NMR. ^{19}F NMR spectroscopy was performed to identify small compounds that bind to JCPyV VP1 pentamers. NMR spectroscopy measurements were carried out by Jonas Aretz and Christoph Rademacher, Max Planck Institute of Colloids and Interfaces, Potsdam. Typical results are shown exemplarily for three different compounds A, B, and C. Binding was observed for compound B and C (bold letters). Two identical sets of spectra were acquired. In the case of binding, a peak shift occurs in the spectrum with JCPyV VP1 (+ JCPyV) compared to the spectrum of the sample without protein (- JCPyV). The figure shown here was adapted from a figure prepared by Jonas Aretz (unpublished data).

So far 10 compounds out of the 40 positive hits have been further analyzed via crystal soaking experiments with JCPyV Mad-1 VP1 crystals (Figure 20).

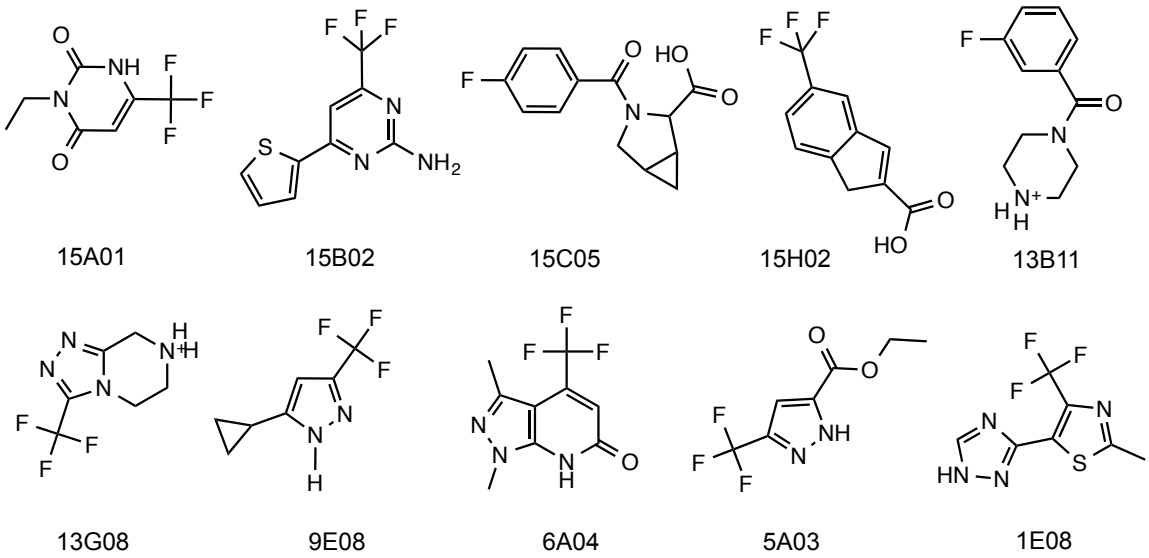


Figure 20: Fragments used for soaking experiments. Displayed are 10 out of 40 double hits identified by ^{19}F and T2-filtered NMR spectroscopy. The fragment screening was performed and analyzed by Christoph Rademacher and Jonas Aretz, Max Planck Institute of Colloids and Interfaces in Potsdam. The figure was prepared using ChemDraw 14.0.0.118 (PerkinElmer, Inc.).

Crystallization and Crystal Soaking Experiments

Crystals were grown using the sitting drop vapor diffusion method, a 4.5 mg/ml JCPyV VP1 solution in 20 mM HEPES pH 7.5, 150 mM NaCl, a reservoir solution containing 0.1 M HEPES pH 7.5, 0.2 M KSCN, 12% (w/v) PEG 3,350, and a microseed solution obtained from previous grown JCPyV VP1 crystals (Neu et al., 2010). The hydrophobicity and thus poor solubility of compounds in aqueous solutions is often a problem during early steps of FBDD. Thus, JCPyV VP1 crystals were tested for their resistance against two common organic solvents, dimethyl sulfoxide (DMSO) and methanol, prior to the crystal soaking experiments. Crystals were incubated overnight in the reservoir solution supplemented with different concentrations of DMSO or methanol, respectively. They were harvested similar to native crystals and tested for diffraction. JCPyV VP1 crystals tolerate up to 30% (v/v) DMSO or 20% (v/v) methanol. At higher concentrations of DMSO or methanol, crystals break to pieces or diffract only to a low resolution (≥ 3.5 Å). Stock solutions of the 10 compounds (Figure 20) were prepared in DMSO and a maximum concentration of 25% DMSO was used for the soaking experiment (Table 4). However, differences in solubility of the fragments and the stability of the crystal in the presence of the respective fragments had influence on the final soaking conditions. Compounds 15B02 and 6A04 are not well soluble in aqueous solutions. At concentrations of

50 mM and above, compound solutions with 15C05, 15H02, and 13G08 destroyed the crystals and the compound concentration was reduced to 25 mM.

Table 4: Soaking experiment with 10 hits from the fragment-based screening. JCPyV VP1 Mad-1 crystals were grown for two days against a reservoir containing 0.1 M HEPES pH 7.5, 0.2 M KSCN, 12% (w/v) PEG 3,350. Crystals were transferred into drops of the harvesting solution supplemented with compound and DMSO as indicated. 15B02 is not well soluble and the soaking solution with 100 mM compound was a saturated solution. Soaking experiments were performed stepwise used in some cases and values for the used solution are listed then with semicolons. Prior to flash freezing in liquid nitrogen, crystals were incubated for 2 sec in a soaking solution supplemented with 30% (v/v) glycerol.

Compound ID	Compound [mM]	DMSO [%]	Duration [h]
15A01	25; 100	25; 25	1; 1
	25	25	approx. 16
15B02	10	25	2
	25; 100	25; 25	1; 1
15C05	25	6.25	approx. 16
15H02	25	25	2
13B11	100	25	approx. 16
13G08	25	6.25	approx. 16
9E08	25; 100	25; 25	1; 1
	100	25	approx. 16
	25	6.25	approx. 16
6A04	20	25	2
5A03	25; 100	25; 25	1; 1
	25	25	approx. 16
1E08	25; 100	25; 25	1; 1
	25	25	approx. 16

Diffraction data sets were collected at beamline X06DA of the SLS and were processed with XDS (Kabsch, 2010). Structures were solved by MR using Phaser MR (McCoy et al., 2007) and the native JCPyV VP1 structure (pdb: 3NXG) without solvent molecules as a model. Rigid body and simulated annealing refinement was carried out with Phenix (Adams et al., 2010). Alternating rounds of model building in Coot (Emsley et al., 2010) and refinement with Refmac5 (Murshudov et al., 1997) were carried out. JCPyV crystals grow in space group C2 with one monomer in the asymmetric unit. Contacts between neighboring VP1 pentamers in the crystal do not block any side of all the monomers simultaneously and consequently, binding sites are likely accessible in the context of the VP1 pentamer. Additional electron density in the simulated annealing electron difference map was observed for all three crystals that were soaked with the fragment 9E08 (Figure 21). The fragment was built using model coordinates and restraints created by PRODRG2 Server (Schuttelkopf and van Aalten, 2004) with subsequent refinement via Refmac5. The compound 9E08 could be built into the electron

density of all three crystals, but only the data set collected from the crystal soaked with 100 mM was used for further refinement (Table 5).

Table 5: Statistics of the data collection and ongoing refinement. Values in parentheses are for the highest resolution shell. The occupancy of the ligand 9E08 is 0.8.

	9E08 (100 mM; 16 h)
Beamline	SLS X06DA
Data collection	
Space group	C2
a, b, c [Å]	49.2, 95.5, 128.5
β [°]	110.2
Resolution [Å]	50-1.85 (1.90-1.85)
Unique reflections	142,639 (10,321)
Total reflections	963,586 (63,607)
R_{meas} [%]	11.3 (110.6)
I/σI	13.26 (1.75)
$CC_{1/2}$ [%]	99.8 (75.2)
Completeness [%]	98.9 (97.3)
Wilson B-factors [Å^2]	30.1
Refinement	
R_{work} / R_{free} [%]	20.4 / 21.3
No. of atoms	
protein	9,879
ligand	60
water	593
B-factor [Å^2]	
protein	25.3
ligand	38.7
water	29.9
R.m.s.d.	
Bond length [Å]	0.007
Bond angles [°]	1.238

The Binding Site for 9E08 is Located Inside the Hydrophobic Cavity of the VP1 Pentamer

The compound 9E08 binds inside the VP1 pentamer in a location next to β-strands E, F, and the ccw G1 of the core CHEFG1-sheet (Figure 21). All five binding sites of the VP1 pentamer in the asymmetric unit are occupied. Non-polar residues F144 and F213 enclose the binding site together with hydrophobic parts of side chains from residues H142, H227ccw, and T229ccw. These residues and the β-strand backbone form a shallow, relatively hydrophobic pocket, but charged interactions or strong hydrogen bonds between 9E08 and VP1 were not observed. The electron density for 9E08 can be explained by ligand occupancies from 0.8 to 1, which result in temperature factor (B-factor) values of 38.7 to 47.4 Å^2, respectively. Thus, B-factors are only little elevated for 9E08 compared to VP1 residues forming the binding pocket. B-factors reflect the relative vibrational motion of atoms and a bound ligand is expected to have similar values as interacting and neighboring protein residues. Side chain atoms possess here B-factors of about 22 Å^2 and main chain atoms of about 18 to 20 Å^2. However, residues of the conserved β-sheet core have in general relatively low B-factors as known from other polyomavirus VP1 structures and thus, B-factor values for the compound are within the expected range.

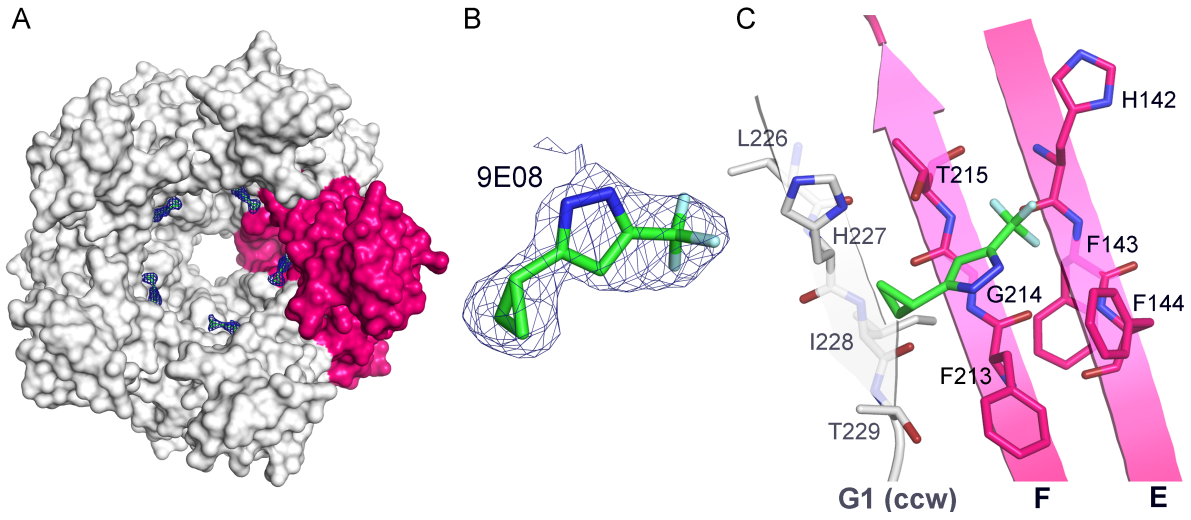


Figure 21: The simulated annealing F_{obs}-F_{calc} omit electron density map shows features of the compound 9E08. (A) View from the bottom of the JCPyV pentamer along the five-fold axis. The VP1 pentamer is shown in surface representation with one monomer highlighted in pink. The simulated annealed omit electron density map is contoured at a σ level of 2.5 with a radius of 2 Å around the ligand 9E08. (C) β-sheets E, F, and the ccw G1 form the binding site for 9E08. The compound 9E08 is shown in stick representation with carbons in green, nitrogens in blue, and fluor in light blue.

The electron density observed for 9E08 is absent in simulated annealing F_{obs}-F_{calc} electron density maps of the native JCPyV VP1 and other crystals soaked with different ligands.

However, it seems that a DMSO molecule partially occupies the binding site in other soaked crystals due to its high concentration in the soaking solution.

In order to further examine the VP1-9E08 interaction in the scenario of the assembled JCPyV virion, we superimposed the complex structure with the MPyV VP1-VP2 complex using Cα atoms and the SSM superimposition tool in Coot (Krissinel and Henrick, 2004) (Figure 22). Molecular details of the JCPyV VP1 and VP2 interactions are not known, but the interaction of VP1 with minor capsid proteins VP2 and VP3 and their exposure in the ER is critical for JCPyV entry (Nelson et al., 2015). Consistently, JCPyV mutants lacking VP2 or VP3 are non-infectious (Gasparovic et al., 2006). The hydrophobic core of VP1 is structurally highly conserved and both structures align well in the region of the 9E08 binding site (Figure 22 C). It becomes clear that the binding site for 9E08 may partially overlap with the binding site for VP2 inside the pentamer cavity (Figure 5 and Figure 22). Since structural data for the interaction of JCPyV VP1-VP2 is not available, a sequence alignment was carried out for VP2 (Figure 22 D). The VP2 amino acid sequence is only partially conserved in the 9E08 binding site. Residues of MPyV VP2 that are involved in backbone-mediated interactions with VP1 are not present in JCPyV VP2 (Figure 5 C). However, the electron density for VP2 in the MPyV VP1-VP2 complex structure did not allow to model side chains in that region and thus it is not possible to make a clear statement what critical role specific side chains play in the interaction between VP1 and VP2. The sequence of the C-terminal part of the VP2 fragment, which was visible and involved in interaction with VP1 in the complex structure, is in contrast well conserved between MPyV and JCPyV (Figure 5 D). Therefore, observations allow only vague speculations about the role of the 9E08 binding site for the JCPyV VP1-VP2 interaction so far.

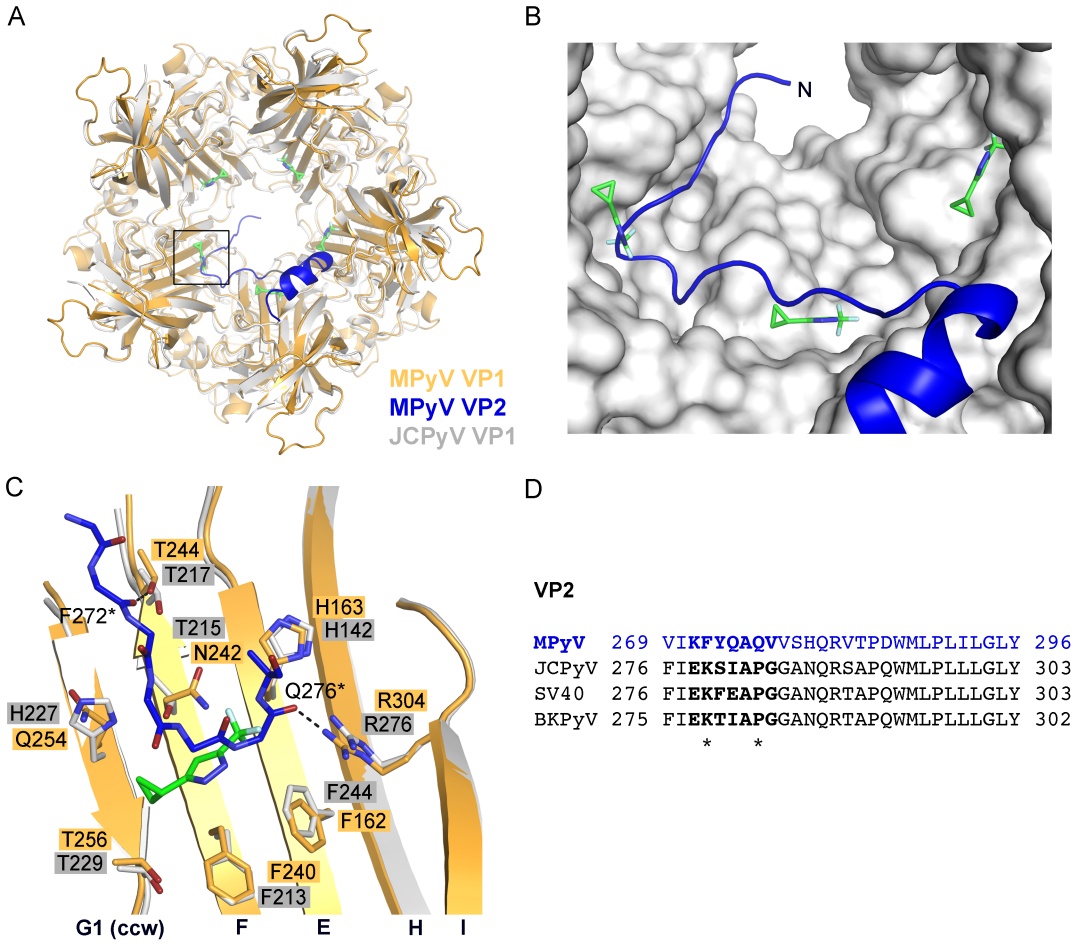


Figure 22: The binding site of 9E08 could partially overlap with the VP2 binding site. (A) Superimposition of the JCPyV VP1-9E08 structure with the MPyV VP1-VP2 structure (pdb: 1CN3) using Cα atoms and the SSM algorithm in Coot. Cartoon representations are colored grey for JCPyV and orange for MPyV, respectively. The part of MPyV VP2 that was visible in the MPyV VP1-VP2 complex structure is shown in blue (see also Figure 5). The compound 9E08 is shown in stick representation with carbons in green, nitrogens in blue and fluor in light blue. (B) The 9E08 binding site on JCPyV VP1 is located in the region of VP1 that contacts the N-terminal part of VP2 in the MPyV VP1-VP2 complex structure. JCPyV is shown in surface representation, MPyV VP2 in cartoon (blue), and 9E08 in sticks. (C) Close-up view of the 9E08 binding site in comparison with the MPyV VP1-VP2 structure. The view is equivalent to Figure 21 C. VP1 side chains of in the binding site and MPyV VP2 residues 271-277 are shown in stick representations. Proteins are colored in orange (MPyV VP1), blue (MPyV VP2), and grey (JCPyV VP1). (D) Sequence alignment of VP2 from the MPyV strain PTA (used for the structural studies), JCPyV (NC_001699), SV40 (NC_001669), and BKPyV (NC_001538). Only the sequence for the VP2 fragment that was visible in the X-ray structure is shown. The part of VP2 shown in panel C is highlighted in bold for MPyV. The positions of the two VP2 amino acids interacting with VP1 via backbone-mediated hydrogen bonds in the case of MPyV are marked with an asterisk in panel C and D.

4. Outlook

The work presented in this thesis has established that JCPyV engages LSTc as a functional receptor motif and that gangliosides do not play a role during JCPyV infection. However, questions where LSTc is linked to and what role its cellular context plays for early step during JCPyV infections remain to be answered in the future. Commonly used techniques to identify virus receptors such as broad genome-wide siRNA screening or virus overlay protein assays would be rational approaches (Cao et al., 1998; Tayyari et al., 2011; Konopka-Anstadt et al., 2014). Pull-down experiments with isolated membrane fractions could be conducted in order to fish out the protein backbone of LSTc glycan. Such experiments using immobilized VP1 pentamers would be technically straightforward, but the low binding affinity of the VP1-glycan interaction could hamper the experimental outcome. Alternatively, similar experiments could be performed with virions or with an antibody against LSTc. Possible receptor candidates could then be identified by mass spectrometry. In addition, a genome-wide siRNA screening could not only help to identify the proteinous backbone for LSTc, but may also shed light on additional viral receptors or cellular factors that are critical for JCPyV infection. A differential approach in kidney and glial cells or even neurons could substantial uncover viral mechanism and cell-specific factors leading to the diseases PML, GCN, and JCVE.

4.1 Role of Serotonin Type 2 Receptors during JCPyV Infection

Further insights into the interaction between JCPyV and 5-HT2Rs could help to answer the question why JCPyV uses a clathrin-dependent pathway, whereas closely related polyomaviruses such as SV40 and BKPyV take a lipid-dependent entry route. JCPyV and 5-HT2AR colocalize 5 min post-infection, suggesting an involvement of 5-HT2ARs in the early entry process (Elphick et al., 2004). 5-HT2Rs are not important for initial viral attachment, but transient direct interactions on the cell surface cannot be completely excluded (Assetta et al., 2013). However, JCPyV infections occur also in the absence of 5-HT2AR in human brain microvascular endothelial cells, suggesting that 5-HTRs function in a redundant manner (Chapagain et al., 2007). In order to investigate if 5-HT2Rs are critical for intracellular signaling events during JCPyV infection, mutations studies, for example at the C-terminus of 5-HT2Rs located in the cytosol, might clarify the role of 5-HT2Rs. Conserved amino acid regions or residues at the C-terminus could be deleted or mutated and effects on JCPyV infectious entry could then be monitored by infection assays and co-localization experiments by fluorescence microscopy.

Available structural-functional data of JCPyV and related polyomaviruses offer a unique opportunity to explore and identify parameters that define the importance of 5-HT2Rs receptor

engagement, and viral entry pathways. Established hypotheses could be tested with engineering viruses that engage different receptors and switch their host and cell specificities and perhaps their entry pathways. Since GM1 and GD1b glycan binding was observed for JCPyV VP1, a receptor retargeting approach towards engagement of the GM1 ganglioside has been tested via site-directed mutagenesis at JCPyV VP1 residue N123 (data not shown). The asparagine at this position in JCPyV VP1, which is a glycine in SV40 and BKPyV VP1, is a key residue of the binding site that makes contacts with both arms of the L-shaped LSTc. Evaluation of mutated VP1 pentamers by crystal structure analysis revealed abolished binding to LSTc in the L-shape conformation and conserved engagement of the GM1 glycan, but supplementation of cells with GM1 ganglioside did not alter cell binding or infection of these mutants. Thus, it cannot be ruled out that there are, in addition to the increased binding affinity for LSTc, other evolutionary constraints for utilizing LSTc, so that this virus-glycan interaction is favored over engagement of GM1 or b-series gangliosides. A reverse approach via site-directed mutagenesis in SV40 and BKPyV VP1 may help to determine if enhanced LSTc binding is sufficient for targeting the viruses to the clathrin-dependent entry pathway and if this binding event is a prerequisite for involvement of 5-HT2Rs in the early steps of infection.

4.2 Development of Antiviral Strategies for JC Polyomavirus

The development of antiviral compounds for JCPyV could focus in the future on several steps during the virus lifecycle. Proceeding with strategies to block the interaction between JCPyV VP1 and LSTc during attachment, competition experiments with LSTc or GM1 glycans by STD NMR spectroscopy could be carried out to further evaluate the 40 hits from the fragment-based screening approach. Although the GM1 ganglioside is not a functional receptor for JCPyV, the non-reducing end of the GM1 glycan binds in a unique groove on the VP1 surface in the JCPyV VP1 complex structure, highlighting a potential binding site for small compounds on the top of VP1. In further experiments, such fragments engaged in this binding site could be chemically linked with positive hits engaged in the Neu5Ac binding pocket. The extension of fragments by adding chemical groups that are pointing into the direction of the core Neu5Ac binding site might also be possible in order to increase the number of specific interactions and hence the binding affinities for compounds. Furthermore, the introduction of a multivalent binding mode may drastically enhance the binding affinities of promising compounds (Rademacher et al., 2011; Spjut et al., 2011). Potent antiviral compounds inhibiting the binding of JCPyV to Neu5Ac on the host cell surface may provide then a novel route for the development of similar antiviral therapies for other viruses that initiate infection through the recognition of sialylated glycans.

However, the fact that the role of the PML-associated VP1 mutations in JCPyV pathogenicity remains still unclear weakens this approach to develop antiviral strategies. It is plausible that the loss of sialic acid binding of PML-associated viruses may be critical for the increased spread to the brain and the subsequent infection of glial cells in a sialic acid-independent manner. Moreover, it is possible that there is an unknown host cell factor, for example a non-sialic acid receptor on glial cells in the CNS, which is important for infection by JCPyV with PML-associated VP1 mutations. This receptor may be not abundantly expressed on cell lines that have been used in available model systems. Consequently, the development of inhibitors against the wild type VP1 variants might be a dead end and other strategies are needed. It was shown in this thesis that the 5-fold pore of the VP1 pentamer is an important structural feature and likely an important interaction platform during the infection. Studies with VP1 mutants interfering with the VP1-VP2 interaction could help to further establish the role of VP2 for infection, possible pointing out alternative strategies for the development of antiviral compounds. Starting from the compound 9E08, chemically related compounds with better solubility could be tested in crystal soaking experiments to validate the binding inside the pore and the addition of chemical groups interacting with VP1 could enhance the binding affinity of the 9E08 fragment. With better binding compounds in hand, cell-based assays are the next step to test their antiviral potentials in cell culture. In contrast to attachment inhibitors, it is necessary that compounds reach the nucleus during virus assembly, facilitating the development of antiviral compounds. Nevertheless, such an assembly inhibitor would be equally effective against other virus variants, since the interaction between VP1 and VP2 is rather conserved between JCPyV genotypes and viral subgroups, but a high-resolution structural analysis is needed to define molecular details of the VP1-VP2 interaction for a rational structure-based approach. Similarly, compounds binding at the C-terminus VP1 may interfere with the correct assembly of VP1 pentamers into the capsid. However, molecular details of VP1 pentamer-VP1 pentamer interactions in the assembled capsid are only available for SV40 and MPyV so far (Liddington et al., 1991; Stehle et al., 1994; Yan et al., 1996). In addition, it remains unclear how the interface between pentamers is involved in virus-host interactions during the infection. The observation that the disease JCPyV GCN is associated with mutations and deletions at the C-terminus of VP1 suggests that canyons between VP1 pentamers may be important during the infection and for the pathogenicity (Dang and Koralnik, 2006; Agnihotri et al., 2014). Hence, further studies are definitely needed to explore the importance of the canyons for JCPyV receptor binding and its applicability as a target for antiviral strategies.

In general, X-ray crystallography has been approved to be very useful during FBDD approaches although several experimental factors could prevent protein-ligand complex formation hampering the experimental outcome (Davies and Tickle, 2012). Crystallization

cocktails differ often significantly from biologically relevant environments, crystal contacts may block the binding sites, or the crystal lattice may not allow conformational change that are necessary for ligand binding during crystal soaking experiments. In some cases co-crystallization experiments can overcome these problems but may require extensive crystallization screening for each ligand due to induced conformational changes upon ligand binding that disrupt crystal contacts or addition of organic solvents such as DMSO, which is necessary to solubilize fragments, may alter the crystallization system.

JCPyV VP1 crystallizes under a relative physiological condition at pH 7.5 (0.1 M HEPES pH 7.5, 0.2 M KSCN, 12% (v/v) PEG 3,350). Chances that binding sites for fragments are accessible in the preformed crystal during soaking experiments are relatively high, because neighboring VP1 pentamers in the crystal do not contact each VP1 monomers in an identical manner. Nevertheless, only one of the ten hits from the screening by NMR was observed in the electron density map, a validation hit rate that is not very high but still comparable with observations from other FBDD projects (Wielens et al., 2013). Different cascades of biophysical methods have been used during FBDD projects with various success of hit validation and hit overlap (Wielens et al., 2013). FBDD strategies include in addition to NMR spectroscopy or fluorescence-based thermal shift assays often a technique to estimate binding affinities such as ITC or SPR (Ciulli and Abell, 2007; Giannetti, 2011). Thus, it is worth to consider implementing such a technique into the FBDD workflow after initial screening by NMR. In contrast, the initial screening step is now more often performed directly by X-ray crystallography (Murray et al., 2007; Davies and Tickle, 2012). The potential of X-ray crystallography for early steps of FBDD was historically underappreciated due to relatively low throughput and the subjective time consuming interpretation of electron density maps, but it does usually not suffer from the problem of false positive such as other screening techniques (Chilingaryan et al., 2012). If hit identification is difficult in poorly defined electron densities due to alternative conformations or low fragment occupancies for example bromide-containing fragments and anomalous scattering could be used to overcome such problems (Tiefenbrunn et al., 2014). In conclusion, FBDD in combination with or carried out by high-throughput X-ray crystallography is a promising strategy for future development of antiviral compounds interfering with several steps during the JCPyV lifecycle in order to develop effective treatments against PML.

4.3 Identification of Glycan Receptor Candidates for Recently Discovered Human Polyomaviruses

This thesis provides structural and functional insights into receptor recognition of five recently discovered polyomaviruses widely circulating within the human population. Clear evidence for

the importance of sialylated glycans during viral infection has been generated for TSPyV, NJPyV, and HPyV12, whereas HPyV6 and HPyV7 engage likely non-sialylated receptors. Initially, polyomaviruses were thought to bind sialic acids in the same general area of VP1. This common binding site is plastic and allows the engagement of the same ligand in different orientations (Ströh and Stehle, 2014). In contrast, a novel sialic acid binding site within the BC2-loop of VP1 has been identified for TSPyV. Subsequent structural studies with HPyV12 VP1 pentamers identified a very similar glycan-binding site. Together with phylogenetic analysis, these structural data suggest that the BC2-loop binding site may also be employed for receptor binding by other polyomaviruses.

In order to identify potentially more complex and specific sialic acid receptors glycan array screening is currently ongoing with TSPyV, HPyV12, and NJPyV VP1 pentamers. Previous glycan array screening approaches with either N-terminal His-tagged TSPyV VP1 pentamers or directly Alexa Fluor 647-labeled VP1 pentamers did not result in significant binding signals on microarrays that were composed of 128 sequence-defined lipid-linked oligosaccharide probes, including 122 sialylated and six neutral probes. Results with Alexa Fluor 647-labelled TSPyV pseudoviruses were similar (Ten Feizi, unpublished data). It is known that recombinant VP1 pentamers form disulfide bond-mediated dimers (“tail-to-tail”) and higher multimers in solution due to the conserved and exposed cysteine in the CD-loop. Disulfide bond formation might result in a scenario where the N-terminal His-tag is less accessible, hampering common glycan array detection systems using an anti-His tag antibody. For the ongoing glycan array screening, cysteine residues were mutated to serine residues in TSPyV (C107), HPyV12 (C119), and NJPyV (C113) VP1 pentamers. The respective mutations are distant from the surface loops of VP1 and should not alter glycan binding.

Based on the structural analyses, it is also tempting to speculate about a broader glycan receptor specificity of TSPyV and HPyV12. The shallow BC2-loop binding groove allows interactions with terminal α2,3-, α2,6-, and likely α2,8-linked sialic acids without steric hindrance. These findings might explain why specific binding events above the background binding level could not be identified on the glycan array. In contrast, cell binding and infection studies have revealed that neither N- nor O-linked glycans are involved in the TSPyV infection, suggesting instead that a specific lipid-based virus receptor with a more complex binding epitope than terminal sialic acids is present on infected cells. The recognition of non-sialylated glycan epitopes in the “old” binding site might be realized due to the plasticity of the VP1 surface loops. A similar switch in receptor specificity between different glycans classes has been described for rotaviruses. The porcine rotavirus CRW-8 binds the GM3 glycan in a conserved overall location on the capsid protein VP8*, whereas the human rotavirus strain HAL1166 engages a neutral tetrasaccharide corresponding to the terminal structure of A-type HBGA in the same area (Yu et al., 2011b; Hu et al., 2012). However, one could only speculate

so far about such a switching scenario in the case of TSPyV and HPyV12 without experimental data from glycan array screening or molecular details from X-ray crystal structures in hand.

4.4 Cell-based Model Systems and Functional Infection Studies

Model systems are clearly required for functional studies of receptor specificity and cell entry mechanisms. However, the pseudovirus system provides already a sensitive method for studying determinants and mechanisms of the infectious entry of newly discovered polyomaviruses, mutant viruses, and virus variants or strains that do not propagate using traditional cell culture methods. Structure-based mutagenesis of pseudoviruses and cell supplementation assays would be the next steps to study the importance of the virus-glycan interaction for NJPyV and HPyV12 infection. A reverse approach, starting from exploring the cell tropism of the virus in order to identify receptor candidates is possible but challenging because systematic data sets about cell-type specific glycans are not available or rather incomplete. Nevertheless, the development of model cell culture systems allowing the propagation of virus in tissue culture is a requirement to study the molecular biology of newly discovered polyomaviruses besides cell attachment and endocytosis. Similar systems are available of JCPyV, including a primary human fetal astrocyte cell culture system (Radhakrishnan et al., 2003), the glial cell line SVGA (Major et al., 1985; Gee et al., 2003), and more recently described oligodendrocyte precursor cells (Darbinyan et al., 2013).

The link to a human disease is not clear for all human polyomaviruses although a persistent infection with these viruses is very common within the healthy human population. There is now clear evidence that the infection with HPyV7 can lead to pruritic, brown plaques on the trunk and extremities of immunosuppressed individuals such as transplant patients (Ho et al., 2014). More recently, and almost five years after its discovery, HPyV7 has also been linked to thymic epithelial tumors (Rennspiess et al., 2015). Thus, an involvement in human diseases and tumors is likely for HPyV6 and has to be considered for other recently discovered human polyomaviruses.

The first confirmed human oncogenic member of the polyomavirus family is MCPyV. Interestingly, mechanisms and characteristics of cellular infections with MCPyV differ in several aspects from human pathogenic polyomviruses such as BKPyV and JCPyV. The genome of MCPyV is integrated into the genome of MCC cells in a large percentage of cases, whereas clonal integration is not observed for other human polyomaviruses. MCPyV engage sialic acids and additionally GAGs during the early stages of the infection. While cellular receptors for HPyV6 and HPyV7 are unknown it remains to be investigated if other skin-tropic polyomaviruses employ similar strategies and require for GAGs attachment. Glycan array screening with pseudoviruses instead of VP1 pentamers samples also the surface between

pentamers in the assembled virion for possible glycan interactions. However, HPyV6 and HPyV7 may also engage proteinous receptors. Screening for infectivity in different cell lines would be required to identify host cells that express potential receptors. Subsequent genome-wide siRNA screening approaches or virus overlay protein assays could be used to identify proteinous receptors for the more recently discovered polyomaviruses (Cao et al., 1998; Tayyari et al., 2011; Konopka-Anstadt et al., 2014).

4.5 Host Tropism, Viral Adaptation, and Pathogenicity

The phylogenetic analysis of the polyomavirus family highlights occurring pairwise relationships of human and monkey viruses, including species such as HPyV9 and LPyV, NJPyV and ChPyV, and TSPyV and OraPyV-Bo (Figure 1). Critical differences in glycan binding specificities of LPyV and HPyV9 have enhanced our understanding of principles that underlie pathogen selectivity based on sialic acid modifications (Khan et al., 2014), but the influence on host specificity remains speculative or unclear in most cases. Further investigations are also needed to understand the molecular basis for switching events in host tropism and viral adaptation and especially pathogenicity. Typically, the polyomavirus infection is latent but asymptomatic in immunocompetent hosts whereas infections can result in severe outcomes during immunosuppression or in an altered pathogenicity in other hosts. For example, SV40 infections are mostly asymptomatic in healthy monkeys, but the virus can lead to tumors in small rodents (Vilchez and Butel, 2004). Similarly, JCPyV induces formation of tumor formation in animal model systems, whereas its reactivation in humans during immunosuppression results in brain diseases such as PML with different pathogenic characteristics (Walker et al., 1973; London et al., 1978; Ohsumi et al., 1986). The resulting lack of a good animal model hampers *in vivo* studies with JCPyV. However, although a humanized mouse model does not recapitulate the disease PML, it may be still useful to study for example the immune response during the infection, viral adaption, and tissue tropism (Tan et al., 2013).

In conclusion, with relatively high sequence homology and conserved architecture across the family, polyomaviruses form a useful platform for analyzing determinants of antigenicity, cell entry, cell tropism and host range, as well as other factors that contribute to viral pathogenesis. Combined data sets including structural data about specific (glycan) receptor engagement and receptor-switching potential of viruses provide hereby a powerful platform for exploring new ways forward in the rational development of antiviral strategies and a more reliable prediction of virus-host interactions, viral evolution, and cross-species transmission where no structural data is at hand. The detailed understanding of underlying principles of glycan-virus and its

impact on cell tropism is also important for new approaches in the design of therapeutic gene vectors.

Overall, an increased understanding how polyomaviruses and in particular JCPyV evolve to evade the host immune system and antibody-mediated neutralization could open the door to clinical developments of vaccines against JCPyV and PML and other diseases caused by human polyomaviruses in the future.

5. References

Abend, J.R., Joseph, A.E., Das, D., Campbell-Cecen, D.B., and Imperiale, M.J. (2009). A truncated T antigen expressed from an alternatively spliced BK virus early mRNA. J Gen Virol *90*, 1238-1245.

Adams, P.D., Afonine, P.V., Bunkoczi, G.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D-Biological Crystallography *66*, 213-221.

Agnihotri, S.P., Dang, X., Carter, J.L., Fife, T.D., Bord, E., Batson, S., and Koralnik, I.J. (2014). JCV GCN in a natalizumab-treated MS patient is associated with mutations of the VP1 capsid gene. Neurology *83*, 727-732.

Ahmed, W., Wan, C., Goonetilleke, A., and Gardner, T. (2010). Evaluating sewage-associated JCV and BKV polyomaviruses for sourcing human fecal pollution in a coastal river in Southeast Queensland, Australia. J Environ Qual *39*, 1743-1750.

Angata, T., and Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. Chem Rev *102*, 439-469.

Asselin, C., Gelinas, C., Branton, P.E., and Bastin, M. (1984). Polyoma middle T antigen requires cooperation from another gene to express the malignant phenotype in vivo. Mol Cell Biol *4*, 755-760.

Assetta, B., Maginnis, M.S., Gracia Ahufinger, I., Haley, S.A., Gee, G.V., Nelson, C.D., O'Hara, B.A., Allen Ramdial, S.A., and Atwood, W.J. (2013). 5-HT2 receptors facilitate JC polyomavirus entry. J Virol *87*, 13490-13498.

Atwood, W.J., and Norkin, L.C. (1989). Class-I major histocompatibility proteins as cell-surface receptors for Simian Virus-40. J Virol *63*, 4474-4477.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Barouch, D.H., and Harrison, S.C. (1994). Interactions among the major and minor coat proteins of polyomavirus. J Virol *68*, 3982-3989.

Bartholomeusz, A., and Locarnini, S. (2006). Hepatitis B virus mutations associated with antiviral therapy. J Med Virol *78 Suppl 1*, S52-55.

Bauer, P.H., Bronson, R.T., Fung, S.C., Freund, R., Stehle, T., Harrison, S.C., and Benjamin, T.L. (1995). Genetic and structural analysis of a virulence determinant in polyomavirus VP1. J Virol *69*, 7925-7931.

Bauer, P.H., Cui, C.Q., Stehle, T., Harrison, S.C., DeCaprio, J.A., and Benjamin, T.L. (1999). Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse. J Virol *73*, 5826-5832.

Bauman, Y., Nachmani, D., Vitenshtein, A., Tsukerman, P., Drayman, N., Stern-Ginossar, N., Lankry, D., Gruda, R., and Mandelboim, O. (2011). An identical miRNA of the human JC and BK polyoma viruses targets the stress-induced ligand ULBP3 to escape immune elimination. Cell Host Microbe *9*, 93-102.

Bellizzi, A., Anzivino, E., Rodio, D.M., Palamara, A.T., Nencioni, L., and Pietropaolo, V. (2013). New insights on human polyomavirus JC and pathogenesis of progressive multifocal leukoencephalopathy. Clin Dev Immunol *2013*, 839719.

Bennett, S.M., Broekema, N.M., and Imperiale, M.J. (2012). BK polyomavirus: emerging pathogen. Microbes Infect *14*, 672-683.

Bennett, S.M., Jiang, M., and Imperiale, M.J. (2013). Role of cell-type-specific endoplasmic reticulum-associated degradation in polyomavirus trafficking. J Virol *87*, 8843-8852.

Benoist, C., and Chambon, P. (1981). In vivo sequence requirements of the SV40 early promotor region. Nature *290*, 304-310.

Blake, S.J., Bokhari, F.F., and McMillan, N.A. (2012). RNA interference for viral infections. Curr Drug Targets *13*, 1411-1420.

Blixt, O., Head, S., Mondala, T.*, et al.* (2004). Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc Natl Acad Sci U S A *101*, 17033-17038.

Bloomgren, G., Richman, S., Hotermans, C.*, et al.* (2012). Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N Engl J Med *366*, 1870-1880.

Bofill-Mas, S., Formiga-Cruz, M., Clemente-Casares, P., Calafell, F., and Girones, R. (2001). Potential transmission of human polyomaviruses through the gastrointestinal tract after exposure to virions or viral DNA. J Virol *75*, 10290-10299.

Bollag, B., Kilpatrick, L.H., Tyagarajan, S.K., Tevethia, M.J., and Frisque, R.J. (2006). JC virus T'135, T'136 and T'165 proteins interact with cellular p107 and p130 in vivo and influence viral transformation potential. J Neurovirol *12*, 428-442.

Bonhaus, D.W., Bach, C., DeSouza, A., Salazar, F.H., Matsuoka, B.D., Zuppan, P., Chan, H.W., and Eglen, R.M. (1995). The pharmacology and distribution of human 5-hydroxytryptamine2B (5-HT2B) receptor gene products: comparison with 5-HT2A and 5-HT2C receptors. Br J Pharmacol *115*, 622-628.

Brady, J.N., Winston, V.D., and Consigli, R.A. (1977). Dissociation of polyoma virus by the chelation of calcium ions found associated with purified virions. J Virol *23*, 717-724.

Brew, B.J., Davies, N.W., Cinque, P., Clifford, D.B., and Nath, A. (2010). Progressive multifocal leukoencephalopathy and other forms of JC virus disease. Nat Rev Neurol *6*, 667-679.

Broekema, N.M., and Imperiale, M.J. (2013). miRNA regulation of BK polyomavirus replication during early infection. Proc Natl Acad Sci U S A *110*, 8200-8205.

Burioni, R., and Clementi, M. (2014). A human monoclonal antibody against the VP1 protein of JC virus. In WO2014002035 A1.

Burkert, O., Kressner, S., Sinn, L., Giese, S., Simon, C., and Lilie, H. (2014). Biophysical characterization of polyomavirus minor capsid proteins. Biol Chem *395*, 871-880.

Burmeister, W.P., Guilligay, D., Cusack, S., Wadell, G., and Arnberg, N. (2004). Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. J Virol *78*, 7727-7736.

Butin-Israeli, V., Ben-nun-Shaul, O., Kopatz, I., Adam, S.A., Shimi, T., Goldman, R.D., and Oppenheim, A. (2011). Simian virus 40 induces lamin A/C fluctuations and nuclear envelope deformation during cell entry. Nucleus *2*, 320-330.

Byres, E., Paton, A.W., Paton, J.C.*, et al.* (2008). Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. Nature *456*, 648-652.

Campanero-Rhodes, M.A., Smith, A., Chai, W.G.*, et al.* (2007). N-glycolyl GM1 ganglioside as a receptor for Simian Virus 40. J Virol *81*, 12846-12858.

Cao, W., Henry, M.D., Borrow, P., Yamada, H., Elder, J.H., Ravkov, E.V., Nichol, S.T., Compans, R.W., Campbell, K.P., and Oldstone, M.B.A. (1998). Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and lassa fever virus. Science *282*, 2079-2081.

Carbone, M., Ascione, G., Chichiarelli, S., Garcia, M.I., Eufemi, M., and Amati, P. (2004). Chromosome-protein interactions in polyomavirus virions. J Virol *78*, 513-519.

Carney, D.W., Nelson, C.D.S., Ferris, B.D., Stevens, J.P., Lipovsky, A., Kazakov, T., DiMaio, D., Atwood, W.J., and Sello, J.K. (2014). Structural optimization of a retrograde trafficking inhibitor that protects cells from infections by human polyoma- and papillomaviruses. Bioorganic & Medicinal Chemistry *22*, 4836-4847.

Carrillo-Tripp, M., Shepherd, C.M., Borelli, I.A., Venkataraman, S., Lander, G., Natarajan, P., Johnson, J.E., Brooks, C.L., 3rd, and Reddy, V.S. (2009). VIPERdb2: an enhanced and web API enabled relational database for structural virology. Nucleic Acids Research *37*, D436-442.

Carter, J.J., Daugherty, M.D., Qi, X., Bheda-Malge, A., Wipf, G.C., Robinson, K., Roman, A., Malik, H.S., and Galloway, D.A. (2013). Identification of an overprinting gene in Merkel Cell polyomavirus provides evolutionary insight into the birth of viral genes. Proc Natl Acad Sci U S A *110*, 12744-12749.

Carter, J.J., Paulson, K.G., Wipf, G.C.*, et al.* (2009). Association of Merkel Cell polyomavirus-specific antibodies with Merkel Cell carcinoma. J Natl Cancer Inst *101*, 1510-1522.

Caruso, M., Belloni, L., Sthandier, O., Amati, P., and Garcia, M.I. (2003). Alpha 4 beta 1 integrin acts as a cell receptor for murine polyomavirus at the postattachment level. J Virol *77*, 3913-3921.

Chapagain, M.L., and Nerurkar, V.R. (2010). Human polyomavirus JC (JCV) infection of human B lymphocytes: a possible mechanism for JCV transmigration across the blood-brain barrier. J Infect Dis *202*, 184-191.

Chapagain, M.L., Verma, S., Mercier, F., Yanagihara, R., and Nerurkar, V.R. (2007). Polyomavirus JC infects human brain microvascular endothelial cells independent of serotonin receptor 2A. Virology *364*, 55-63.

Chen, C.J., Cox, J.E., Azarm, K.D., Wylie, K.N., Woolard, K.D., Pesavento, P.A., and Sullivan, C.S. (2014). Identification of a polyomavirus microRNA highly expressed in tumors. Virology *476C*, 43-53.

Chen, H., Viel, S., Ziarelli, F., and Peng, L. (2013). F-19 NMR: a valuable tool for studying biological events. Chemical Society Reviews *42*, 7971-7982.

Chen, X.S., Stehle, T., and Harrison, S.C. (1998). Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry. EMBO J *17*, 3233-3240.

Chen, Y., Cheng, G., and Mahato, R.I. (2008). RNAi for treating hepatitis B viral infection. Pharm Res *25*, 72-86.

Childs, R.A., Palma, A.S., Wharton, S.*, et al.* (2009). Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. Nat Biotechnol *27*, 797-799.

Chilingaryan, Z., Yin, Z., and Oakley, A.J. (2012). Fragment-based screening by protein crystallography: successes and pitfalls. Int J Mol Sci *13*, 12857-12879.

Cho, J.A., Chinnapen, D.J., Aamar, E., te Welscher, Y.M., Lencer, W.I., and Massol, R. (2012). Insights on the trafficking and retro-translocation of glycosphingolipid-binding bacterial toxins. Front Cell Infect Microbiol *2*, 51.

Chou, H.H., Hayakawa, T., Diaz, S., Krings, M., Indriati, E., Leakey, M., Paabo, S., Satta, Y., Takahata, N., and Varki, A. (2002). Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. Proc Natl Acad Sci U S A *99*, 11736-11741.

Chou, H.H., Takematsu, H., Diaz, S., Iber, J., Nickerson, E., Wright, K.L., Muchmore, E.A., Nelson, D.L., Warren, S.T., and Varki, A. (1998). A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. Proc Natl Acad Sci U S A *95*, 11751-11756.

Ciulli, A., and Abell, C. (2007). Fragment-based approaches to enzyme inhibition. Curr Opin Biotechnol *18*, 489-496.

Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003). A 'rule of three' for fragment-based lead discovery? Drug Discov Today *8*, 876-877.

Coric, P., Saribas, A.S., Abou-Gharbia, M., Childers, W., White, M.K., Bouaziz, S., and Safak, M. (2014). Nuclear magnetic resonance structure revealed that the human polyomavirus JC virus agnoprotein contains an alpha-helix encompassing the Leu/Ile/Phe-rich domain. J Virol *88*, 11000-11000.

Dalianis, T., and Hirsch, H.H. (2013). Human polyomaviruses in disease and cancer. Virology *437*, 63-72.

Dalyot-Herman, N., Ben-nun-Shaul, O., Gordon-Shaag, A., and Oppenheim, A. (1996). The Simian Virus 40 packaging signal ses is composed of redundant DNA elements which are partly interchangeable. J Mol Biol *259*, 69-80.

Dang, X., and Koralnik, I.J. (2006). A granule cell neuron-associated JC virus variant has a unique deletion in the VP1 gene. J Gen Virol *87*, 2533-2537.

Dang, X., Vidal, J.E., Oliveira, A.C., Simpson, D.M., Morgello, S., Hecht, J.H., Ngo, L.H., and Koralnik, I.J. (2012a). JC virus granule cell neuronopathy is associated with VP1 C terminus mutants. J Gen Virol *93*, 175-183.

Dang, X., Wuthrich, C., Gordon, J., Sawa, H., and Koralnik, I.J. (2012b). JC virus encephalopathy is associated with a novel agnoprotein-deletion JCV variant. PLoS One *7*, e35793.

Daniels, R., Rusan, N.M., Wadsworth, P., and Hebert, D.N. (2006). SV40VP2 and VP3 insertion into ER membranes is controlled by the capsid protein VP1: Implications for DNA translocation out of the ER. Mol Cell *24*, 955-966.

Daniels, R., Sadowicz, D., and Hebert, D.N. (2007). A very late viral protein triggers the lytic release of SV40. PLoS Pathog *3*, e98.

Darbinyan, A., Kaminski, R., White, M.K., Darbinian-Sarkissian, N., and Khalili, K. (2013). Polyomavirus JC infection inhibits differentiation of oligodendrocyte progenitor cells. J Neurosci Res *91*, 116-127.

Davies, T.G., and Tickle, I.J. (2012). Fragment screening using X-ray crystallography. Top Curr Chem *317*, 33-59.

De Gascun, C.F., and Carr, M.J. (2013). Human polyomavirus reactivation: disease pathogenesis and treatment approaches. Clin Dev Immunol *2013*, 373579.

DeCaprio, J.A., and Garcea, R.L. (2013). A cornucopia of human polyomaviruses. Nat Rev Microbiol *11*, 264-276.

Delbue, S., Elia, F., Carloni, C., Tavazzi, E., Marchioni, E., Carluccio, S., Signorini, L., Novati, S., Maserati, R., and Ferrante, P. (2012). JC virus load in cerebrospinal fluid and transcriptional control region rearrangements may predict the clinical course of progressive multifocal leukoencephalopathy. J Cell Physiol *227*, 3511-3517.

Deng, L., Song, J., Gao, X., Wang, J., Yu, H., Chen, X., Varki, N., Naito-Matsui, Y., Galan, J.E., and Varki, A. (2014). Host adaptation of a bacterial toxin from the human pathogen Salmonella Typhi. Cell *159*, 1290-1299.

Diotti, R.A., Nakanishi, A., Clementi, N., Mancini, N., Criscuolo, E., Solforosi, L., and Clementi, M. (2013). JC polyomavirus (JCV) and monoclonal antibodies: friends or potential foes? Clin Dev Immunol *2013*, 967581.

Drayman, N., Glick, Y., Ben-nun-Shaul, O., Zer, H., Zlotnick, A., Gerber, D., Schueler-Furman, O., and Oppenheim, A. (2013). Pathogens use structural mimicry of native host ligands as a mechanism for host receptor engagement. Cell Host Microbe *14*, 63-73.

Du Pasquier, R.A., Corey, S., Margolin, D.H., Williams, K., Pfister, L.A., De Girolami, U., Mac Key, J.J., Wuthrich, C., Joseph, J.T., and Koralnik, I.J. (2003). Productive infection of cerebellar granule cell neurons by JC virus in an HIV+ individual. Neurology *61*, 775-782.

Dugan, A.S., Gasparovic, M.L., and Atwood, W.J. (2008). Direct correlation between sialic acid binding and infection of cells by two human polyomaviruses (JC virus and BK virus). J Virol *82*, 2560-2564.

Duncavage, E.J., and Pfeifer, J.D. (2011). Human Polyomaviruses 6 and 7 (HPyV6/7) are not detectable in Merkel Cell carcinoma. Modern Pathology *24*, 115A-115A.

Eash, S., and Atwood, W.J. (2005). Involvement of cytoskeletal components in BK virus infectious entry. J Virol *79*, 11734-11741.

Eash, S., Querbes, W., and Atwood, W.J. (2004). Infection of vero cells by BK virus is dependent on caveolae. J Virol *78*, 11583-11590.

Ehlers, B., and Moens, U. (2014). Genome analysis of non-human primate polyomaviruses. Infect Genet Evol *26*, 283-294.

Elphick, G.F., Querbes, W., Jordan, J.A.*, et al.* (2004). The human polyomavirus, JCV, uses serotonin receptors to infect cells. Science *306*, 1380-1383.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Engel, S., Heger, T., Mancini, R., Herzog, F., Kartenbeck, J., Hayer, A., and Helenius, A. (2011). Role of endosomes in Simian Virus 40 entry and infection. J Virol *85*, 4198-4211.

Erickson, K.D., Bouchet-Marquis, C., Heiser, K., Szomolanyi-Tsuda, E., Mishra, R., Lamothe, B., Hoenger, A., and Garcea, R.L. (2012). Virion assembly factories in the nucleus of polyomavirus-infected cells. PLoS Pathog *8*.

Erickson, K.D., Garcea, R.L., and Tsai, B. (2009). Ganglioside GT1b is a putative host cell receptor for the Merkel Cell polyomavirus. J Virol *83*, 10275-10279.

Erlanson, D.A. (2012). Introduction to fragment-based drug discovery. Top Curr Chem *317*, 1-32.

Ermis, U., Weis, J., and Schulz, J.B. (2013). Case reports of PML in patients treated for psoriasis. N Engl J Med *369*, 1081.

Ernst, B., and Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. Nature Reviews Drug Discovery *8*, 661-677.

Ewers, H., Romer, W., Smith, A.E.*, et al.* (2010). GM1 structure determines SV40-induced membrane invagination and infection. Nat Cell Biol *12*, 11-18; sup pp 11-12.

Feltkamp, M.C., Kazem, S., van der Meijden, E., Lauber, C., and Gorbalenya, A.E. (2013). From Stockholm to Malawi: recent developments in studying human polyomaviruses. J Gen Virol *94*, 482-496.

Feng, H., Shuda, M., Chang, Y., and Moore, P.S. (2008). Clonal integration of a polyomavirus in human Merkel Cell carcinoma. Science *319*, 1096-1100.

Ferenczy, M.W., Marshall, L.J., Nelson, C.D., Atwood, W.J., Nath, A., Khalili, K., and Major, E.O. (2012). Molecular biology, epidemiology, and pathogenesis of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. Clin Microbiol Rev *25*, 471-506.

Freund, R., Garcea, R.L., Sahli, R., and Benjamin, T.L. (1991). A single-amino-acid substitution in polyomavirus VP1 correlates with plaque size and hemagglutination behavior. J Virol *65*, 350-355.

Frisque, R.J. (1983). Nucleotide sequence of the region encompassing the JC virus origin of DNA replication. J Virol *46*, 170-176.

Frisque, R.J., Bollag, B., Tyagarajan, S.K., and Kilpatrick, L.H. (2003). T' proteins influence JC virus biology. J Neurovirol *9 Suppl 1*, 15-20.

Frisque, R.J., Bream, G.L., and Cannella, M.T. (1984). Human polyomavirus JC virus genome. J Virol *51*, 458-469.

Fromm, M., and Berg, P. (1983). Simian Virus 40 early- and late-region promoter functions are enhanced by the 72-base-pair repeat inserted at distant locations and inverted orientations. Mol Cell Biol *3*, 991-999.

Fujinaga, Y., Wolf, A.A., Rodighiero, C., Wheeler, H., Tsai, B., Allen, L., Jobling, M.G., Rapoport, T., Holmes, R.K., and Lencer, W.I. (2003). Gangliosides that associate with lipid rafts mediate transport of cholera and related toxins from the plasma membrane to endoplasmic reticulm. Molecular Biology of the Cell *14*, 4783-4793.

Gardner, S.D., Field, A.M., Coleman, D.V., and Hulme, B. (1971). New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet *1*, 1253-1257.

Gasparovic, M.L., Gee, G.V., and Atwood, W.J. (2006). JC virus minor capsid proteins VP2 and VP3 are essential for virus propagation. J Virol *80*, 10858-10861.

Gee, G.V., Manley, K., and Atwood, W.J. (2003). Derivation of a JC virus-resistant human glial cell line: implications for the identification of host cell factors that determine viral tropism. Virology *314*, 101-109.

Geiger, R., Andritschke, D., Friebe, S., Herzog, F., Luisoni, S., Heger, T., and Helenius, A. (2011). BAP31 and BiP are essential for dislocation of SV40 from the endoplasmic reticulum to the cytosol. Nat Cell Biol *13*, 1305-1314.

Giannetti, A.M. (2011). From Experimental Design to Validated Hits: A Comprehensive Walk-through of Fragment Lead Identification Using Surface Plasmon Resonance. Fragment-Based Drug Design: Tools, Practical Approaches, and Examples *493*, 169-218.

Gilbert, J., and Benjamin, T. (2004). Uptake pathway of polyomavirus via ganglioside GD1a. J Virol *78*, 12259-12267.

Gilbert, J., Ou, W., Silver, J., and Benjamin, T. (2006). Downregulation of protein disulfide isomerase inhibits infection by the mouse polyomavirus. J Virol *80*, 10868-10870.

Gilbert, J.M., and Benjamin, T.L. (2000). Early steps of polyomavirus entry into cells. J Virol *74*, 8582-8588.

Gillock, E.T., Rottinghaus, S., Chang, D., Cai, X., Smiley, S.A., An, K., and Consigli, R.A. (1997). Polyomavirus major capsid protein VP1 is capable of packaging cellular DNA when expressed in the baculovirus system. J Virol *71*, 2857-2865.

Giorda, K.M., Raghava, S., Zhang, M.W., and Hebert, D.N. (2013). The viroporin activity of the minor structural proteins VP2 and VP3 is required for SV40 propagation. Journal of Biological Chemistry *288*, 2510-2520.

Goodwin, E.C., Atwood, W.J., and DiMaio, D. (2009). High-throughput cell-based screen for chemicals that inhibit infection by simian virus 40 and human polyomaviruses. J Virol *83*, 5630-5639.

Goodwin, E.C., Lipovsky, A., Inoue, T.*, et al.* (2011). BiP and multiple DNAJ molecular chaperones in the endoplasmic reticulum are required for efficient Simian Virus 40 infection. MBio *2*, e00101-00111.

Gorelik, L., Reid, C., Testa, M.*, et al.* (2011). Progressive multifocal leukoencephalopathy (PML) development is associated with mutations in JC virus capsid protein VP1 that change its receptor specificity. Journal of Infectious Diseases *204*, 103-114.

Gray, F., Bazille, C., Adle-Biassette, H., Mikol, J., Moulignier, A., and Scaravilli, F. (2005). Central nervous system immune reconstitution disease in acquired immunodeficiency syndrome patients receiving highly active antiretroviral treatment. J Neurovirol *11 Suppl 3*, 16-22.

Griffith, J.D. (1975). Chromatin structure: deduced from a minichromosome. Science *187*, 1202-1203.

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.

Hajduk, P.J., Galloway, W.R., and Spring, D.R. (2011). Drug discovery: A question of library design. Nature *470*, 42-43.

Hamza, I.A., Jurzik, L., Stang, A., Sure, K., Uberla, K., and Wilhelm, M. (2009). Detection of human viruses in rivers of a densly-populated area in Germany using a virus adsorption elution method optimized for PCR analyses. Water Research *43*, 2657-2668.

Han, L., Kitova, E.N., Tan, M., Jiang, X., Pluvinage, B., Boraston, A.B., and Klassen, J.S. (2015). Affinities of human histo-blood group antigens for norovirus capsid protein complexes. Glycobiology *25*, 170-180.

Hayer, A., Stoeber, M., Ritz, D., Engel, S., Meyer, H.H., and Helenius, A. (2010). Caveolin-1 is ubiquitinated and targeted to intralumenal vesicles in endolysosomes for degradation. Journal of Cell Biology *191*, 615-629.

Henrich, T.J., and Kuritzkes, D.R. (2013). HIV-1 entry inhibitors: recent development and clinical use. Curr Opin Virol *3*, 51-57.

Hidari, K.I., Shimada, S., Suzuki, Y., and Suzuki, T. (2007). Binding kinetics of influenza viruses to sialic acid-containing carbohydrates. Glycoconj J *24*, 583-590.

Ho, J., Jedrych, J.J., Feng, H.*, et al.* (2014). Human Polyomavirus 7-associated pruritic rash and viremia in transplant recipients. J Infect Dis.

Holthofer, H., Reivinen, J., and Miettinen, A. (1994). Nephron segment and cell-type specific expression of gangliosides in the developing and adult kidney. Kidney Int *45*, 123-130.

Hoque, K.M., Azim, M.F., Mia, M.R., Kayesh, R., Ali, M.H., Islam, M.J., Shakil, S.K., and Shuvra, T.M. (2012). Design of potential siRNA molecules for T antigen gene silencing of Merkel Cell Polyomavirus. Bioinformation *8*, 924-930.

Houben, R., Shuda, M., Weinkam, R., Schrama, D., Feng, H., Chang, Y., Moore, P.S., and Becker, J.C. (2010). Merkel Cell polyomavirus-infected Merkel Cell carcinoma cells require expression of viral T antigens. J Virol *84*, 7064-7072.

Hu, L.Y., Crawford, S.E., Czako, R., Cortes-Penfield, N.W., Smith, D.F., Le Pendu, J., Estes, M.K., and Prasad, B.V.V. (2012). Cell attachment protein VP8*of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature *485*, 256-U147.

Hughes, M.P., Hardee, M.E., Cornelius, L.A., Hutchins, L.F., Becker, J.C., and Gao, L. (2014). Merkel Cell carcinoma: epidemiology, target, and therapy. Curr Dermatol Rep *3*, 46-53.

Imamura, T., Okamoto, M., Nakakita, S.*, et al.* (2014). Antigenic and receptor binding properties of enterovirus 68. J Virol *88*, 2374-2384.

Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., and Suzuki, A. (1998). The molecular basis for the absence of N-glycolylneuraminic acid in humans. J Biol Chem *273*, 15866-15871.

Ishizu, K.I., Watanabe, H., Han, S.I., Kanesashi, S.N., Hoque, M., Yajima, H., Kataoka, K., and Handa, H. (2001). Roles of disulfide linkage and calcium ion-mediated interactions in assembly

and disassembly of virus-like particles composed of Simian Virus 40 VP1 capsid protein. J Virol *75*, 61-72.

Ishov, A.M., and Maul, G.G. (1996). The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition. Journal of Cell Biology *134*, 815-826.

Ison, M.G. (2011). Antivirals and resistance: influenza virus. Curr Opin Virol *1*, 563-573.

Jao, C.C., Weidman, M.K., Perez, A.R., and Gharakhanian, E. (1999). Cys9, Cys104 and Cys207 of simian virus 40 VP1 are essential for inter-pentamer disulfide-linkage and stabilization in cell-free lysates. J Gen Virol *80 (Pt 9)*, 2481-2489.

Jay, G., Nomura, S., Anderson, C.W., and Khoury, G. (1981). Identification of the SV40 agnogene product: a DNA binding protein. Nature *291*, 346-349.

Jensen, P.N., and Major, E.O. (2001). A classification scheme for human polyomavirus JCV variants based on the nucleotide sequence of the noncoding regulatory region. J Neurovirol *7*, 280-287.

Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. Assay and Drug Development Technologies *9*, 342-353.

Jiang, Z.G., Cohen, J., Marshall, L.J., and Major, E.O. (2010). Hexadecyloxypropyl-cidofovir (CMX001) suppresses JC virus replication in human fetal brain SVG cell cultures. Antimicrob Agents Chemother *54*, 4723-4732.

Jobling, M.G., Yang, Z., Kam, W.R., Lencer, W.I., and Holmes, R.K. (2012). A single native ganglioside GM1-binding site is sufficient for cholera toxin to bind to cells and complete the intoxication pathway. MBio *3*.

Johne, R., Buck, C.B., Allander, T., Atwood, W.J., Garcea, R.L., Imperiale, M.J., Major, E.O., Ramqvist, T., and Norkin, L.C. (2011). Taxonomical developments in the family Polyomaviridae. Archives of Virology *156*, 1627-1634.

Jordan, J.B., Poppe, L., Xia, X.Y., Cheng, A.C., Sun, Y., Michelsen, K., Eastwood, H., Schnier, P.D., Nixey, T., and Zhong, W.G. (2012). Fragment based drug discovery: practical implementation based on F-19 NMR spectroscopy. J Med Chem *55*, 678-687.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kassem, A., Schopflin, A., Diaz, C., Weyers, W., Stickeler, E., Werner, M., and Zur Hausen, A. (2008). Frequent detection of Merkel Cell polyomavirus in human Merkel Cell carcinomas and identification of a unique deletion in the VP1 gene. Cancer Res *68*, 5009-5013.

Kawamoto, E., Nakahashi, S., Okamoto, T., Imai, H., and Shimaoka, M. (2012). Anti-integrin therapy for multiple sclerosis. Autoimmune Dis *2012*, 357101.

Kawano, M., Inoue, T., Tsukamoto, H.*, et al.* (2006). The VP2/VP3 minor capsid protein of Aimian Virus 40 promotes the in vitro assembly of the major capsid protein VP1 into particles. Journal of Biological Chemistry *281*, 10164-10173.

Kawano, M.A., Xing, L., Tsukamoto, H., Inoue, T., Handa, H., and Cheng, R.H. (2009). Calcium bridge triggers capsid disassembly in the cell entry process of Simian Virus 40. J Biol Chem *284*, 34703-34712.

Kazem, S., van der Meijden, E., and Feltkamp, M.C. (2013). The trichodysplasia spinulosa-associated polyomavirus: virological background and clinical implications. Apmis *121*, 770-782.

Kazem, S., van der Meijden, E., Kooijman, S.*, et al.* (2012). Trichodysplasia spinulosa is characterized by active polyomavirus infection. Journal of Clinical Virology *53*, 225-230.

Kazem, S., van der Meijden, E., Wang, R.C., Rosenberg, A.S., Pope, E., Benoit, T., Fleckman, P., and Feltkamp, M.C. (2014). Polyomavirus-associated trichodysplasia spinulosa involves hyperproliferation, pRB phosphorylation and upregulation of p16 and p21. PLoS One *9*, e108947.

Kean, J.M., Rao, S., Wang, M., and Garcea, R.L. (2009). Seroepidemiology of human polyomaviruses. PLoS Pathog *5*, e1000363.

Khan, Z.M., Liu, Y., Neu, U., Gilbert, M., Ehlers, B., Feizi, T., and Stehle, T. (2014). Crystallographic and glycan microarray analysis of Human Polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate. J Virol *88*, 6100-6111.

Knowles, W.A., Pipkin, P., Andrews, N., Vyse, A., Minor, P., Brown, D.W., and Miller, E. (2003). Population-based study of antibody to the human polyomaviruses BKV and JCV and the simian polyomavirus SV40. J Med Virol *71*, 115-123.

Komagome, R., Sawa, H., Suzuki, T., Suzuki, Y., Tanaka, S., Atwood, W.J., and Nagashima, K. (2002). Oligosaccharides as receptors for JC virus. J Virol *76*, 12992-13000.

Konopka-Anstadt, J.L., Mainou, B.A., Sutherland, D.M., Sekine, Y., Strittmatter, S.M., and Dermody, T.S. (2014). The Nogo receptor NgR1 mediates infection by mammalian reovirus. Cell Host Microbe *15*, 681-691.

Koralnik, I.J., Wuthrich, C., Dang, X., Rottnek, M., Gurtman, A., Simpson, D., and Morgello, S. (2005). JC virus granule cell neuronopathy: A novel clinical syndrome distinct from progressive multifocal leukoencephalopathy. Annals of Neurology *57*, 576-580.

Korup, S., Rietscher, J., Calvignac-Spencer, S., Trusch, F., Hofmann, J., Moens, U., Sauer, I., Voigt, S., Schmuck, R., and Ehlers, B. (2013). Identification of a novel human polyomavirus in organs of the gastrointestinal tract. PLoS One *8*, e58021.

Kosukegawa, A., Arisaka, F., Takayama, M., Yajima, H., Kaidow, A., and Handa, H. (1996). Purification and characterization of virus-like particles and pentamers produced by the expression of SV40 capsid proteins in insect cells. Biochimica Et Biophysica Acta-General Subjects *1290*, 37-45.

Krauzewicz, N., Streuli, C.H., Stuart-Smith, N., Jones, M.D., Wallace, S., and Griffin, B.E. (1990). Myristylated polyomavirus VP2: role in the life cycle of the virus. J Virol *64*, 4414-4420.

Krissinel, E., and 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.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

Lane, D.P., and Crawford, L.V. (1979). T-antigen is bound to a host protein in SV40-transformed cells. Nature *278*, 261-263.

Laude, H.C., Jonchere, B., Maubec, E.*, et al.* (2010). Distinct Merkel Cell polyomavirus molecular features in tumour and non tumour specimens from patients with Merkel Cell carcinoma. PLoS Pathog *6*, e1001076.

Lee, S., Paulson, K.G., Murchison, E.P., Afanasiev, O.K., Alkan, C., Leonard, J.H., Byrd, D.R., Hannon, G.J., and Nghiem, P. (2011). Identification and validation of a novel mature microRNA encoded by the Merkel Cell polyomavirus in human Merkel Cell carcinomas. J Clin Virol *52*, 272-275.

Li, J.J., and Kelly, T.J. (1984). Simian Virus 40 DNA replication in vitro. Proc Natl Acad Sci U S A *81*, 6973-6977.

Li, P.P., Nakanishi, A., Shum, D., Sun, P.C.K., Salazar, A.M., Fernandez, C.F., Chan, S.W., and Kasamatsu, H. (2001). Simian Virus 40 VP1 DNA-binding domain is functionally separable from the overlapping nuclear localization signal and is required for effective virion formation and full viability. J Virol *75*, 7321-7329.

Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L., and Harrison, S.C. (1991). Structure of Simian Virus 40 at 3.8-A resolution. Nature *354*, 278-284.

Lilley, B.N., Gilbert, J.M., Ploegh, H.L., and Benjamin, T.L. (2006). Murine polyomavirus requires the endoplasmic reticulum protein Derlin-2 to initiate infection. J Virol *80*, 8739-8744.

Lim, E.S., Reyes, A., Antonio, M.*, et al.* (2013). Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. Virology *436*, 295-303.

Lingwood, D., and Simons, K. (2010). Lipid rafts as a membrane-organizing principle. Science *327*, 46-50.

Linzer, D.I., and Levine, A.J. (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. Cell *17*, 43-52.

Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev *46*, 3-26.

London, W.T., Houff, S.A., Madden, D.L.*, et al.* (1978). Brain tumors in owl monkeys inoculated with a human polyomavirus (JC virus). Science *201*, 1246-1249.

Lopez, P.H.H., and Schnaar, R.L. (2009). Gangliosides in cell recognition and membrane protein regulation. Current Opinion in Structural Biology *19*, 549-557.

Lou, Z., Sun, Y., and Rao, Z. (2014). Current progress in antiviral strategies. Trends Pharmacol Sci *35*, 86-102.

Low, J.A., Magnuson, B., Tsai, B., and Imperiale, M.J. (2006). Identification of gangliosides GD1b and GT1b as receptors for BK virus. J Virol *80*, 1361-1366.

Maginnis, M.S., Haley, S.A., Gee, G.V., and Atwood, W.J. (2010). Role of N-linked glycosylation of the 5-HT2A receptor in JC virus infection. J Virol *84*, 9677-9684.

Magnuson, B., Rainey, E.K., Benjamin, T., Baryshev, M., Mkrtchian, S., and Tsai, B. (2005). ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding. Mol Cell *20*, 289-300.

Major, E.O. (2010). Progressive multifocal leukoencephalopathy in patients on immunomodulatory therapies. Annu Rev Med *61*, 35-47.

Major, E.O., Miller, A.E., Mourrain, P., Traub, R.G., de Widt, E., and Sever, J. (1985). Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc Natl Acad Sci U S A *82*, 1257-1261.

Mancuso, C.E., Gabay, M.P., Steinke, L.M., and Vanosdol, S.J. (2010). Peramivir: an intravenous neuraminidase inhibitor for the treatment of 2009 H1N1 influenza. Ann Pharmacother *44*, 1240-1249.

Mannhold, R., Poda, G.I., Ostermann, C., and Tetko, I.V. (2009). Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. J Pharm Sci *98*, 861-893.

Martin, J.D., King, D.M., Slauch, J.M., and Frisque, R.J. (1985). Differences in regulatory sequences of naturally occurring JC virus variants. J Virol *53*, 306-311.

Matthews, M.R., Wang, R.C., Reddick, R.L., Saldivar, V.A., and Browning, J.C. (2011). Viral-associated trichodysplasia spinulosa: a case with electron microscopic and molecular detection of the trichodysplasia spinulosa-associated human polyomavirus. J Cutan Pathol *38*, 420-431.

Mayer, M., and Meyer, B. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angewandte Chemie-International Edition *38*, 1784-1788.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. Journal of Applied Crystallography *40*, 658-674.

McQuaig, S.M., Scott, T.M., Lukasik, J.O., Paul, J.H., and Harwood, V.J. (2009). Quantification of human polyomaviruses JC Virus and BK Virus by TaqMan quantitative PCR and comparison to other water quality indicators in water and fecal samples. Appl Environ Microbiol *75*, 3379-3388.

Mishra, N., Pereira, M., Rhodes, R.H., An, P., Pipas, J.M., Jain, K., Kapoor, A., Briese, T., Faust, P.L., and Lipkin, W.I. (2014). Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. J Infect Dis.

Moens, U., Van Ghelue, M., and Ehlers, B. (2014). Are human polyomaviruses co-factors for cancers induced by other oncoviruses? Reviews in Medical Virology *24*, 343-360.

Monaco, M.C., Atwood, W.J., Gravell, M., Tornatore, C.S., and Major, E.O. (1996). JC virus infection of hematopoietic progenitor cells, primary B lymphocytes, and tonsillar stromal cells: implications for viral latency. J Virol *70*, 7004-7012.

Monaco, M.C., Jensen, P.N., Hou, J., Durham, L.C., and Major, E.O. (1998). Detection of JC virus DNA in human tonsil tissue: evidence for site of initial viral infection. J Virol *72*, 9918-9923.

Moreland, R.B., Montross, L., and Garcea, R.L. (1991). Characterization of the DNA-binding properties of the polyomavirus capsid protein VP1. J Virol *65*, 1168-1176.

Muller, B., and Krausslich, H.G. (2009). Antiviral strategies. Handb Exp Pharmacol, 1-24.

Murray, C.W., Callaghan, O., Chessari, G., Cleasby, A., Congreve, M., Frederickson, M., Hartshorn, M.J., McMenamin, R., Patel, S., and Wallis, N. (2007). Application of fragment screening by X-ray crystallography to beta-secretase. J Med Chem *50*, 1116-1123.

Murray, C.W., Verdonk, M.L., and Rees, D.C. (2012). Experiences in fragment-based drug discovery. Trends Pharmacol Sci *33*, 224-232.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Nakanishi, A., Li, P.P., Qu, Q., Jafri, Q.H., and Kasamatsu, H. (2007). Molecular dissection of nuclear entry-competent SV40 during infection. Virus Res *124*, 226-230.

Nakanishi, A., Nakamura, A., Liddington, R., and Kasamatsu, H. (2006). Identification of amino acid residues within Simian Virus 40 capsid proteins VP1, VP2, and VP3 that are required for their interaction and for viral infection. J Virol *80*, 8891-8898.

Nakanishi, A., Shum, D., Morioka, H., Otsuka, E., and Kasamatsu, H. (2002). Interaction of the VP3 nuclear localization signal with the importin alpha-2/beta heterodimer directs nuclear entry of infecting Simian Virus 40. J Virol *76*, 9368-9377.

Nelson, C.D., Carney, D.W., Derdowski, A., Lipovsky, A., Gee, G.V., O'Hara, B., Williard, P., DiMaio, D., Sello, J.K., and Atwood, W.J. (2013). A retrograde trafficking inhibitor of ricin and Shiga-like toxins inhibits infection of cells by human and monkey polyomaviruses. MBio *4*, e00729-00713.

Nelson, C.D., Ströh, L.J., Gee, G.V., O'Hara, B.A., Stehle, T., and Atwood, W.J. (2015). Modulation of a pore in the capsid of JC polyomavirus reduces infectivity and prevents exposure of the minor capsid proteins. J Virol.

Nelson, C.D.S., Derdowski, A., Maginnis, M.S., O'Hara, B.A., and Atwood, W.J. (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.

Neu, U., Allen, S.A., Blaum, B.S*., et al.* (2013a). A structure-guided mutation in the major capsid protein retargets BK polyomavirus. PLoS Pathog *9*, e1003688.

Neu, U., Hengel, H., Blaum, B.r.S*., et al.* (2012). Structures of Merkel Cell polyomavirus VP1 complexes define a sialic acid binding site required for infection. PLoS Pathog *8*, e1002738.

Neu, U., Khan, Z.M., Schuch, B., Palma, A.S., Liu, Y., Pawlita, M., Feizi, T., and Stehle, T. (2013b). Structures of B-lymphotropic polyomavirus VP1 in complex with oligosaccharide ligands. PLoS Pathog *9*, e1003714.

Neu, U., Maginnis, M.S., Palma, A.S., Ströh, L.J., Nelson, C.D., Feizi, T., Atwood, W.J., and 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.

Neu, U., Wang, J., Macejak, D., Garcea, R.L., and Stehle, T. (2011). Structures of the major capsid proteins of the human Karolinska Institutet and Washington University polyomaviruses. J Virol *85*, 7384-7392.

Neu, U., Woellner, K., Gauglitz, G., and Stehle, T. (2008). Structural basis of GM1 ganglioside recognition by simian virus 40. Proc Natl Acad Sci U S A *105*, 5219-5224.

Nicol, J.T., Leblond, V., Arnold, F., Guerra, G., Mazzoni, E., Tognon, M., Coursaget, P., and Touze, A. (2014). Seroprevalence of human Malawi polyomavirus. J Clin Microbiol *52*, 321-323.

Nicol, J.T., Robinot, R., Carpentier, A., Carandina, G., Mazzoni, E., Tognon, M., Touze, A., and Coursaget, P. (2013). Age-specific seroprevalences of Merkel Cell polyomavirus, Human

Polyomaviruses 6, 7, and 9, and Trichodysplasia Spinulosa-associated polyomavirus. Clin Vaccine Immunol *20*, 363-368.

Nicol, J.T., Touze, A., Robinot, R., Arnold, F., Mazzoni, E., Tognon, M., and Coursaget, P. (2012). Seroprevalence and cross-reactivity of human polyomavirus 9. Emerg Infect Dis *18*, 1329-1332.

Nilsson, E.C., Storm, R.J., Bauer, J.*, et al.* (2011). The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. Nature Medicine *17*, 105-U279.

Nilsson, O., and Svennerholm, L. (1982). Characterization and quantitative-determination of gangliosides and neutral glycosphingolipids in human-liver. Journal of Lipid Research *23*, 327-334.

O'Donnell, C.D., Kovacs, M., Akhtar, J., Valyi-Nagy, T., and Shukla, D. (2010). Expanding the role of 3-O sulfated heparan sulfate in herpes simplex virus type-1 entry. Virology *397*, 389-398.

O'Hara, B.A., Rupasinghe, C., Yatawara, A., Gaidos, G., Mierke, D.F., and Atwood, W.J. (2014). Gallic acid-based small-molecule inhibitors of JC and BK polyomaviral infection. Virus Res *189*, 280-285.

Ohsumi, S., Motoi, M., and Ogawa, K. (1986). Induction of undifferentiated tumors by JC virus in the cerebrum of rats. Acta Pathol Jpn *36*, 815-825.

Okada, Y., Sawa, H., Endo, S., Orba, Y., Umemura, T., Nishihara, H., Stan, A.C., Tanaka, S., Takahashi, H., and Nagashima, K. (2002). Expression of JC virus agnoprotein in progressive multifocal leukoencephalopathy brain. Acta Neuropathol *104*, 130-136.

Okomo-Adhiambo, M., Fry, A.M., Su, S.*, et al.* (2015). Oseltamivir-resistant influenza A(H1N1)pdm09 viruses, United States, 2013-14. Emerg Infect Dis *21*, 136-141.

Oppenheim, A., Sandalon, Z., Peleg, A., Shaul, O., Nicolis, S., and Ottolenghi, S. (1992). A cis-acting DNA signal for encapsidation of Simian Virus 40. J Virol *66*, 5320-5328.

Osswald, S.S., Kulick, K.B., Tomaszewski, M.M., and Sperling, L.C. (2007). Viral-associated trichodysplasia in a patient with lymphoma: a case report and review. J Cutan Pathol *34*, 721-725.

Padgett, B.L., Walker, D.L., ZuRhein, G.M., Eckroade, R.J., and Dessel, B.H. (1971). Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet *1*, 1257-1260.

Padler-Karavani, V., Yu, H., Cao, H., Chokhawala, H., Karp, F., Varki, N., Chen, X., and Varki, A. (2008). Diversity in specificity, abundance, and composition of anti-Neu5Gc antibodies in normal humans: potential implications for disease. Glycobiology *18*, 818-830.

Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr *62*, 439-450.

Palma, A.S., Feizi, T., Childs, R.A., Chai, W., and Liu, Y. (2014). The neoglycolipid (NGL)-based oligosaccharide microarray system poised to decipher the meta-glycome. Curr Opin Chem Biol *18*, 87-94.

Parton, R.G., and del Pozo, M.A. (2013). Caveolae as plasma membrane sensors, protectors and organizers. Nat Rev Mol Cell Biol *14*, 98-112.

Pelkmans, L., Kartenbeck, J., and 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.

Peretti, A., FitzGerald, P.C., Bliskovsky, V., Pastrana, D.V., and Buck, C.B. (2015). Genome sequence of a fish-associated polyomavirus, Black Sea Bass (Centropristis striata) Polyomavirus 1. Genome Announc *3*.

Pho, M.T., Ashok, A., and Atwood, W.J. (2000). JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. J Virol *74*, 2288-2292.

Pietropaolo, V., Videtta, M., Fioriti, D., Mischitelli, M., Arancio, A., Orsi, N., and Degener, A.M. (2003). Rearrangement patterns of JC virus noncoding control region from different biological samples. J Neurovirol *9*, 603-611.

Prokazova, N.V., Samovilova, N.N., Gracheva, E.V., and Golovanova, N.K. (2009). Ganglioside GM3 and its biological functions. Biochemistry (Mosc) *74*, 235-249.

Qian, M., Cai, D., Verhey, K.J., and Tsai, B. (2009). A lipid receptor sorts polyomavirus from the endolysosome to the endoplasmic reticulum to cause infection. PLoS Pathog *5*, e1000465.

Qian, M., and Tsai, B. (2010). Lipids and proteins act in opposing manners to regulate polyomavirus infection. J Virol *84*, 9840-9852.

Qu, Q.M., Sawa, H., Suzuki, T., Semba, S., Henmi, C., Okada, Y., Tsuda, M., Tanaka, S., Atwood, W.J., and Nagashima, K. (2004). Nuclear entry mechanism of the human polyomavirus JC virus-like particle - Role of importins and the nuclear pore complex. Journal of Biological Chemistry *279*, 27735-27742.

Querbes, W., Benmerah, A., Tosoni, D., Di Fiore, P.P., and Atwood, W.J. (2004). A JC virus-induced signal is required for infection of glial cells by a clathrin- and eps15-dependent pathway. J Virol *78*, 250-256.

Querbes, W., O'Hara, B.A., Williams, G., and Atwood, W.J. (2006). Invasion of host cells by JC virus identifies a novel role for caveolae in endosomal sorting of noncaveolar ligands. J Virol *80*, 9402-9413.

Rademacher, C., Guiard, J., Kitov, P.I., Fiege, B., Dalton, K.P., Parra, F., Bundle, D.R., and Peters, T. (2011). Targeting norovirus infection-multivalent entry inhibitor design based on NMR experiments. Chemistry *17*, 7442-7453.

Radhakrishnan, S., Otte, J., Enam, S., Del Valle, L., Khalili, K., and Gordon, J. (2003). JC virus-induced changes in cellular gene expression in primary human astrocytes. J Virol *77*, 10638-10644.

Raghava, S., Giorda, K.M., Romano, F.B., Heuck, A.P., and Hebert, D.N. (2011). The SV40 late protein VP4 is a viroporin that forms pores to disrupt membranes for viral release. PLoS Pathog *7*.

Raghava, S., Giorda, K.M., Romano, F.B., Heuck, A.P., and Hebert, D.N. (2013). SV40 late protein VP4 forms toroidal pores to disrupt membranes for viral release. Biochemistry *52*, 3939-3948.

Randhawa, P., Zeng, G., Bueno, M.*, et al.* (2014). Inhibition of large T antigen ATPase activity as a potential strategy to develop anti-polyomavirus JC drugs. Antiviral Res *112*, 113-119.

Reid, C.E., Li, H., Sur, G.*, et al.* (2011). Sequencing and analysis of JC virus DNA from natalizumab-treated PML patients. J Infect Dis *204*, 237-244.

Reiss, K., Stencel, J.E., Liu, Y., Blaum, B.S., Reiter, D.M., Feizi, T., Dermody, T.S., and Stehle, T. (2012). The GM2 glycan serves as a functional coreceptor for serotype 1 reovirus. PLoS Pathog *8*, e1003078.

Rennspiess, D., Pujari, S., Keijzers, M.*, et al.* (2015). Detection of human polyomavirus 7 in human thymic epithelial tumors. J Thorac Oncol *10*, 360-366.

Riley, M.I., Yoo, W., Mda, N.Y., and Folk, W.R. (1997). Tiny T antigen: an autonomous polyomavirus T antigen amino-terminal domain. J Virol *71*, 6068-6074.

Rinaldo, C.H., Gosert, R., Bernhoff, E., Finstad, S., and Hirsch, H.H. (2010). 1-O-hexadecyloxypropyl cidofovir (CMX001) effectively inhibits polyomavirus BK replication in primary human renal tubular epithelial cells. Antimicrob Agents Chemother *54*, 4714-4722.

Rinaldo, C.H., and Hirsch, H.H. (2013). The human polyomaviruses: from orphans and mutants to patchwork family. Apmis *121*, 681-684.

Rinaldo, C.H., Traavik, T., and Hey, A. (1998). The agnogene of the human polyomavirus BK is expressed. J Virol *72*, 6233-6236.

Rinaldo, C.H., Tylden, G.D., and Sharma, B.N. (2013). The human polyomavirus BK (BKPyV): virological background and clinical implications. Apmis *121*, 728-745.

Rishton, G.M. (2003). Nonleadlikeness and leadlikeness in biochemical screening. Drug Discov Today *8*, 86-96.

Sadeghi, M., Aaltonen, L.M., Hedman, L., Chen, T., Soderlund-Venermo, M., and Hedman, K. (2014). Detection of TS polyomavirus DNA in tonsillar tissues of children and adults: Evidence for site of viral latency. J Clin Virol *59*, 55-58.

Sahli, R., Freund, R., Dubensky, T., Garcea, R., Bronson, R., and Benjamin, T. (1993). Defect in entry and altered pathogenicity of a polyoma virus mutant blocked in VP2 myristylation. Virology *192*, 142-153.

Salunke, D.M., Caspar, D.L., and Garcea, R.L. (1986). Self-assembly of purified polyomavirus capsid protein VP1. Cell *46*, 895-904.

Samraj, A.N., Laubli, H., Varki, N., and Varki, A. (2014). Involvement of a non-human sialic acid in human cancer. Front Oncol *4*, 33.

Saribas, A.S., White, M.K., and Safak, M. (2012). JC virus agnoprotein enhances large T antigen binding to the origin of viral DNA replication: evidence for its involvement in viral DNA replication. Virology *433*, 12-26.

Sariyer, I.K., Saribas, A.S., White, M.K., and Safak, M. (2011). Infection by agnoprotein-negative mutants of polyomavirus JC and SV40 results in the release of virions that are mostly deficient in DNA content. Virology Journal *8*.

Sastre-Garau, X., Peter, M., Avril, M.F.*, et al.* (2009). Merkel Cell carcinoma of the skin: pathological and molecular evidence for a causative role of MCV in oncogenesis. J Pathol *218*, 48-56.

Sato, C., and Kitajima, K. (2013). Disialic, oligosialic and polysialic acids: distribution, functions and related disease. J Biochem *154*, 115-136.

Sauvage, V., Foulongne, V., Cheval, J.*, et al.* (2011). Human polyomavirus related to African green monkey lymphotropic polyomavirus. Emerg Infect Dis *17*, 1364-1370.

Schelhaas, M., Malmstrom, J., Pelkmans, L., Haugstetter, J., Ellgaard, L., Grunewald, K., and Helenius, A. (2007). Simian Virus 40 depends on ER protein folding and quality control factors for entry into host cells. Cell *131*, 516-529.

Schippling, S., Kempf, C., Buchele, F.*, et al.* (2013). JC virus granule cell neuronopathy and GCN-IRIS under natalizumab treatment. Annals of Neurology *74*, 622-626.

Schnaar, R.L., Suzuki, A., and Stanley, P. (2009). Glycosphingolipids. In Essentials of Glycobiology, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi*, et al.*, eds. (Cold Spring Harbor (NY)).

Schowalter, R.M., and Buck, C.B. (2013). The Merkel Cell polyomavirus minor capsid protein. PLoS Pathog *9*, e1003558.

Schowalter, R.M., Pastrana, D.V., and Buck, C.B. (2011). Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel Cell polyomavirus infectious entry. PLoS Pathog *7*, e1002161.

Schowalter, R.M., Pastrana, D.V., Pumphrey, K.A., Moyer, A.L., and Buck, C.B. (2010). Merkel Cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe *7*, 509-515.

Schrama, D., Buck, C.B., Houben, R., and Becker, J.C. (2012). No evidence for association of HPyV6 or HPyV7 with different skin cancers. J Invest Dermatol *132*, 239-241.

Schuttelkopf, A.W., and van Aalten, D.M. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr *60*, 1355-1363.

Scola, N., Wieland, U., Silling, S., Altmeyer, P., Stucker, M., and 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.

Scuda, N., Hofmann, J., Calvignac-Spencer, S., Ruprecht, K., Liman, P., Kuhn, J., Hengel, H., and Ehlers, B. (2011). A novel human polyomavirus closely related to the african green monkey-derived lymphotropic polyomavirus. J Virol *85*, 4586-4590.

Scuda, N., Madinda, N.F., Akoua-Koffi, C.*, et al.* (2013). Novel polyomaviruses of nonhuman primates: genetic and serological predictors for the existence of multiple unknown polyomaviruses within the human population. PLoS Pathog *9*.

Seguin, S.P., Ireland, A.W., Gupta, T., Wright, C.M., Miyata, Y., Wipf, P., Pipas, J.M., Gestwicki, J.E., and Brodsky, J.L. (2012). A screen for modulators of large T antigen's ATPase activity uncovers novel inhibitors of Simian Virus 40 and BK virus replication. Antiviral Res *96*, 70-81.

Seo, G.J., Fink, L.H., O'Hara, B., Atwood, W.J., and Sullivan, C.S. (2008). Evolutionarily conserved function of a viral microRNA. J Virol *82*, 9823-9828.

Shah, K., and Nathanson, N. (1976). Human exposure to SV40: review and comment. Am J Epidemiol *103*, 1-12.

Shah, K.V. (2007). SV40 and human cancer: a review of recent data. Int J Cancer *120*, 215-223.

Shayman, J.A., and Radin, N.S. (1991). Structure and function of renal glycosphingolipids. Am J Physiol *260*, F291-302.

Shen, P.S., Enderlein, D., Nelson, C.D.S.*, et al.* (2011). The structure of avian polyomavirus reveals variably sized capsids, non-conserved inter-capsomere interactions, and a possible location of the minor capsid protein VP4. Virology *411*, 142-152.

Shishido-Hara, Y., Ichinose, S., Higuchi, K., Hara, Y., and Yasui, K. (2004). Major and minor capsid proteins of human polyomavirus JC cooperatively accumulate to nuclear domain 10 for assembly into virions. J Virol *78*, 9890-9903.

Shuda, M., Feng, H., Kwun, H.J., Rosen, S.T., Gjoerup, O., Moore, P.S., and Chang, Y. (2008). T antigen mutations are a human tumor-specific signature for Merkel Cell polyomavirus. Proc Natl Acad Sci U S A *105*, 16272-16277.

Sievers, F., Wilm, A., Dineen, D.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology *7*.

Simon, K., Cameron, T., RUSHE, M., Caravella, J., and KAYNOR, G.C. (2013). JCV neutralizing antibodies. In WO2013142300 A3.

Spjut, S., Qian, W., Bauer, J., Storm, R., Frangsmyr, L., Stehle, T., Arnberg, N., and Elofsson, M. (2011). A potent trivalent sialic acid inhibitor of adenovirus type 37 infection of human corneal cells. Angew Chem Int Ed Engl *50*, 6519-6521.

Stehle, T., Gamblin, S.J., Yan, Y., and Harrison, S.C. (1996). The structure of Simian Virus 40 refined at 3.1 A resolution. Structure *4*, 165-182.

Stehle, T., and Harrison, S.C. (1996). Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. Structure *4*, 183-194.

Stehle, T., and Harrison, S.C. (1997). High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. EMBO J *16*, 5139-5148.

Stehle, T., and Khan, Z.M. (2014). Rules and exceptions: sialic acid variants and their role in determining viral tropism. J Virol *88*, 7696-7699.

Stehle, T., Yan, Y., Benjamin, T.L., and Harrison, S.C. (1994). Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. Nature *369*, 160-163.

Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. Journal of Applied Crystallography *41*, 641-643.

Stergiou, L., Bauer, M., Mair, W., Bausch-Fluck, D., Drayman, N., Wollscheid, B., Oppenheim, A., and Pelkmans, L. (2013). Integrin-mediated signaling induced by Simian Virus 40 leads to transient uncoupling of cortical actin and the plasma membrane. PLoS One *8*.

Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K., Palese, P., Paulson, J.C., and Wilson, I.A. (2006). Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. J Mol Biol *355*, 1143-1155.

Streuli, C.H., and Griffin, B.E. (1987). Myristic acid is coupled to a structural protein of polyoma virus and SV40. Nature *326*, 619-622.

Ströh, L.J., Maginnis, M.S., Blaum, B.S.*, et al.* (2015). The Greater Affinity of JC Polyomavirus Capsid for alpha2,6-Linked Lactoseries Tetrasaccharide c than for Other Sialylated Glycans Is a Major Determinant of Infectivity. J Virol *89*, 6364-6375.

Ströh, L.J., and Stehle, T. (2014). Glycan engagement by viruses: receptor switches and specificity. Annual Review of Virology *1*, 285-306.

Sturgill, E.R., Aoki, K., Lopez, P.H.H.*, et al.* (2012). Biosynthesis of the major brain gangliosides GD1a and GT1b. Glycobiology *22*, 1289-1301.

Sullivan, C.S., Grundhoff, A.T., Tevethia, S., Pipas, J.M., and Ganem, D. (2005). SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature *435*, 682-686.

Sullivan, C.S., Sung, C.K., Pack, C.D., Grundhoff, A., Lukacher, A.E., Benjamin, T.L., and Ganem, D. (2009). Murine polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection. Virology *387*, 157-167.

Suzuki, T., Orba, Y., Okada, Y., Sunden, Y., Kimura, T., Tanaka, S., Nagashima, K., Hall, W.W., and Sawa, H. (2010). The human polyoma JC virus agnoprotein acts as a viroporin. PLoS Pathog *6*.

Svennerholm, L., Bostrom, K., Fredman, P., Mansson, J.E., Rosengren, B., and Rynmark, B.M. (1989). Human brain gangliosides: developmental changes from early fetal stage to advanced age. Biochim Biophys Acta *1005*, 109-117.

Sweetser, M.T., Dawson, K.T., and Bozic, C. (2013). Manufacturer's response to case reports of PML. N Engl J Med *368*, 1659-1661.

Szklarczyk, O.M., Gonzalez-Segredo, N., Kukura, P., Oppenheim, A., Choquet, D., Sandoghdar, V., Helenius, A., Sbalzarini, I.F., and Ewers, H. (2013). Receptor concentration and diffusivity control multivalent binding of SV40 to membrane bilayers. PLoS Comput Biol *9*, e1003310.

Tan, C.S., Broge, T.A., Jr., Seung, E., Vrbanac, V., Viscidi, R., Gordon, J., Tager, A.M., and Koralnik, I.J. (2013). Detection of JC virus-specific immune responses in a novel humanized mouse model. PLoS One *8*, e64313.

Tan, C.S., and Koralnik, I.J. (2010). Progressive multifocal leukoencephalopathy and other disorders caused by JC virus: clinical features and pathogenesis. Lancet Neurol *9*, 425-437.

Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. Proc Natl Acad Sci U S A *100*, 12045-12050.

Taube, S., Perry, J.W., McGreevy, E., Yetming, K., Perkins, C., Henderson, K., and Wobus, C.E. (2012). Murine noroviruses bind glycolipid and glycoprotein attachment receptors in a strain-dependent manner. J Virol *86*, 5584-5593.

Tayyari, F., Marchant, D., Moraes, T.J., Duan, W.M., Mastrangelo, P., and Hegele, R.G. (2011). Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. Nature Medicine *17*, 1132-U1140.

Tiefenbrunn, T., Forli, S., Happer, M., Gonzalez, A., Tsai, Y.S., Soltis, M., Elder, J.H., Olson, A.J., and Stout, C.D. (2014). Crystallographic Fragment-Based Drug Discovery: Use of a Brominated Fragment Library Targeting HIV Protease. Chemical Biology & Drug Design *83*, 141-148.

Tornatore, C., Berger, J.R., Houff, S.A., Curfman, B., Meyers, K., Winfield, D., and Major, E.O. (1992). Detection of JC virus-DNA in peripheral lymphocytes from patients with and without progressive multifocal leukoencephalopathy. Annals of Neurology *31*, 454-462.

Trowbridge, P.W., and Frisque, R.J. (1995). Identification of three new JC virus proteins generated by alternative splicing of the early viral mRNA. J Neurovirol *1*, 195-206.

Tsai, B., Gilbert, J.M., Stehle, T., Lencer, W., Benjamin, T.L., and Rapoport, T.A. (2003). Gangliosides are receptors for murine polyoma virus and SV40. EMBO J *22*, 4346-4355.

Unterstab, G., Gosert, R., Leuenberger, D., Lorentz, P., Rinaldo, C.H., and Hirsch, H.H. (2010). The polyomavirus BK agnoprotein co-localizes with lipid droplets. Virology *399*, 322-331.

van der Meijden, E., Bialasiewicz, S., Rockett, R.J., Tozer, S.J., Sloots, T.P., and Feltkamp, M.C. (2013). Different serologic behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 polyomaviruses found on the skin. PLoS One *8*, e81078.

van der Meijden, E., Janssens, R.W., Lauber, C., Bouwes Bavinck, J.N., Gorbalenya, A.E., and Feltkamp, M.C. (2010). Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLoS Pathog *6*, e1001024.

van der Meijden, E., Kazem, S., Burgers, M.M., Janssens, R., Bouwes Bavinck, J.N., de Melker, H., and Feltkamp, M.C. (2011). Seroprevalence of trichodysplasia spinulosa-associated polyomavirus. Emerg Infect Dis *17*, 1355-1363.

Van Loy, T., Thys, K., Ryschkewitsch, C., Lagatie, O., Monaco, M.C., Major, E.O., Tritsmans, L., and Stuyver, L.J. (2015). JC virus quasispecies analysis reveals a complex viral population underlying progressive multifocal leukoencephalopathy and supports viral dissemination via the hematogenous route. J Virol *89*, 1340-1347.

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.

Varki, A. (2007). Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. Nature *446*, 1023-1029.

Varki, A., and Schauer, R. (2009). Sialic acids. In Essentials of Glycobiology, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi*, et al.*, eds. (Cold Spring Harbor (NY)).

Varki, N.M., and Varki, A. (2007). Diversity in cell surface sialic acid presentations: implications for biology and disease. Lab Invest *87*, 851-857.

Vilchez, R.A., and Butel, J.S. (2004). Emergent human pathogen Simian Virus 40 and its role in cancer. Clin Microbiol Rev *17*, 495-508, table of contents.

Viscidi, R.P., and Clayman, B. (2006). Serological cross reactivity between polyomavirus capsids. Polyomaviruses and Human Diseases *577*, 73-84.

Waga, S., and Stillman, B. (1994). Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. Nature *369*, 207-212.

Walczak, C.P., Ravindran, M.S., Inoue, T., and Tsai, B. (2014). A cytosolic chaperone complexes with dynamic membrane J-proteins and mobilizes a nonenveloped virus out of the endoplasmic reticulum. PLoS Pathog *10*, e1004007.

Walczak, C.P., and Tsai, B. (2011). A PDI family network acts distinctly and coordinately with ERp29 to facilitate polyomavirus infection. J Virol *85*, 2386-2396.

Walker, D.L., Padgett, B.L., ZuRhein, G.M., Albert, A.E., and Marsh, R.F. (1973). Human papovavirus (JC): induction of brain tumors in hamsters. Science *181*, 674-676.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Wattjes, M.P., Richert, N.D., Killestein, J., de Vos, M., Sanchez, E., Snaebjornsson, P., Cadavid, D., and Barkhof, F. (2013). The chameleon of neuroinflammation: magnetic resonance imaging characteristics of natalizumab-associated progressive multifocal leukoencephalopathy. Mult Scler *19*, 1826-1840.

Wielens, J., Headey, S.J., Rhodes, D.I.*, et al.* (2013). Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification? J Biomol Screen *18*, 147-159.

Winn, M.D., Ballard, C.C., Cowtan, K.D.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallographica Section D-Biological Crystallography *67*, 235-242.

Wuthrich, C., Cheng, Y.M., Joseph, J.T., Kesari, S., Beckwith, C., Stopa, E., Bell, J.E., and Koralnik, I.J. (2009a). Frequent infection of cerebellar granule cell neurons by polyomavirus JC in progressive multifocal leukoencephalopathy. J Neuropathol Exp Neurol *68*, 15-25.

Wuthrich, C., Dang, X., Westmoreland, S., McKay, J., Maheshwari, A., Anderson, M.P., Ropper, A.H., Viscidi, R.P., and Koralnik, I.J. (2009b). Fulminant JC virus encephalopathy with productive infection of cortical pyramidal neurons. Annals of Neurology *65*, 742-748.

Wuthrich, C., and Koralnik, I.J. (2012). Frequent infection of cortical neurons by JC virus in patients with progressive multifocal leukoencephalopathy. J Neuropathol Exp Neurol *71*, 54-65.

Wychowski, C., Benichou, D., and Girard, M. (1987). The intranuclear location of Simian Virus 40 polypeptide VP2 and polypeptide VP3 depends on a specific amino acid sequence. J Virol *61*, 3862-3869.

Xie, H., Lee, L., Caramuta, S., Hoog, A., Browaldh, N., Bjornhagen, V., Larsson, C., and Lui, W.O. (2014). MicroRNA expression patterns related to Merkel Cell polyomavirus infection in human Merkel Cell carcinoma. J Invest Dermatol *134*, 507-517.

Xiong, X., Coombs, P.J., Martin, S.R.*, et al.* (2013). Receptor binding by a ferret-transmissible H5 avian influenza virus. Nature *497*, 392-396.

Yamaguchi, H., Kobayashi, S., Maruyama, J., Sasaki, M., Takada, A., Kimura, T., Sawa, H., and Orba, Y. (2014). Role of the C-terminal region of vervet monkey polyomavirus 1 VP1 in virion formation. J Vet Med Sci *76*, 637-644.

Yamashita, M. (2010). Laninamivir and its prodrug, CS-8958: long-acting neuraminidase inhibitors for the treatment of influenza. Antivir Chem Chemother *21*, 71-84.

Yan, Y., Stehle, T., Liddington, R.C., Zhao, H., and Harrison, S.C. (1996). Structure determination of Simian Virus 40 and murine polyomavirus by a combination of 30-fold and 5-fold electron-density averaging. Structure *4*, 157-164.

Yogo, Y., Kitamura, T., Sugimoto, C., Ueki, T., Aso, Y., Hara, K., and Taguchi, F. (1990). Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J Virol *64*, 3139-3143.

Yu, R.K., Nakatani, Y., and Yanagisawa, M. (2009). The role of glycosphingolipid metabolism in the developing brain. Journal of Lipid Research *50 Suppl*, S440-445.

Yu, R.K., Tsai, Y.T., Ariga, T., and Yanagisawa, M. (2011a). Structures, biosynthesis, and functions of gangliosides--an overview. J Oleo Sci *60*, 537-544.

Yu, X., Coulson, B.S., Fleming, F.E., Dyason, J.C., von Itzstein, M., and Blanchard, H. (2011b). Novel structural insights into rotavirus recognition of ganglioside glycan receptors. J Mol Biol *413*, 929-939.

Zerrahn, J., Knippschild, U., Winkler, T., and Deppert, W. (1993). Independent expression of the transforming amino-terminal domain of SV40 large I antigen from an alternatively spliced third SV40 early mRNA. EMBO J *12*, 4739-4746.

Zila, V., Difato, F., Klimova, L., Huerfano, S., and Forstova, J. (2014). Involvement of microtubular network and its motors in productive endocytic trafficking of mouse polyomavirus. PLoS One *9*, e96922.

6. Appendix

6.1 Phylogenetic Analysis

Abbreviation	Full Name	NCBI Reference
AfricanElephantPyV	African elephant polyomavirus 1	NC_022519
ApanPyV1	Ateles paniscus polyomavirus 1	NC_019853
BatPyV-A1055	Bat polyomavirus isolate A1055	JQ958886
BatPyV-A504	Bat polyomavirus isolate A504	JQ958890
BatPyV-AT7	Bat polyomavirus isolate AT7	JQ958892
BatPyV-B0454	Bat polyomavirus isolate B0454	JQ958888
BatPyV-B1130	Bat polyomavirus isolate B1130	JQ958893
BatPyV-C1109	Bat polyomavirus isolate C1109	JQ958889
BatPyV-CardiodermaPyV1-KY336	Cardioderma polyomaviurs isolate KY336	NC_020067
BatPyV-ChaerephonPyV1-KY397	Chaerephon polyomavirus 1 isolate KY397	NC_020065
BatPyV-EidolonPyV1-KY270	Eidolon polyomavirus 1 isolate KY270	NC_020068
BatPyV-MiniopterusPyV-KY369	Miniopterus polyomavirus isolate KY369	NC_020069
BatPyV-OtomopsPyV1-KY156	Otomops polyomavirus 2 isolate KY156	NC_020066
BatPyV-OtomopsPyV1-KY157	Otomops polyomavirus 1 isolate KY157	NC_020071
BatPyV-PteronotusPyV-GTM203	Pteronotus polyomavirus isolate GTM203	NC_020070
BatPyV-R104	Bat polyomavirus isolate R104	JQ958887
BatPyV-R266	Bat polyomavirus isolate R266	JQ958891
BatPyV-TbPyV1	Tadarida brasiliensis polyomavirus 1	AJA41149.1
BatPyV-TbPyV2	Tadarida brasiliensis polyomavirus 2	AJA41153.1
BbPyV	Butcherbird polyomavirus isolate AWH19840	NC_023008
BFDPyV	Budgerigar fledgling disease polyomavirus	AY672646
BKPyV	BK Polyomavirus	NC_001538
BoPyV	Bovine polyomavirus	NC_001442
CalbPyV1	Cebus albifrons polyomavirus 1	NC_019854
CaPyV	Canary polyomavirus	NC_017085
CeryPyV1	Cercopithecus erythrotis polyomavirus 1	JX159985
ChPyV-Azzie	Chimpanzee polyomavirus complete genome	FR692336
ChPyV-Bob	Chimpanzee polyomavirus complete genome,	FR692334

	isolate Bob	
ChPyV-Tanu	Chimpanzee polyomavirus complete genome, isolate Tanu	FR692335
CrowPyV	Crow polyomavirus	NC_007922
CSLPyV	California sea lion polyomavirus 1	YP_003429322.1
DolphinPyV1	Dolphin polyomavirus 1 isolate DPyV-1/Trachea/2010	KC594077
EPyV	Equine polyomavirus	NC_017982
FiPyV	Finch polyomavirus	NC_007923
GgorgPyV1	Gorilla gorilla gorilla polyomavirus 1 isolate 5766	HQ385752
GhPyV	Goose hemorrhagic polyomavirus isolate Toulouse Goose 2000	HQ681903
HaPyV	Hamster polyomavirus	YP_009111410.1
HPyV12	Human Polyomavirus 12	NC_020106
HPyV6	Human Polyomavirus 6	NC_014406
HPyV7	Human Polyomavirus 7	NC_014407
HPyV9	Human Polyomavirus 9	NC_015150
JCPyV	JC Polyomavirus	NC_001699
KIPyV	Karolinska Institute Polyomavirus	NC_009238
LPyV	Monkey B-lymphotropic papovavirus	M30540
MCPyV	Merkel Cell Polyomavirus	NC_010277
MfasPyV1	Macaca fascicularis polyomavirus 1	NC_019851
MPtPyV	Murine pneumotropic virus strain Kilham clone pKV (37-1)	EF186666
MPyV-PTA	Murine polyomavirus strain PTA	U27812
MWPyV	Malawi Polyomavirus	NC_018102
MyoPyV	Myotis polyomavirus VM-2008	NC_011310
NJPyV	New Jersey Polyomavirus	NC_024118
OraPyV-Bo	Orangutan polyomavirus complete genome, isolate Bo	NC_013439
OraPyV-PI	Orangutan polyomavirus complete genome, isolate Pi	FN356901
PrufPyV1	Piliocolobus rufomitratus polyomavirus 1	JX159984
PtrosPyV2	Pan troglodytes schweinfurthii polyomavirus 2	NC_019858
PtrovPyV1a	Pan troglodytes verus polyomavirus 1a isolate 6444	HQ385746
PtrovPyV1b	Pan troglodytes verus polyomavirus 1b isolate 6520	HQ385747
PtrovPyV2a	Pan troglodytes verus polyomavirus 2a isolate 6512	HQ385748

PtrovPyV2c	Pan troglodytes verus polyomavirus 2c isolate 5924	HQ385749
PtrovPyV3	Pan troglodytes verus polyomavirus 3	NC_019855
PtrovPyV4	Pan troglodytes verus polyomavirus 4	NC_019856
PtrovPyV5	Pan troglodytes verus polyomavirus 5	NC_019857
SA12	Simian virus 12	DQ435829
SqPyV	Squirrel monkey polyomavirus	NC_009951
SsciPyV1	Saimiri sciureus polyomavirus 1	JX159989
STLPyV	Saint Louis Polyomavirus	NC_020890
SV40	Simian virus 40	NC_001669
TSPyV	*Trichodysplasia spinulosa*- associated Polyomavirus	NC_014361
VmPyV1	Vervet monkey polyomavirus 1	NC_019844
VmPyV2	Vervet monkey polyomavirus 2	AB767299
VmPyV3	Vervet monkey polyomavirus 3	AB767297
WUPyV	Washington University Polyomavirus	NC_009539
YbPyV1	Yellow baboon polyomavirus 1	AB767294
YbPyV2	Yellow baboon polyomavirus 2	AB767295

6.2 Author Contributions

Ongoing research

LS designed, performed, and analyzed structural and functional experiments. LS and TS provided conception of the project. Bernhard Ehlers (RKI, Berlin) provided the DNA for HPyV12 VP1 and Mischra Nischay (Columbia University) provided NJPyV VP1 cDNA. Florian Wedekink (IFIB, Tübingen) cloned the expression construct for NJPyV VP1, purified, and crystallized the protein under supervision of LS. Bärbel Blaum (IFIB, Tübingen) performed and analyzed NMR experiments with NJPyV and HPyV12 VP1. Christoph Rademacher and Jonas Aretz (Max Planck Institute of Colloids and Interfaces, Potsdam) performed fragment-based screening for JCPyV by NMR. LS planned and analyzed soaking experiments with JCPyV VP1 crystals. Jan-Marten Schmidt (IFIB, Tübingen) performed several soaking experiments, processed, and analyzed three data sets under supervision of LS.

Ströh L, Gee G, Blaum B, Dugan A, Feltkamp M, Atwood W, Stehle T. *Trichodysplasia spinulosa*-associated polyomavirus employs a displaced binding site on VP1 to engage sialylated glycolipids.

LS provided conception of the project, designed, performed, and analyzed structural and functional experiments. GG purified pseudoviruses, performed infection studies with the mutant TSPyV pseudoviruses, and designed and carried out the A549 infection studies. BB performed NMR. AS designed and cloned the pseudovirus constructs. MF provided TSPyV cDNA. LS wrote the bulk of the manuscript together with TS. TS provided conception of the project. WA and TS provided intellectual contributions in experimental design and data interpretation, and contributed to writing the manuscript.

Ströh L*, Maginnis M*, Blaum B, Nelson C, Neu U, Gee G, O'Hara B, Motamedi N, DiMaio D, Atwood W, Stehle T. Increased affinity of JC polyomavirus capsid for LSTc over other sialylated glycans is a major determinant of infectivity.

*MS and LS contributed equally to this work. LS provided conception of the project, designed and performed experiments, and wrote the manuscript together with MM. In further detail, LS carried out cloning, expression and protein purification, crystallographic studies and structural analysis. MM performed all functional cell-based experiments. UN and BB provided intellectual input, and BB performed NMR. CN, GG, and NM provided experimental advice and reagents. DD provided intellectual input and reagents. WA and TS provided conception of the project, intellectual contributions in experimental design and data interpretation, and contributed to writing the manuscript.

Nelson C, **Ströh L**, Gee G, O'Hara B, Stehle T, Atwood W. Modulation of a pore in the capsid of JC polyomavirus reduces infectivity and prevents exposure of the minor capsid proteins, 2015

LS carried out cloning, expression and protein purification, crystallographic studies, and structural analysis. LS did the homology modeling and contributed to experimental design, data interpretation, and writing of the manuscript.

Ströh L, Neu U, Blaum B, Buch M, Garcea R, Stehle T. Structure analysis of the major capsid proteins of the human polyomaviruses 6 and 7 reveals an obstructed sialic acid binding site, 2014

LS designed and carried out expression, protein purification, flow cytometry experiments, crystallographic studies, and structure analysis. LS wrote the manuscript together with TS. BB performed the NMR experiments. MB cloned the expression constructs under the supervision of UN. RG provided the cDNA of HPyV6 and HPyV7 VP1.

Ströh L and Stehle T. Glycan engagement by viruses: receptor switches and specificity, 2014

LS and TS provided conceptions for the review and wrote the manuscript together. LS prepared the figures.

Maginnis M*, **Ströh L***, Gee, G, O'Hara B, Derdowski A, Stehle T, Atwood W. Progressive multifocal leukoencephalopathy-associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding, 2013

*MS and LS contributed equally to this work. LS performed protein purification, crystallization, and structural analysis. MM carried out functional studies together with GG and BH. LS and MM designed experiments, conducted data interpretation and wrote the manuscript together. GG, BH and AA provided reagents. WA and TS were involved in the project conception, provided intellectual contributions in data interpretation, and contributed to writing the manuscript.

Neu U, Allen S, Blaum B, Liu Y, Frank M, Palma A, **Ströh L**, Feizi T, Peters T, Atwood W, Stehle T. A Structure-guided mutation in the major capsid protein retargets BK polyomavirus, 2013

LS performed model building of the native and complex structures, prepared Table 1 for the manuscript and deposited final structural models and structure factor amplitudes to the PDB database.

Forouzan D, Ammelburg M, Hobel C, **Ströh L**, Sessler N, Martin J, Lupas A. The archaeal proteasome is regulated by a network of AAA ATPases, 2012

LS designed, performed, and analyzed SPR experiments.

Neu U, Maginnis M, Palma A, **Ströh L**, Nelson C, Feizi T, Atwood A, Stehle T. Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif, 2010

LS carried out expression, protein purification and crystallization for the VP1-LSTc glycan complex structure. UN and LS performed crystal soaking experiments together. Data collection, data processing, structure determination, and model building were carried out by UN and/or LS under the guidance of UN.

6.3 Acknowledgments

- Prof. Dr. Thilo Stehle danke ich besonders für die kompetente und stets optimistische Betreuung meiner Promotion, für die Förderung meiner wissenschaftlichen Weiterbildung und das Ermöglichen meiner persönlichen und wissenschaftlichen Freiheit.
- I thank Prof Dr. Walter Atwood for his helpful questions, for his scientific advice and support, and the great opportunity to join his group at Brown University for five month in order to work with “real” cells and viruses.
- I thank Melissa Maginnis for her enthusiasm about our projects, all inspiring discussions and conversations we had in “real life”, on skype, and per email. It was a great pleasure and fun working with you!
- Allen Mitgliedern des AK Stehles insbesondere Bärbel, Georg, Yinglan, Micha, Melanie, Volker, MicBuc, Manuel, Felix, Alex, Irmi und Niki danke ich für ihre Hilfsbereitschaft, die vielen produktiven Diskussionen, die gemeinsamen Obst-, Kaffee- und Kuchenpausen und die gute Arbeitsatmosphäre. Ich danke besonders Ulla für die gute und stets motivierende Zusammenarbeit.
- I thank Gretchen, Bethany, Aaron, Stacy, Christian, Stephen and Sheila for answering all my questions, their help inside and outside the Atwood lab and the great time I spent in Rhode Island and at the Brown University.
- Zudem danke ich den ehemaligen Labor- und Institutsmitgliedern JoBa, Sebastian, Dirk, Ulla, Kerstin, Karo, Riki, Denise, Yvonne und Hanna für ihre ständige Hilfs- und Dikussionsbereitschaft am Institut sowie für die gemeinsamen Aktivitäten welche die nötige Abwechslung in den Laboralltag und nach Tübingen gebracht haben.
- I thank Prof. Dr. Robert Garcea, Prof. Dr. Daniel DiMaio, Dr. Bernhard Ehlers and Dr. Mariet Feltkamp for the collaboration, helpful questions and discussions.
- Außerdem danke ich Anna-Lisa, Markus, Jan-Marten, Florian, Anne, Viktoria und Steffi für die gute Zusammenarbeit als HiWi, Arbeitsgruppenpraktikanten oder Bachelorstudenten.
- Ich danke Dr. Anja Drescher für die lehrreichen drei Wochen mit dem Biacore T-200 und Prof. Dr. Niklas Arnberg sowie Lars Frängsmyr für die Möglichkeit das T-100 in Umea zu nutzen.
- Zusätzlich möchte ich mich bei Solli, Fridi und Anni, Sonja, Ruth, Christopher, Rike, Tina, Anna, Tobi, Steff, Fabi, Adam und Basti für die Freundschaft bedanken. Ich freue mich sehr, dass wir die letzten Jahre in Kontakt geblieben sind. Ihr habt mir oft sehr geholfen Tübungen hinter mir zu lassen und einfach mal abzuschalten.
- Meinen Eltern und Jon danke ich für die immerwährende Unterstützung meiner Doktorarbeit in “weit weg vom Land der Horizonte”.
- Alex, ich bin so froh, dass es dich gibt!

PLOS | PATHOGENS

RESEARCH ARTICLE

Trichodysplasia spinulosa-Associated Polyomavirus Uses a Displaced Binding Site on VP1 to Engage Sialylated Glycolipids

Luisa J. Ströh[1], Gretchen V. Gee[2], Bärbel S. Blaum[1], Aisling S. Dugan[3], Mariet C. W. Feltkamp[4], Walter J. Atwood[2]*, Thilo Stehle[1,5]*

1 Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany, **2** Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, United States of America, **3** Department of Natural Sciences, Assumption College, Worcester, Massachusetts, United States of America, **4** Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands, **5** Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America

* Walter_Atwood@brown.edu (WJA); thilo.stehle@uni-tuebingen.de (TS)

Citation: Ströh LJ, Gee GV, Blaum BS, Dugan AS, Feltkamp MCW, Atwood WJ, et al. (2015) *Trichodysplasia spinulosa*-Associated Polyomavirus Uses a Displaced Binding Site on VP1 to Engage Sialylated Glycolipids. PLoS Pathog 11(8): e1005112. doi:10.1371/journal.ppat.1005112

Editor: James Pipas, University of Pittsburgh, UNITED STATES

Received: April 7, 2015

Accepted: July 27, 2015

Published: August 24, 2015

Data Availability Statement: All relevant data are included in the paper and its Supporting Information files except for coordinates and structure factor amplitudes which are deposited with the RCSB data bank under PDB accession numbers 4U5Z (unliganded TSPyV VP1), 4U60 (TSPyV VP1-GM1 glycan), 4U61 (TSPyV VP1-6'SL) and 4U62 (TSPyV VP1-3'SL).

Funding: This work was supported by contract research "Glykobiologie / Glykomik" of the Baden-Württemberg Foundation (TS). The funders had no

Abstract

Trichodysplasia spinulosa-associated Polyomavirus (TSPyV) was isolated from a patient suffering from *trichodysplasia spinulosa*, a skin disease that can appear in severely immunocompromised patients. While TSPyV is one of the five members of the polyomavirus family that are directly linked to a human disease, details about molecular recognition events, the viral entry pathway, and intracellular trafficking events during TSPyV infection remain unknown. Here we have used a structure-function approach to shed light on the first steps of TSPyV infection. We established by cell binding and pseudovirus infection studies that TSPyV interacts with sialic acids during attachment and/or entry. Subsequently, we solved high-resolution X-ray structures of the major capsid protein VP1 of TSPyV in complex with three different glycans, the branched GM1 glycan, and the linear trisaccharides α2,3- and α2,6-sialyllactose. The terminal sialic acid of all three glycans is engaged in a unique binding site on TSPyV VP1, which is positioned about 18 Å from established sialic acid binding sites of other polyomaviruses. Structure-based mutagenesis of sialic acid-binding residues leads to reduction in cell attachment and pseudovirus infection, demonstrating the physiological relevance of the TSPyV VP1-glycan interaction. Furthermore, treatments of cells with inhibitors of N-, O-linked glycosylation, and glycosphingolipid synthesis suggest that glycolipids play an important role during TSPyV infection. Our findings elucidate the first molecular recognition events of cellular infection with TSPyV and demonstrate that receptor recognition by polyomaviruses is highly variable not only in interactions with sialic acid itself, but also in the location of the binding site.

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.

Author Summary

Viruses engage receptors on their host cell to initiate entry and infection. Members within a virus family are known to target different tissues and hosts and exploit different pathogenic mechanisms due to critical changes in receptor specificity. The human *Trichodysplasia spinulosa*-associated Polyomavirus (TSPyV) is known to cause a rare skin disease in immunocompromised individuals. The pathogenic mechanism includes hyperproliferation of inner root sheath cells, but molecular determinants underlying the infection and the associated disease remain unknown. Here we applied a structural and functional approach to investigate the recognition events during early infection steps of the virus. We found that TSPyV engages sialic acid receptors but employs a novel binding site on the capsid that is shifted in comparison to other structurally characterized polyomaviruses. Cell-based studies demonstrate the relevance of the observed interaction for attachment and infection and suggest that glycolipids, rather than N- and O-linked glycoproteins, are important for infection. Our findings demonstrate exemplarily that receptor recognition by (polyoma-) viruses is highly variable not only in interactions with sialic acids, but also in the location of the binding site on the capsid, providing insights about structural determinants of receptor and host specificity and evolution of these viruses.

Introduction

Trichodysplasia spinulosa-associated Polyomavirus (TSPyV) was discovered in 2010 in facial skin samples from a patient with the skin disease *trichodysplasia spinulosa* (TS) [1]. TS is characterized by the development of folliculocentric papules and keratin spines, predominantly localized to the face and less frequently on the trunk and limbs of immunocompromised individuals [1, 2–4, 5]. The pathogenic mechanism of TSPyV during the symptomatic infection includes uncontrolled hyperproliferation of inner root sheath (IRS) cells, but molecular determinants underlying TSPyV infection and disease remain largely unknown [2,6,7]. Electron microscopic studies have confirmed the presence of icosahedral viral particles in affected hair follicles [2,4,8,9]. However, high viral loads have been detected only in TS patients, whereas samples from the skin, plucked eyebrows, serum/plasma, and urine of immunocompetent and-compromised individuals were mostly negative for TSPyV DNA [1,6,10]. In contrast, seroprevalence values of about 70% within the human population suggest that initial infections with TSPyV occur during childhood [11–13], and thus persistent infections at undetectable levels or in undiscovered latent extracutaneous reservoirs are very likely [10,14]. The detection of TSPyV DNA in tonsillar samples from healthy individuals indicates that the virus infects lymphoid tissue establishing a persistent infection [10,14]. Viral shedding and spreading from this persistent site may then be crucial for transmission and reactivation during immunosuppression [14].

Among the recently discovered human members of the growing polyomavirus family, TSPyV and the carcinogenic Merkel Cell Polyomavirus (MCPyV) have gained particular attention due to clear links to a human disease or human cancer, respectively [15,16]. Furthermore, the skin-tropic Human Polyomavirus 7 (HPyV7) has recently been associated with thymic epithelial tumors [17]. While these three viruses share skin tropism, characteristics of the infection and pathogenicity seem to differ. For example, MCPyV is clonally integrated in the host cell genome in the majority of the neuroendocrine Merkel cell carcinomas (MCC) [15], but there is no evidence for genomic integration of TSPyV to date.

TSPyV has a 5232-nucleotide dsDNA genome encapsulated in its non-enveloped icosahedral capsid made up of the proteins VP1, VP2 and VP3 [1]. X-ray crystallographic studies of the pentameric major capsid proteins (VP1s) from several polyomaviruses have revealed a conserved jell-roll fold topology [18–27]. On the outer surface of the virion, structurally distinct loops connect the β-sheet core, and these loops are chiefly responsible for viral antigenicity. They also form a virus-host interaction platform that contributes to host range, cell tropism, viral spread, and pathogenicity. Although sialylated glycans are functional receptors for several polyomaviruses, the role and importance of these glycans for infectious entry remain unknown for other more recently discovered family members. The engagement of non-sialylated receptor types has been suggested in several cases [22,27].

Sialic acids are commonly connected by either α2,3- or α2,6-linkages to a galactose (Gal), by an α2,6-linkage to N-acetylgalactosamine (GalNAc) or via α2,8-linkages to one another. They are abundantly expressed on N- or O-linked glycoproteins as well as on gangliosides, and several chemical modifications are known [28]. The predominant sialic acid species in humans is α-5-N-acetylneuraminic acid (Neu5Ac), which is a central building block of cell surface receptors for many human viruses [29,30]. In contrast, α-5-N-glycolylneuraminic acid (Neu5Gc), the predominant type of sialic acid in many other mammals, cannot be synthesized by humans due to a species-specific inactivating deletion in the gene encoding the hydroxylase that converts CMP-Neu5Ac to CMP-Neu5Gc [31,32]. However, Neu5Gc can be metabolically incorporated into human tissues from dietary sources [33] and its role in receptor engagement by viruses and in defining viral tropism is only beginning to emerge [34].

The same binding region at the surface of VP1 is employed for the interaction with terminal sialic acids in all structurally investigated sialic acid-engaging polyomaviruses so far, but amino acid differences in or near the core binding pocket can modulate the recognition of specific sialylated glycan receptors or receptor motifs [30]. Subtle VP1 amino acid changes in the binding pocket can have a critical impact on infection and viral pathogenicity [35,36,24,26]. This is illustrated by a single amino residue mutation in the binding pocket of the human BK Polyomavirus (BKPyV), which allows for a switch in the ganglioside receptor specificity [24]. Affinity is also critical, as the closely related JC Polyomavirus (JCPyV) binds several ganglioside motifs, including GM1 and GD1b, but the increased affinity for the α2,6-linked lactoseries tetrasaccharide c (LSTc) is crucial for its function as a receptor [21,37]. Murine Polyomavirus (MPyV), MCPyV, B-lymphotropic Polyomavirus (LPyV), and Human Polyomavirus 9 (HPyV9) recognize sialic acids in distinct orientations within the conserved location of the binding pocket on VP1 by applying different interaction strategies [19,23,25,26]. In contrast, Human Polyomaviruses 6 (HPyV6) and 7 possess elongated VP1 surface loops that obstruct the Neu5Ac-binding region, indicating that these viruses utilize non-sialylated receptors [27].

To enhance our understanding of the cell and host tropism of TSPyV, and ultimately its pathogenicity, we aimed to shed light on molecular determinants and principles during the early steps of infection. Following cell attachment studies, we determined high-resolution X-ray structures of TSPyV VP1 pentamers alone and in complex with glycans bearing α2,3- or α2,6-linked terminal Neu5Ac. TSPyV engages sialic acid receptors but the virus uses a novel binding site on VP1 that is shifted in comparison to all other structurally characterized polyomaviruses so far. Cell-based studies demonstrate the relevance of this binding site for cell attachment and infectivity, and also suggest that glycolipids rather than N- and O-linked glycoproteins play an important role during TSPyV infection. In conclusion, our findings highlight the complexity and plasticity of virus-glycan receptor interactions, and provide clues about the determinants of host specificity and evolution of TSPyV.

Results

Sialic acids promote TSPyV VP1 cell attachment

To determine if TSPyV interacts with sialic acids on cell surfaces, we first examined binding of unassembled TSPyV VP1 pentamers to cultured human cells by flow cytometry (Fig 1A and S1 Fig). For these experiments, TSPyV VP1 was recombinantly expressed in *E. coli*. The VP1 protein forms the characteristic homopentamer but cannot assemble further into viral capsids due to N- and C-terminal truncations of the expression construct. Cells were mock treated or incubated with *Clostridium perfringens* neuraminidase type V to remove terminal α2,3-, α2,6-, and α2,8-linked sialic acids from the cell surface prior to the flow cytometry experiment. The binding signal of TSPyV VP1 pentamers to enzyme-treated cells is significantly reduced compared to binding to mock-treated cells (Fig 1A and S1 Fig). BKPyV VP1 pentamers were used as positive control for neuraminidase-sensitive cell attachment (S1 Fig).

Sialic acids are important for TSPyV pseudovirus infection

Next, we investigated the role of sialic acids in TSPyV infection. To our knowledge, a cell culture system for growing TSPyV is not available. However, virus-like particles (VLPs) or reporter vector particles, known also as pseudoviruses (PsV), have been generated for other polyomaviruses and are useful model systems for infection studies reflecting the viral tropisms [38–40,36]. We therefore developed a TSPyV pseudovirus system (TSPsV), which uses the TSPyV structural proteins expressed in the producer cells HEK293TT to package a reporter plasmid coding for secreted luciferase that can be detected by relative luminescence units (RLU) for quantification of PsV infection [39,40]. A mock control for PsV infections was generated by transforming the reporter plasmid together with a control plasmid instead of plasmids coding for structural proteins into producer cells. First, multiple common cell lines were screened for efficient transduction of the reporter plasmid to identify suitable cell lines for TSPsV infection studies (S2 Fig). Significant transduction in comparison to the mock transfection control was seen for HEK293 cells and the human glial cell line SVGA, but not for HeLa, monkey kidney (Vero) and Chinese Hamster Ovary (CHO) cells. TSPsV cell attachment and transduction in SVGA and HEK293 is sensitive to pretreatment with neuraminidase (Fig 1B–1D). Together with the cell binding analysis, the PsV infection studies suggest that TSPyV engages sialic acids on the cell surface during viral attachment and prior to infectious entry.

Crystal structure of TSPyV VP1 pentamers in complex with α2,3- and α2,6-linked sialylated glycan motifs

To establish a platform for understanding the specificity of the interaction with sialic acid, we solved crystal structures of TSPyV VP1 alone and in complex with three different sialylated compounds, the branched α2,3-linked GM1 pentasaccharide and the linear glycans α2,3- and α2,6-sialyllactose (3'SL and 6'SL, respectively) (Table 1 and Fig 2). TSPyV VP1 pentamers were co-crystallized in the presence of either 10 mM GM1 glycan or 10 mM 6'SL, while preformed native TSPyV VP1 pentamer crystals were derivatized by soaking them in a 10 mM 3'SL solution for complex formation. As is typical for polyomaviruses [18–27], the TSPyV VP1 monomer adopts the iconic jelly-roll fold with a conserved core comprising two β-sheets that are formed by strands B, I, D, G and C, H, E, F, respectively (Fig 2A). Extensive loops connect the strands on the top of the VP1 pentamer and are named BC-, DE-, HI- and EF-loops according to the strands they connect. The long BC-loop folds into two different directions on top of the protein and is divided for clarity into BC1-loop, BC-linker and BC2-loop [18,22]. In the unbiased difference (F_{obs}-F_{calc}) electron density map of the glycan complex structures, we observed

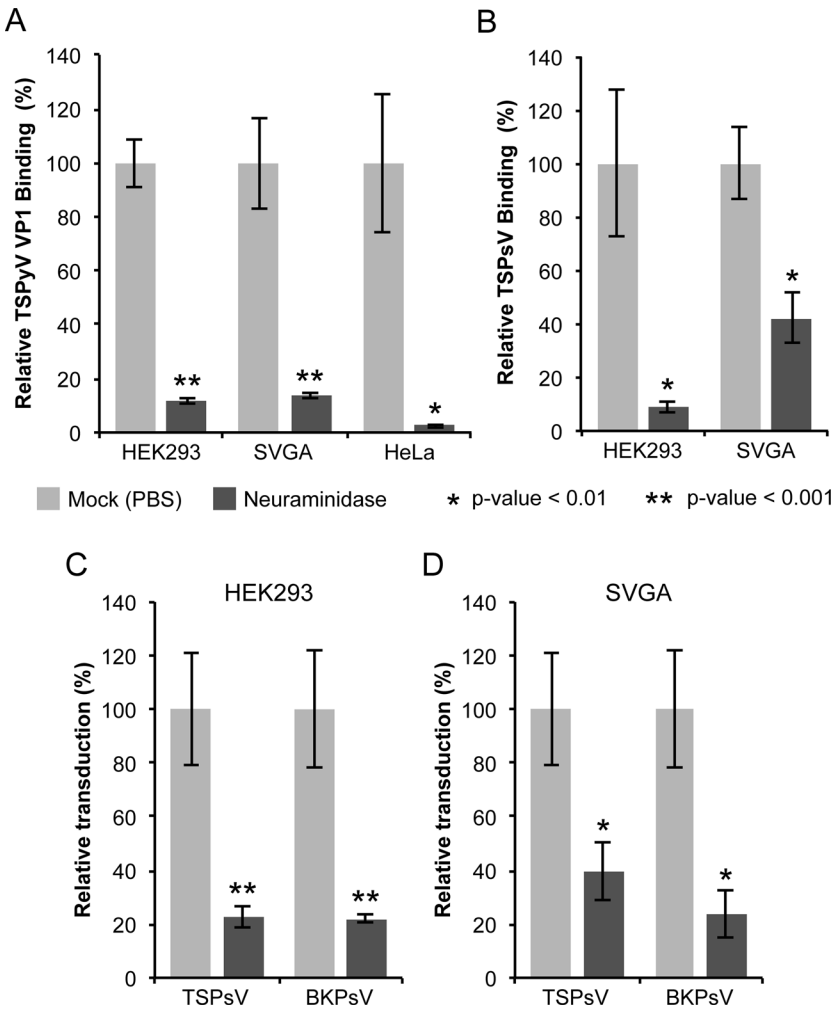


Fig 1. Cell surface glycans featuring terminal sialic acids promote TSPyV attachment and pseudovirus infection. (A) Binding of Alexa Fluor 488-conjugated TSPyV VP1 pentamers to mock (PBS) or with 1 U/ml *Clostridium perfringens* type V neuraminidase pre-treated cells was analysed by flow cytometry. 30,000 gated events were measured for each sample. The average from three independent experiments is shown, and error bars indicate the standard deviation. Data was standardized to signals of mock treated cells alone. S1 Fig shows non-standardized data of individual experimental replicates. (B) Attachment of Alexa Fluor 488-conjugated TSPsV to mock (PBS) or with 1 U/ml *Clostridium perfringens* type V neuraminidase pre-treated HEK293 and SVGA cells. The analysis was carried out by flow cytometry with 10,000 events measured for each sample. The average from three replicates is shown, and error bars indicate the standard deviation. (C-D) TSPyV pseudovirus (TSPsV) infections of HEK293 and SVGA cells (mock-treated cells and cells incubated with *Clostridium perfringens* neuraminidase for 30 min prior to PsV infection). BK Polyomavirus pseudovirus (BKPsV) was used as positive control for a neuraminidase-sensitive infection. The efficiency of the PsV infection was measured by transduction of the reporter plasmid phGluc coding for a secreted form of *Gaussia* luciferase 72 h post infection by detection of the secreted luciferase (measured in relative light units, RLUs using the BioLux *Gaussia* Luciferase Assay Kit). PsV experiments were done in quadruplicate and repeated three times and the average relative transduction is shown. Error bars indicate standard deviations and statistic analysis was performed using the two-tailed unpaired t test. The data was standardized to mock PsV transfections. The mock PsV control was generated by harvesting HEK293TT cells transfected with control plasmid instead of the capsid expression plasmids and then purifying according to the PsV purification protocol to measure background signal and non-specific transfer of luciferase to infected cells.

doi:10.1371/journal.ppat.1005112.g001

well-defined electron density for the terminal Neu5Ac residues of 3'SL, 6'SL and the GM1 glycan. The highest occupancy was obtained for GM1 in all five binding sites of the VP1 pentamer

Table 1. Crystallographic data collection and structure refinement statistics.

	TSPyV VP1	TSPyV VP1—GM1 glycan	TSPyV VP1–6'SL	TSPyV VP1–3'SL
PDB code	4U5Z	4U60	4U61	4U62
Data collection				
Space group	$P2_1$	$P2_12_12$	$P2_12_12_1$	$P2_1$
Unit cell				
a, b, c [Å]	64.10, 153.64, 143.98	144.99, 152.05, 67.97	139.56, 146.35, 151.54	66.24, 152.84, 147.24
β [°]	91.83	90.00	90.00	92.34
Resolution [Å]	30–2.10 (2.16–2.10)	50–1.50 (1.54–1.50)	30–1.65 (1.69–1.65)	40–1.55 (1.59–1.55)
Unique reflections	158,871 (11,115)	238,315 (17,410)	368,928 (27,122)	412,698 (29,151)
Total reflections	749,166 (48,901)	1,312,931 (92,629)	3,423,638 (224,558)	1,584,158 (106,332)
R_{meas} [%]	9.0 (94.9)	8.1 (88.2)	9.5 (98.6)	6.2 (70.8)
I/σI	12.0 (1.7)	16.0 (2.0)	15.8 (2.3)	13.4 (2.2)
$CC_{1/2}$ [%]	99.8 (65.2)	99.9 (66.3)	99.9 (75.3)	99.9 (76.3)
Completeness [%]	98.3 (92.8)	99.4 (99.0)	99.7 (100.0)	97.8 (93.5)
Wilson B factors [$Å^2$]	44.1	21.5	27.3	29.2
Refinement				
R_{work}/R_{free} [%]	20.5/24.9	15.8/18.0	16.5/18.5	15.5/18.0
No. of atoms				
Protein	20,375	10,611	20,747	20,892
Glycans	-	251	146	63
Water	551	1,637	2,267	2,355
Other solvent molecules*	-	74	24	46
Average B factors [$Å^2$]				
Protein	43.0	17.3	23.5	27.5
Glycans	-	24.2	35.5	38.4
Water	38.5	26.9	30.6	32.4
Other solvent molecules*	-	33.2	34.0	37.0
RMSD				
Bond length [Å]	0.008	0.009	0.008	0.009
Bond angles [°]	1.273	1.419	1.309	1.356
Ramachandran plot				
Favored [%]	96.0	95.7	97.0	97.1
Allowed [%]	4.0	4.3	3.0	2.9

Values for the highest resolution shell are given in parentheses. R_{free} was calculated with 5% of the data. RMSD, root-mean-square deviation. The Ramachandran plot statistics were calculated with MolProbity [69].

*Other solvent molecules are glycerol and 1,2-ethandiol.

doi:10.1371/journal.ppat.1005112.t001

in the asymmetric unit, and so additional glycan residues besides Neu5Ac could be built unambiguously into well-defined electron density (Fig 2B). Nevertheless, three of the five binding sites were not considered for further structural analysis because the ligand participates in crystal contacts. In the remaining two sites of the VP1 pentamer, the Neu5Ac residue and the GalNAcβ1-3(Neu5Acα2–3)Gal trisaccharide (Fig 2B and 2C), respectively, are well-defined by electron density and distant from crystal contacts. These two binding sites are therefore discussed below in detail.

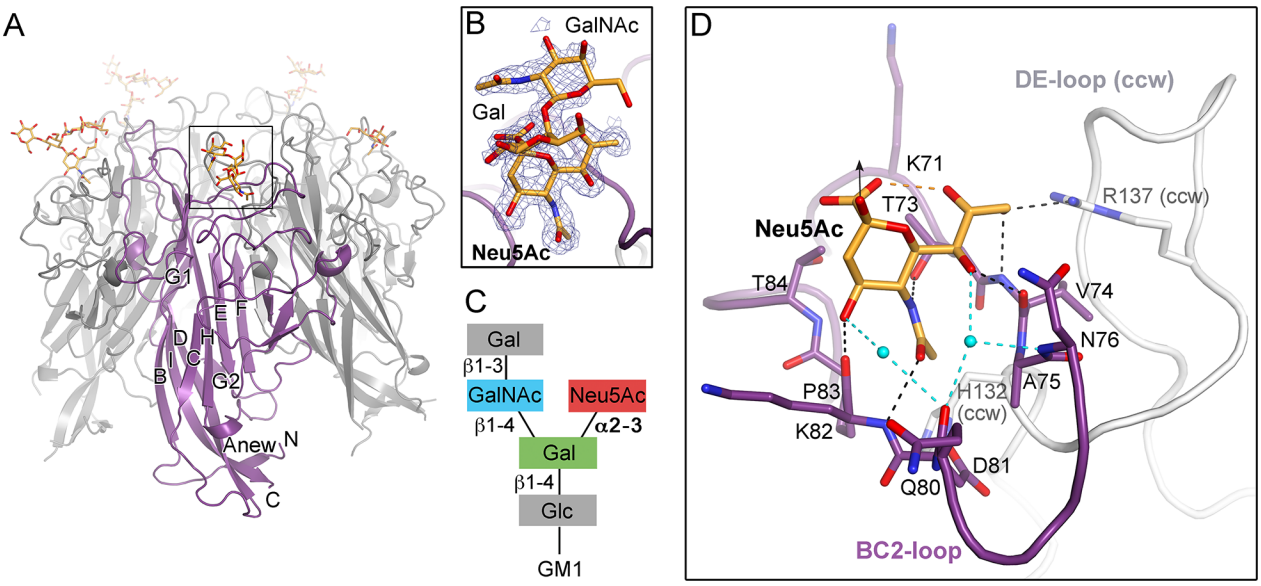


Fig 2. TSPyV VP1 specifically binds to terminal Neu5Ac. (A) Crystal structure of the unassembled TSPyV VP1 pentamer in complex with the GM1 glycan. The TSPyV VP1 pentamer is shown in cartoon representation with one monomer depicted in purple. Glycan residues are shown in stick representations and glycan atoms are coloured according to the atom type, with oxygen in red, nitrogen in blue, and carbon in orange. (B) Close-up view of the Neu5Ac binding site. The simulated annealing omit difference map contoured at 3.0 σ is shown with a radius of 2.0 Å around glycan residues Neu5Ac-α2,3-Gal-β1,4GalNAc of the GM1 glycan. (C) Schematic representation of the GM1 glycan. Glycan rings built into electron density in panel B are highlighted in colours. (D) Interaction between TSPyV VP1 and the terminal Neu5Ac residue of the GM1 glycan in a binding site that is formed by the BC2-loop. Side chains and backbone atoms of VP1 interacting with the glycan are shown in stick representation, and atoms are coloured as in panel A. Water molecules are shown as spheres, hydrogen bonds between VP1 and Neu5Ac are drawn as black dashed lines, water-mediated contacts between VP1 and Neu5Ac are depicted in cyan, and intramolecular Neu5Ac hydrogen bonds are shown in orange.

doi:10.1371/journal.ppat.1005112.g002

Specific recognition of terminal sialic acids

The N-acetyl group of Neu5Ac inserts into a shallow cavity that is built by the BC2-loop on top of a VP1 monomer, thereby forming hydrogen bonds via its NH-group towards the hydroxyl side chain of threonine 73 and via its oxygen with the backbone amino group of lysine 82 (Fig 2D). The side chain of the arginine residue 137 from the DE-loop of the counterclockwise (ccw) VP1 monomer interacts with the glycerol chain of Neu5Ac. The backbone amino and carbonyl groups of residue V74 recognize the glycerol chain at one side, and the carbonyl group of residue K82 contacts the O4 of Neu5Ac on the other side of the BC2-loop. Residue H132 of the ccw monomer forms the bottom of the shallow binding groove so that the methyl group of Neu5Ac is recognized via hydrophobic interactions by the non-polar part of the ccw H132 side and by BC2-loop residues A75 and P83. Additionally, water-mediated hydrogen bonds are found between O4 and the glycerol chain of Neu5Ac and the VP1 backbone.

The VP1 surface area that is buried by the Neu5Ac interaction comprises 197 $Å^2$ [41]. Because of the inherent flexibility of glycosidic linkages, complex glycans do not have a single, well-defined three-dimensional structure in solution and the mobility of the sugar rings are reflected in their temperature factors (B-factors). The average B-factor of the Neu5Ac ring of GM1 is in the same range as those of the neighboring protein residues, indicating that the occupancy is near 1.0 for the ligand. In contrast, the GalNAc and the branched Gal ring have elevated B-factors. This is in agreement with their lack of contacts to protein residues and resultant higher mobility. Neu5Ac moieties of 3'SL, 6'SL, and the branched GM1 glycan are bound in an identical manner in all three glycan-VP1 complex structures. Interactions with glycan residues other than the terminal Neu5Ac were not observed in any of the three complex

structures. However, the branched "ganglioside core" GalNAcβ1-3(Neu5Acα2–3)Gal is likely more restricted in its conformational mobility than the linear glycan ligands 3'SL and 6'SL, and this could explain the better defined density for the GM1 glycan.

We next performed saturation transfer difference (STD) NMR spectroscopy experiments to provide evidence for the observed interaction between VP1 and the GM1 glycan in solution (S3 Fig). In this approach, saturation from a macromolecule, such as the VP1 pentamer, is transferred to a bound small-molecule ligand, such as a glycan, within a distance of up to 5 Å from the protein. Thus, this technique can be used to identify the protein-bound parts of the ligand [42]. We found that TSPyV VP1 interacts with the GM1 glycan in solution. Protons of the engaged glycerol chain (H7 and H9) and the methyl group (Me) of the N-acetyl chain of Neu5Ac are readily identified in the STD NMR difference spectrum. These findings are in good agreement with the crystal structure analysis, where mainly interactions with the terminal Neu5Ac and in particular with its glycerol and acetyl groups were observed. Protons of the α2,3-linked Gal and the GalNAc residues receive comparably little magnetization, and these protons are also not engaged by VP1 in the crystal structure.

The sialic acid binding site on TSPyV VP1 mediates cell attachment and infection

In order to verify the functionality of the binding site on TSPyV VP1 for sialic acid-dependent cell binding, mutations in the sialic acid binding site were introduced into the unassembled TSPyV VP1 pentamer construct. To analyze the importance of threonine 73, which recognizes the N-acetyl group of Neu5Ac, this residue was mutated to alanine (T73A) in order to remove the hydrogen bond, and to glutamine (T73Q) and glutamic acid (T73E) in order to introduce steric hindrance. In addition, residues 71, 84, and 137 were also mutated (K71L, T84A, and R137A). The design of relevant VP1 mutations was limited because the TSPyV VP1-Neu5Ac interactions are mainly mediated by the protein backbone.

We used a thio-labeling strategy to standardize the labeling efficiency of TSPyV VP1 wild type and mutants for the flow cytometry experiment. TSPyV possesses a conserved cysteine, C107, within the CD-loop of VP1. This cysteine lies at the base of the pentamer, distant from the ligand binding site, and while it is the only solvent-exposed cysteine in the pentamer available for thio-labeling, it faces towards the interior and is not exposed in the assembled virus. For all mutants, the binding to HEK293 and SVG-A was significantly reduced (Fig 3A). Next, three of the mutations (K71L, T73E, and T84A) were introduced into TSPsV particles. All three mutations reduce cell binding of PsV and infectivity in HEK293 and SVGA cells (Fig 3B and 3C). The PsV infection assay reveals small differences between the mutants with T73E having the largest influence on infectivity in both cell lines tested. This finding is in accord with the VP1-Neu5Ac interaction seen in our structures. Residue T73 recognizes the N-acetyl group of Neu5Ac and has to be considered a central component of the binding site. Although none of the selected mutations abolishes PsV infection completely, each one of them reduces binding of VP1 pentamers and PsVs, indicating that the identified binding site on TSPyV clearly mediates attachment to sialylated glycans on cells, and thus defines an important event during early steps of TSPyV infection.

A new location of the sialic binding site on the polyomavirus VP1 protein

To date, X-ray structures of seven different polyomavirus VP1 proteins in complex with sialylated glycans have been solved [19–21,23–26]. The human JCPyV, BKPyV, MCPyV and HPyV9 viruses, but also the two simian viruses LPyV and SV40 and the murine virus MPyV engage sialic acids of their specific glycan receptor motifs in the same general area of VP1

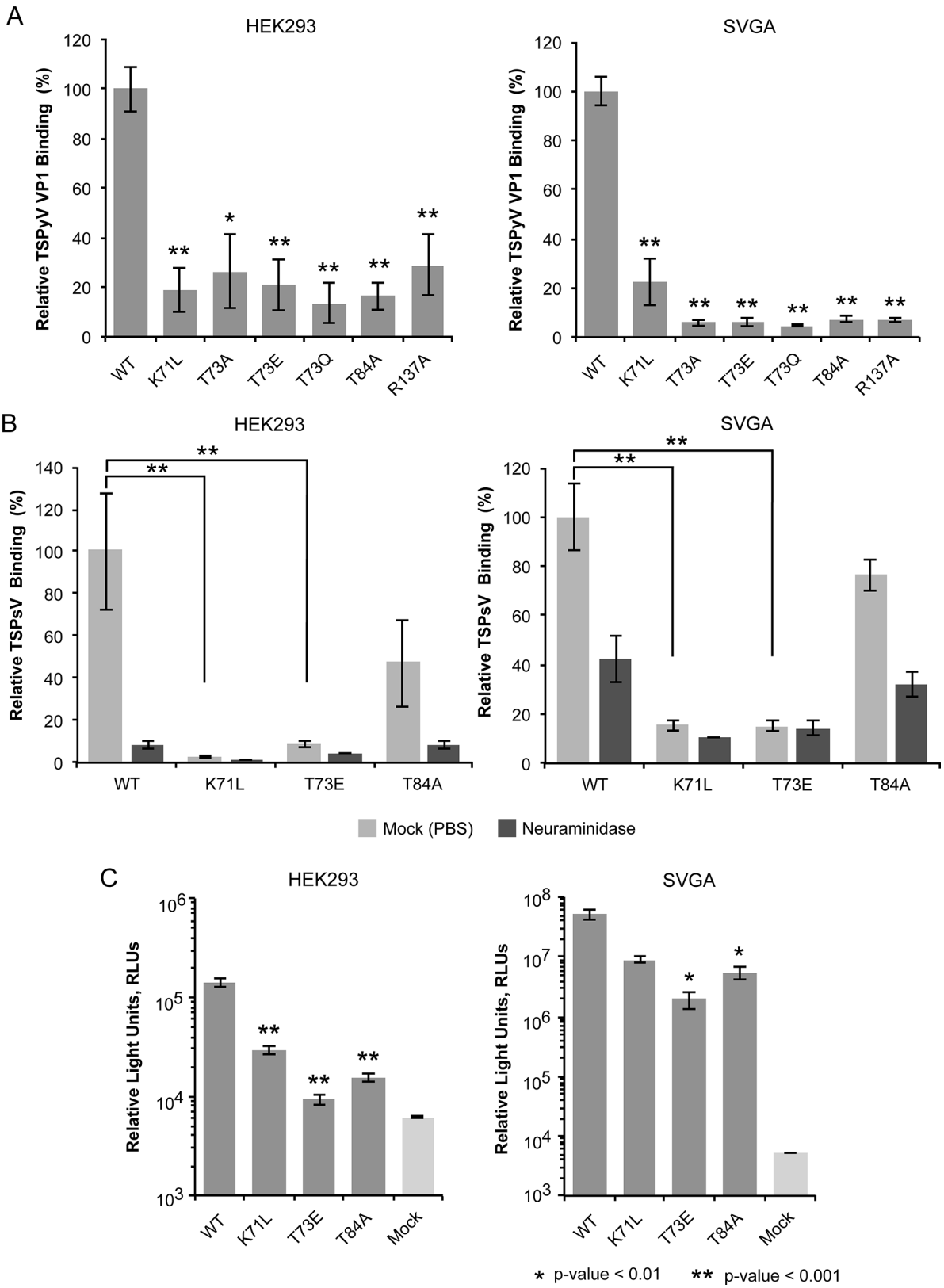


Fig 3. Cell binding and PsV infection of TSPyV mutants. (A) Cell binding analysis of TSPyV wild type (WT) and mutant VP1 pentamers to HEK293 and SVGA cells by flow cytometry. VP1 pentamers were covalently thio-labelled using Alexa Fluor 488 C5 maleimide. The histogram depicts the relative

fluorescence signal of mutants compared to the binding signal of wild type pentamers. Data was measured in three independent experiments and was standardized for the background signal from cells alone. Error bars indicate standard deviations. (B) WT and mutant TSPsV binding (Alexa Fluor 488 amine-conjugated) to HEK293 and SVGA cells (mock-treated cells and cells incubated with *Clostridium perfringens* neuraminidase for 30 min prior to PsV infection) examined by flow cytometry. Relative fluorescence signals of mutants compared to the binding signal of WT PsV were plotted and error bars indicate standard deviations for three experimental replicates. Labeled PsV content was normalized by western blotting detection with purified PAB597, a hybridoma supernatant that produces a monoclonal antibody against JCPyV VP1, which cross-reacts with TSPyV VP1. The two-tailed unpaired t test was performed for binding of WT and mutant TSPsVs to mock-treated cells. (C) TSPsV infection of HEK293 and SVGA cells assayed 72 h post infection by detection of the secreted luciferase due to transduction of the reporter plasmid phGluc. The luciferase was quantified using a BioLux *Gaussia* Luciferase Assay Kit. The secreted luciferase catalyzes a photo-oxidation that emits light, which is measured in relative light units, RLUs. RLUs are given in logarithmic scale. Mock PsV infections were done with a control sample obtained according the PsV purification protocol from cells only transfected with phGluc and control plasmid measure background and nonspecific transfer of luciferase. PsV experiments were done in quadruplicate and repeated two times. Statistic analysis was performed using the two-tailed unpaired t test.

doi:10.1371/journal.ppat.1005112.g003

(Fig 4). This shallow binding site is located at the interface between two monomers, with contributions from the DE-, HI- and BC1-loops and the BC-linker of one monomer and ccw BC2- and EF-loops from the neighboring VP1 chain. In contrast, TSPyV engages Neu5Ac in a unique, exposed binding site that lies about 18 Å away from the sialic-acid binding groove of all other polyomaviruses. This new binding site is formed by the BC2-loop of a single VP1 monomer on top of the pentamer. The comparison of all seven structures reveals that TSPyV VP1 does not differ drastically in its amino acid sequence, length and fold of the surface loops. Full-length TSPyV VP1 shares 52–61% amino sequence identity with these seven other VP1 proteins. BKPyV and JCPyV VP1 are most closely related to TSPyV VP1 in terms of the overall VP1 structures as well as the amino acid sequences with 52 and 54% VP1 sequence identities, respectively. The TSPyV VP1 structure superposes on the BKPyV and JCPyV VP1 structures with small root-mean-square deviation (RMSD) values of about 0.8 Å (calculated for Cα atoms of the VP1 monomers).

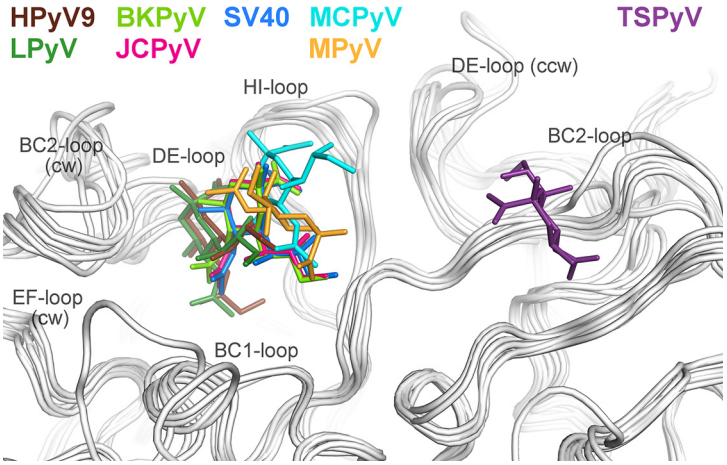


Fig 4. Unique location of the TSPyV Neu5Ac binding site. The location of the Neu5Ac binding site on TSPyV VP1 is shifted compared to binding sites of polyomaviruses SV40, BKPyV, JCPyV, MCPy, LPyV, HPyV9 and MPyV, which all engage sialylated glycans by employing a Neu5Ac binding site in a conserved location on VP1. For clarity, only terminal Neu5Ac residues are shown in stick representations and are coloured according as assigned for each virus. TSPyV VP1-GM1 glycan (purple), SV40 VP1-GM1 glycan (blue; pdb 3BWR), BKPyV VP1-GD3 glycan (light green; pdb 4MJ0), JCPyV VP1-LSTc pentasaccharide (pink; pdb 3NXD), MCPyV VP1-GD1a glycan (cyan; pdb 4FMJ), LPyV-α2,3-sialyllactose (dark green; pdb 4MBY), HPyV9 VP1-α2,3-Neu5Gc-sialyllactose (brown; pdb 4POT), MPyV VP1-Neu5Ac-α2,3-Gal-β1,3-[α2,6-Neu5Ac]–GlcNAc-β1,3-Gal-β1,4-Glc (orange; pdb 1VPS).

doi:10.1371/journal.ppat.1005112.g004

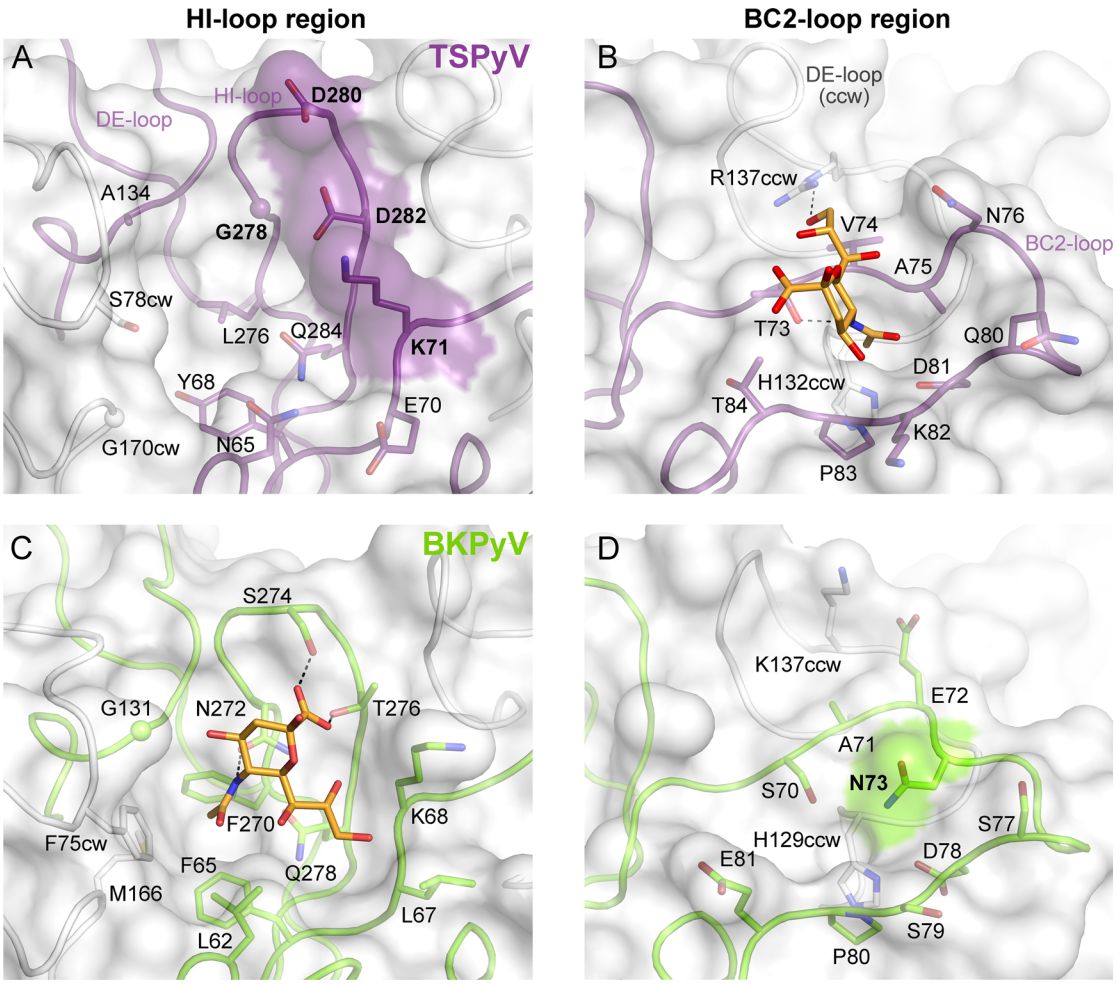


Fig 5. Comparison of the Neu5Ac binding sites of HI- loop and BC2- loop regions. (A, C) HI-loop regions of TSPyV VP1 and of BKPyV VP1 in complex with the GD3 glycan (pdb 4MJ0), are shown in the same orientation. The terminal Neu5Ac of GD3 is show in orange sticks. Residues that do not allow interaction with Neu5Ac in a manner similar to BKPyV are colored on the surface of VP1. Direct hydrogen bonds between BKPyV VP1 side chain residues and Neu5Ac are depicted as dashed lines. (B, D) BC2-loop regions of TSPyV VP1 and BKPyV VP1 are shown. Direct hydrogen bonds between side chain residues of TSPyV VP1 and the terminal Neu5Ac of the GM1 glycan and an intramolecular BKPyV VP1 hydrogen bond are shown as black dashed lines. The structures were aligned using their Cα atoms. One VP1 monomer is highlighted in purple for TSPyV and light green for BKPyV, respectively.

doi:10.1371/journal.ppat.1005112.g005

Structural basis for the relocation of the sialic binding site on VP1

To understand the molecular basis for the relocation of the Neu5Ac binding site we compared the TSPyV VP1-GM1 complex structure with known VP1-glycan complex structures in more detail. Usually, the terminal sialic acids is recognized by HI-loop residues of VP1 via direct interactions with the protruding functional groups of the Neu5Ac, the carboxylate group, the N-acetyl group and the glycerol chain [30]. In TSPyV VP1, only small differences in structure and amino acid sequence within the HI-loop binding region are sufficient to switch the location of the Neu5Ac binding site to the BC2-loop region as shown in an exemplary manner for the comparison with the BKPyV VP1-GD3 glycan structure (Fig 5). The groove forming the binding site for the terminal sialic acid in VP1 of BKPyV and other polyomaviruses next to the HI-loop is still present on TSPyV VP1. This groove is often hydrophobic at the bottom and recognizes the methyl group of N-acetyl chain of Neu5Ac [21,23–25] through a hydrophobic patch formed by residues L62, F65, F270 and F75cw of BKPyV VP1. With Y68 and L276, the

equivalent region of TSPyV VP1 is similarly hydrophobic (Fig 5A and 5C). However, while the carboxylate group of Neu5Ac is recognized by two hydrogen bonds from side chains of residues S274 and T276 in BKPyV VP1, similar interactions are not possible in TSPyV as the recognition of the sialic acid carboxylate group appears to be electrostatically hindered through residues D280 and D282. In addition, residue G278 in TSPyV, which substitutes for residue N272 of BKPyV, cannot recognize the N-acetyl group of Neu5Ac in the HI-binding loop. Residue K71 may also interfere with binding of Neu5Ac in the HI-loop region of TSPyV VP1, although the K71 side chain has elevated temperature factors. On the other hand, the shifted Neu5Ac-binding site in TSPyV VP1 is unable to bind sialic acid in BKPyV VP1 as residue N73 forms a hydrogen bonds with residue S70, thereby widening the BC2-loop and interfering with Neu5Ac binding. Hence, subtle differences seem to modulate the recognition of Neu5Ac both in the HI-loop region and also in the BC2-loop region (Fig 5B and 5D).

Glycolipids are important for TSPyV infection

Sialic acids are abundantly expressed on N- or O-linked glycoproteins, and they are also key components of glycolipids and in particular gangliosides. In order to elucidate the cellular scaffold of the sialic acids engaged by TSPyV we treated HEK293 cells with inhibitors interfering with the synthesis of these molecular scaffolds for terminal sialic acids and monitored TSPyV pentamers binding (Fig 6). Tunicamycin inhibits the synthesis of N-linked glycans, Benzyl-α-N-acetylgalactosamine (BenzylGalNAc) specifically blocks the synthesis of O-linked glycans, and D,L-threo-phenyl-2-hexadecanoylamino-3-morpholino-propanol (PPMP) is a structural analogue of a ceramide and inhibits the UDP-glucose-ceramide glucosyltransferase leading to decreased glycosphingolipid expression on cells. The effects of the cell treatment were assessed with two positive controls: (i) cholera toxin B subunit, which binds GM1 gangliosides, and (i) the lectins Concanavalin A (ConA) and *Maackia amurensis* II (MALII), which bind to N- and O-glycans. In addition, the toxicity of either drug was monitored using a cell proliferation assay, confirming that drug treatments did not reduce the cell viability. Tunicamycin and BenzylGalNAc did not affect cell binding of TSPyV pentamers, whereas a significant reduction of attachment to HEK293 cells was observed after treatment with PPMP. Similar flow cytometry experiments were carried out with SVGA cells. Tunicamycin and BenzylGalNAc treatments did not reduced TSPyV VP1 pentamer binding to SVGA cells. However, cholera toxin B subunit and TSPyV VP1 pentamer binding to PPMP-treated SVGA cells was also not significantly reduced, suggesting that inhibition of the glycosphingolipid expression by PPMP is not very effective in this cell line.

TSPsV transduction in PPMP-treated HEK293 cells was significantly impaired, suggesting that glycolipids are important for TSPyV infection of these cells (Fig 7). Additionally, the human lung cell line A549 could be infected with TSPsV resulting in a reduction after cell treatment with PPMP (S4 Fig). However, we could not observe reduced binding of TSPyV VP1 pentamer to A549 cells after PPMP treatment. Hence, these experiments establish that a glycolipid is likely important for viral entry, but TSPyV might also bind to other sialylated glycans on the cell surface during initial cell attachment, possibly in a cell type-specific manner.

Phylogenetic and structural analysis highlights the close relationship to Orangutan Polyomavirus

The phylogenetic analysis based on VP1 amino acid sequence reveals that TSPyV is most closely related to two simian polyomaviruses, Bornean Orangutan Polyomavirus (OraPyV1) and Ateles paniscus polyomavirus 1 (ApanPyV1) [1,43]. TSPyV shares VP1 amino acid sequence identities of 79% and 69% with OraPyV and ApPyV1, respectively. To investigate

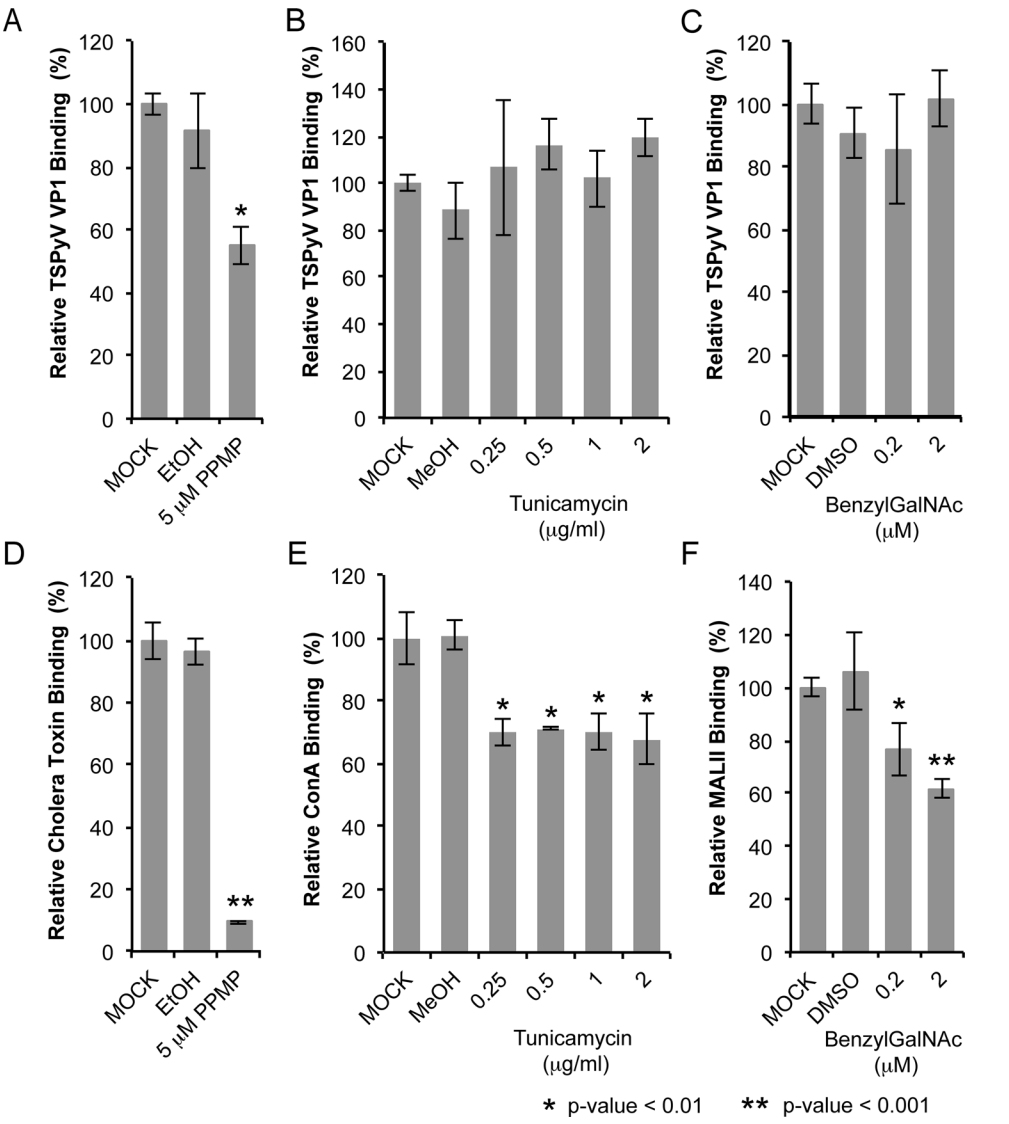


Fig 6. TSPyV binds to sialic acid containing glycolipids. Cell binding analysis of thio-labelled Alexa Fluor 488 TSPyV VP1 pentamers to HEK293 cells treated with inhibitors interfering with the synthesis of glycosphingolipids, N. and O-linked glycoproteins. Cells were treated with (A, D) PPMP, (B, E) tunicamycin, (C, F) BenzylGalNAc, and the respective carrier controls. The binding of TSPyV pentamers was standardized to binding signals to mock (PBS) treated cells. 30,000 gated events were measured for each sample. Cell binding of cholera toxin B subunit and the lectins ConA and MALII was used to monitor the effect of the inhibitors on the cellular expression of glycans. Data was measured in three independent experiments. Error bars indicate standard deviations.

doi:10.1371/journal.ppat.1005112.g006

common features and differences between capsids of these three viruses, we compared their VP1 sequences and mapped conservations onto the molecular surface of TSPyV VP1 (Fig 8). Differences occur predominantly on the top surface of the VP1 pentamer, which is especially accessible to antibodies in the context of the assembled virion. VP1 residues within the BC2-loop binding site are highly conserved in OraPyV (Fig 8A) but the HI-loop region differs in the human and in the monkey virus. The comparison of VP1 proteins from TSPyV, OraPyV and ApanPyV1 highlights that BC2-loop residues T73, V74 and P83 are conserved in contrast to the rest of the upper VP1 surface. In addition, DE-loop residue R137, which interacts with

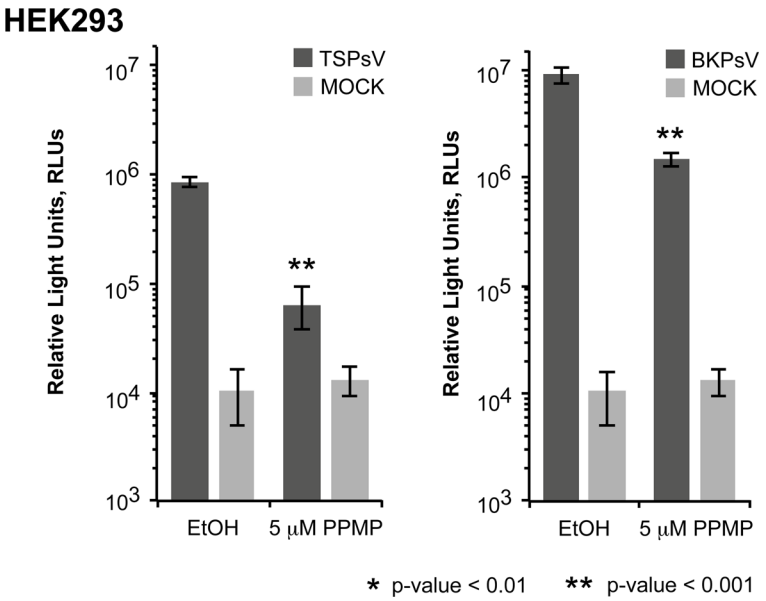


Fig 7. Glycolipids are important for TSPsV infection. TSPsV infection of HEK293 pre-treated with PPMP or the carrier control for 6 days was assayed 72 h post infection by quantification of the secreted luciferase due to transduction of the reporter plasmid phGluc. BKPsV were used as a positive control for a ganglioside-dependent infection. Mock PsV infections were done with a control sample obtained according the PsV protocol from cells only transfected with phGluc and control plasmid to test for background signal of the luciferase assay for uninfected cells. PsV experiments were done in quadruplicate and repeated three times. The results from one representative quadruplicate experiment is shown here as an example. Statistic analysis was performed using the two-tailed unpaired t test.

doi:10.1371/journal.ppat.1005112.g007

the glycerol chain of Neu5Ac in TSPyV VP1, is also present in OraPyV and ApanPyV1 VP1. Consequently, based on the structure-based sequence alignment, we predict that TSPyV shares its sialic acid binding site with OraPyV and ApanPyV1.

Discussion

Although a large number of new polyomavirus family members have been identified in recent years [16], details about their receptor-binding specificities and cell attachment strategies are not known with the exception of MCPyV and HPyV9 [23,26]. We demonstrate here that the skin-tropic TSPyV binds sialylated glycans in a binding site formed by the BC2-loop on top of the VP1 protein that is not utilized by any other structurally characterized member of the polyomavirus family. This new site is highly exposed to solvent and, apart from its location, also unique in that the sialic acid carboxylate group is not engaged by the protein, in contrast to all other known virus-sialic acid interactions [44]. The binding site has functional relevance, as site-directed mutations decrease binding of VP1 pentamers and PsVs to cells as well as PsV transduction. However, since the introduced mutations did not abolish PsV transduction completely, we cannot rule out that additional glycan residues on sialylated oligosaccharides, a different type of glycans such as glycosaminoglycans, or proteinous receptors are involved in the initial steps of TSPyV infection.

Our structures show that TSPyV VP1 can interact with terminal sialic acids in α2,3- and α2,6-linkages. Due to the exposed location of the binding site, glycans bearing terminal α2,8-linked sialic acid may be engaged in a similar manner. It is likely that TSPyV recognizes a more complex glycan motif or receptor similar to other polyomavirus members on its target cells to mount an infection. This hypothesis is supported by the observation that TSPsV

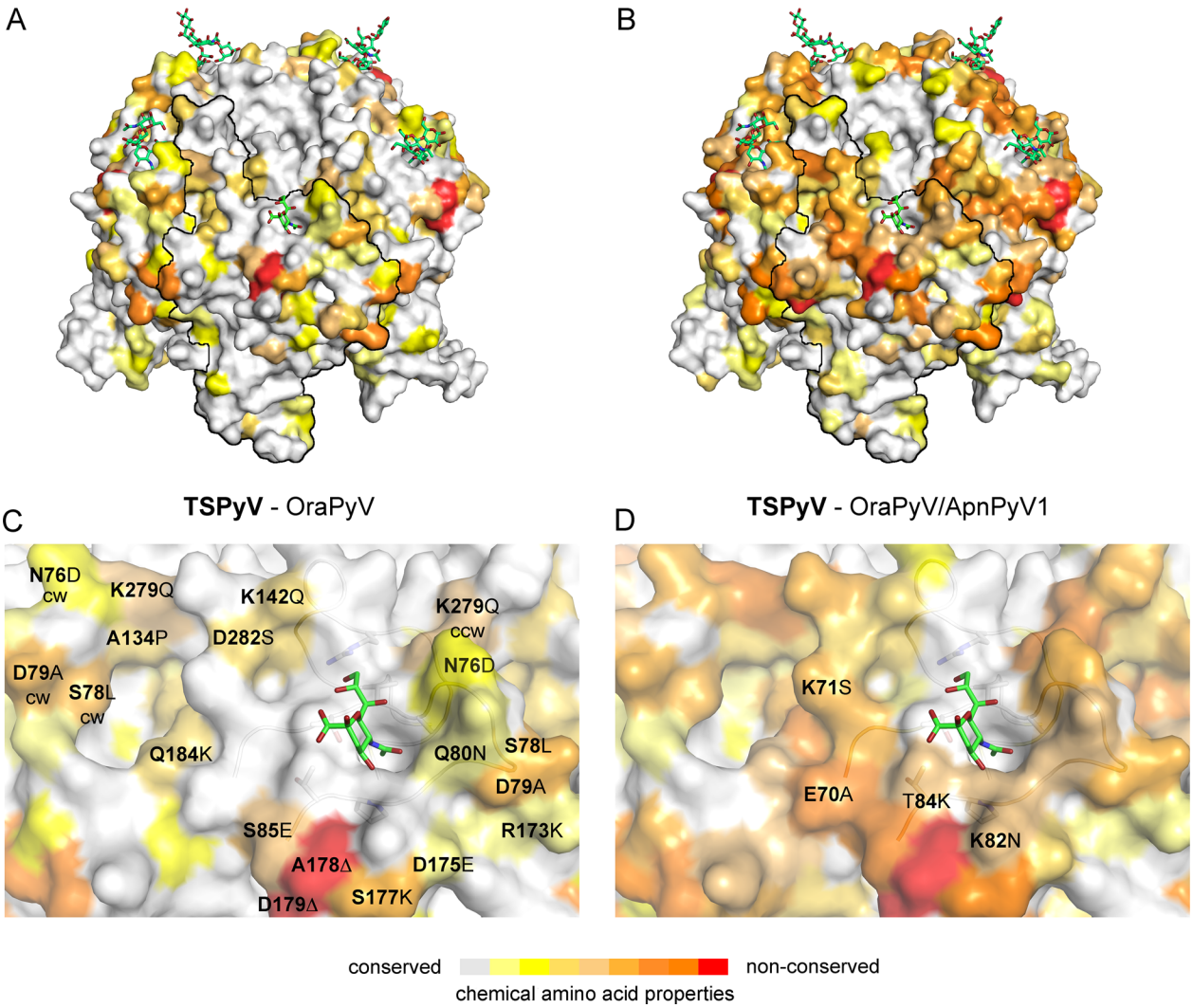


Fig 8. Conserved binding site for terminal Neu5Ac build up by the BC2-loop. Mapping amino acid differences between (A, C) TSPyV VP1 and OraPyV VP1 and (B, D) TSPyV, OraPyV and ApanPyV based on the TSPyV VP1-GM1 complex structure. Amino acid differences between the two or three closely related viruses are coloured according their conservations on the surface of TSPyV VP1. An amino acid sequence alignment was done with Clustal Omega. JalView [67] was also used to assign values for the respective conservation of the chemical acid amino acid properties from 10 (conserved; grey) to 0 (non-conserved; red). Values of 2 to 0 are colored in red. (C-D) Close-up views of the HI- and BC2-loop regions. The backbone of the BC2- and the ccw DE-loops are depicted in cartoon representations. Side chains of conserved key amino acids within the BC2-binding site are shown in sticks in panel C.

doi:10.1371/journal.ppat.1005112.g008

transduction of HEK293 and A549 cells is reduced after interfering with the glycolipid synthesis by drug treatment. However, the cell treatment with PPMP did only reduce binding of TSPyV VP1 pentamers to HEK293, but not to A549 cells suggesting that different glycans may support cell attachment in different cells lines. To identify the complete glycan receptor(s) for TSPyV, further studies are required such as glycan array screening or the use of cutaneous cell lines or primary cells for infection studies. Importantly, functional data are needed to define the role of glycolipids for the TSPyV infection of IRS cells resulting in hyperproliferation and TS in immunocompromised patients [7]. Compared to normal hair follicles, TS hair follicles seem devoid of normal hair shafts and papilla but include instead large numbers of eosinophilic, trichohyalin-positive IRS cells [45]. The composition of cell surface glycans in different skin cells, for example the expression of mostly non-sialylated glycoconjugates in hair follicles

during the keratinization process, has been investigated by lectin histochemistry, but these studies are often incomplete [46–48]. In addition, quantitative data and detailed information about the prevalence of sialic acid-containing glycoconjugates and chemical modifications of sialic acids would advance an understanding of the tropism of TSPyV for IRS cells. With regard to sialic acids, keratinization has also been shown to include reduction of these glycan residues in the plasma membrane of feline hair cortex during the differentiation [49]. Thus, it remains to be investigated if additional factors besides sialic acids support TSPyV infection of these cells. However, it is also important to realize that TSPyV is able to infect other, non-cutaneous cell types, for example tonsillar tissue [14].

Previous functional and structural studies established that polyomaviruses have a plastic binding site for Neu5Ac in a highly conserved region on top of VP1 that can engage the Neu5Ac ring in at least four different orientations [30]. Exceedingly small alterations in or near this binding site of polyomavirus VP1 have been found to modulate receptor specificity, which can render cells refractory to infection [35,24] or are critical for pathogenicity [50,51,36]. The case of TSPyV provides an example of a different strategy for modulating receptor interactions, as the typical binding site is no longer accessible and a new, alternate site has been developed. This new site is likely also present in related polyomaviruses. Examples for virus families in which members use different sites to engage a similar glycan receptor epitope are rare and include reoviruses [52,53] and adenoviruses [54–56]. While the physiologic reason for such alternate binding sites is not clear in either of these cases, virus-glycan interactions seem well suited for such shifts as they are usually of low affinity and depend on only a few specific contacts that can easily be modulated.

Many polyomaviruses form 'pairs' in which a human and a non-human version are most similar in sequence and therefore most likely closely linked in evolution [43,57]. A recent example is the HPyV9-LPyV pair, which displays subtle differences in sialic acid specificity [25,26]. In most cases, however, we lack knowledge about how such pairs differ in structure and receptor recognition, and how differences might relate to host tropism [34]. The architecture and the location of the BC2-loop binding region in TSPyV is likely conserved in OraPyV and ApnPyV1, and it is tempting to speculate that sequence differences near the conserved HI-loop binding region modulate binding to more complex, and perhaps subtly differing, sialylated glycan structures expressed on cells in the respective hosts.

Our analysis of TSPyV glycan binding properties provides a useful platform for further studies to define role of glycans for the TSPyV entry pathway, cell and host tropism as well as pathogenicity in order to understand structural determinants of receptor and host switching events. Such studies will inform not only polyomavirus research but also help to provide guidelines for rationalizing the receptor-binding specificities and tropism of other glycan-binding viruses.

Materials and Methods

Cell culture

Cells were maintained in a humidified 37°C CO_2 chamber in Dulbecco's Modified Eagle Media (DMEM) or Minimum Essential Media (MEM) supplemented with 1% penicillin/streptomycin and 10% heat-inactivated fetal bovine serum (FBS). SVGA cells have been derived from the original SVG human glial cell line established by SV40 transformation of human fetal glial cells [58]. HEK293 cells (human embryonic kidney cell line), HeLa cells (human epithelial cell line derived from cervical carcinoma), Vero cells (monkey kidney epithelial cell line), CHO cells (epithelial cell line derived from hamster ovaries), and A549 cells (human lung epithelial) were purchased from the American Type Culture Collection (ATCC). HEK293TT cells are HEK

cells transformed with two copies of SV40 large T antigen (a gift from Dan DiMaio, Yale University).

Protein expression, purification and fluorescence labeling

cDNA sequence (GenBank accession code ADK12664) coding for amino acids 31–304 of TSPyV VP1 was amplified by PCR and cloned into the *Nhe*I and *Bam*HI restriction sites of pET28b (Novagen). To be consistent with the previously described polyomavirus VP1 pentamer structures and to facilitate structural comparisons, the amino acid numbering of TSPyV VP1 (amino acids 30–303) used here in this study excludes the N-terminal methionine. The N- and C-terminal truncated TSPyV VP1 expression construct possesses an N-terminal hexahistidine tag (His-tag) followed by a thrombin cleavage site. Site-directed mutagenesis was done using 5' to 3' primers as follows (mismatched nucleotides are highlighted in boldface):

K71L: CTAAGGACAACTATGGTTACAGTGAA**CT**AGTAACTGTTGCTAACAGCAG, CTGCTGTTAGCAACAGTTACT**AG**TTCACTGTAACCATAGTTGTCCTTAG; T74A: GGTTACAGTGAAAAAGTA**G**CTGTTGCTAACAGCAGTGACCAG, CTGGTCACTGCTGTTAGCAACAG**C**TACTTTTTCACTGTAACC;
T73E: GGTTACAGTGAAAAAGTA**GAA**GTTGCTAACAGCAGTGACCAGGACAAGC, GCTTGTCCTGGTCACTGCTGTTAGCAAC**TTC**TACTTTTTCACTGTAACC; T74Q: GGTTACAGTGAAAAAGTA**CAG**GTTGCTAACAGCAGTGACCAGGACAAGC, GCTTGTCCTGGTCACTGCTGTTAGCAAC**CTG**TACTTTTTCACTGTAACC T85A: CAGTGACCAGGACAAGCCT**G**CTTCTGGAGAGATAC, GTATCTCTCCAGAAG**C**AGGCTTGTCCTGGTCACTG;
R137A: CTGGTAAATGTTCATATGGCTACTAAA**GC**AATGTATGATGACAAAGGTATTG, CAATACCTTTGTCATCATACATT**GC**TTTAGTAGCCATATGAACATTTACCAG

Expression was performed in *E. coli* BL21 (DE3), and His-tagged VP1 was purified by nickel affinity chromatography from the crude cell extract. For crystallization, the His-tag was cleaved off in solution using thrombin (GE Healthcare). Six non-native residues (GSHMAS) are present at the N-terminus of VP1 after thrombin cleavage. His-tagged and cleaved wild type and mutant VP1 pentamers were further purified by gel filtration (Superdex 200). Prior to flow cytometry experiments, His-tagged VP1 pentamers were labeled with Alexa Fluor 488 C5 maleimide (Invitrogen). VP1 pentamers (1 mg/ml) were incubated for 16–18 h with 80 mM dithiothreitol (DTT) at 4°C. Excess DTT was removed using 2 x 5 ml HiTrap Desalting columns (GE Healthcare) and then the dye (10 mM in 20 mM HEPES pH 7.0, 150 mM NaCl) was added dropwise by gently mixing to the protein solution (0.2–0.3 mg/ml in 20 mM HEPES pH 7.0, 150 mM NaCl) to give a 10 molar excess of the dye. The reaction was incubated for 18 h at 4°C. 10 mM DTT were added and excess of the dye and DTT were removed by desalting. The labeling efficiency was determined by absorption according to the manufacturer's protocol. The molecular ratio was about 1: 4 (VP1 pentamer: dye) for wild type and mutant VP1 pentamers.

Flow cytometry experiments

Cells that were 80–90% confluent were detached non-enzymatically from flasks by incubation with Cellstripper (Corning) for 10–20 min at 37°C. Cells were washed twice with PBS. Next, 10^6 cells were suspended in 75 μl of labeled wild type and mutant TSPyV VP1 pentamer solution (20 μg/ml in PBS) for incubation on ice for 2 h with agitation every 20 min. Cells were washed twice in PBS and then fixed in 500 μl of 1% paraformaldehyde. Analysis was done using a BD FACSCalibur (Becton, Dickinson and Company) flow cytometer equipped with a 488-nm excitation line. Between 30,000 to 50,000 gated events were measured for each sample. Data were analyzed using FlowJo software (Tree Star, Inc.).

The flow cytometry experiments with the PsV were performed on purified PsV labeled with Alexa Fluor 488 carboxylic acid succinimidyl ester (Invitrogen). After labeling, PsVs were

dialyzed into 10 mM Tris pH 7.4, 50 mM NaCl, 0.01 mM $CaCl_2$ using 10,000 MWCO Slide-A-Lyzer dialysis cassettes (Life Technologies) to remove excess dye.

Labeled PsV preparations were normalized for VP1 content by Western blot. 6 x 10^5 cells were mixed with equal amounts of labeled PsV as determined by Western Blot (about 50 μg τοταλ προτειν, measured via absorption at 280 nm) in 50 μl allowed to bind on ice for 30 minutes, washed and resuspended in PBS and then analyzed as pentamers counting 10,000 events. Experiments were performed in triplicate.

Fluorescein-labeled Concanavalin A (ConA, Vector Laboratories, USA) was added at a concentration of 1 μg/ml in 100 μl PBS to 10^6 cells and was incubated on ice for 1 h with agitation every 20 min. Biotin-labeled *Maackia amurensis* II lectin, (MALII; Vector Laboratories, USA) at a concentration of 1 μg/ml was incubated with 10^6 cells in 100 μl PBS on ice for 1 h. In the case of MALII, cells were washed 2x in PBS and detection of MALII was carried out with Alexa Fluor 488-labeled Streptavidin (488-Strep, Invitrogen), which was added at 2 μg/ml to about 10^6 cells in 100 μl PBS for 30 min. Mock treated cells were incubated with 488-Strep as a negative control. The cholera toxin subunit B Alexa Fluor 488 conjugate (Invitrogen) was added at 10–20 μg/ml in 100 μl PBS to 10^6 cells and was incubated with 10^6 cells in 100 μl PBS on ice for 2 h. Prior to measurement by flow cytometry, cells were washed again twice with PBS and were then fixed in 500 μl of PBS, 1% paraformaldehyde.

PsV production

For expression in the human derived cell line HEK293TT, codon optimization of TSPyV VP1, VP2 and VP3 genes was performed according the guidelines from the National Cancer Institute Center for Cancer Research Lab of Cellular Oncology Technical Files (http://home.ccr.cancer.gov/LCO/production.asp). DNA was synthesized by BlueHeron Biotechnology, LLC (Bothell, WA, USA). The VP1 gene was subcloned into the pwP vector in place of the MPyV VP1 gene and VP2 and VP3 genes into the ph2p vectors in place of the respective MPyV genes. Sequences are based on VP1, VP2 and VP3 sequences (NCBI reference YP_003800006, YP_003800004, YP_003800005) from the full length TSPyV genome (NCBI reference NC_014361). The luciferase reporter vector phGluc expresses a secreted form of *Gaussia* luciferase under control of the EF1α promoter. The parent PsV plasmids were obtained from AddGene (Cambridge, MA, USA). Site-directed mutagenesis of PsV VP1 was done using 5' to 3' primers as follows (mismatched nucleotides are highlighted in boldface):

K71L: GATAATTACGGCTATTCCGAG**CT**GGTGACCGTCGCCAATTCATC;
GATGAATTGGCGACGGTCACC**AG**CTCGGAATAGCCGTAATTATC;
T73E: CTATTCCGAGAAGGTG**GAA**GTCGCCAATTCATCCGATC;
GATCGGATGAATTGGCGAC**TTC**CACCTTCTCGGAATAG;
T84A: CGATCAAGATAAACCC**G**CCAGCGGCGAAATCCCCAC;
GTGGGGATTTCGCCGCTGG**C**GGGTTTATCTTGATCG;

TSPsV were purified by a sucrose gradient followed by a CsCl gradient as described earlier [59].

Titers of purified PsV were determined according the encapsidated reported plasmid phGluc for infections studies. PsV preparations were treated with DNAse I (New England Bioloabs) for 60 min. DNAse I was then inactivated by 75°C for 10 min and protected phGluc was extracted using the DNeasy Blood and Tissue kit (Qiagen). The titer of packaged reporter plasmid was determined by TaqMan quantitative PCR (Applied Biosystems) and a standard curve for serial dilutions of phGluc.

For PsV binding studies, labeled PsV content was normalized by Western blotting detection with purified PAB597, a hybridoma supernatant that produces a monoclonal antibody against JCPyV VP1, which we found to cross-react with TSPyV VP1.

PsV infections

For PsV infection assay cells were seeded to 6,000–7,000 cells per well in a 96-well plate 18 h prior the infection. PsV Infections were performed by adding 5.5 x 10^5 or 5.5 x 10^6 PsV particles to each well in 35 μl serum free media for 2 h at 37°C with 30 min agitations. Cells were washed twice with PBS and media was the replaced with phenol red free media with FBS. Mock PsV controls were generated by transfecting 293TT cells with control plasmid and the phGluc reporter plasmid in a 7:1 ratio and then purifying according to the PsV purification protocol. PsV infection was measured after 3 days by detection of secreted luciferase according to the manufacturer's instructions (BioLux *Gaussia* Luciferase Assay Kit, New England Biolabs) using an opaque 96-well microplate in a GloMax Multi-Detection System Luminometer (Promega) equipped with an autoinjector. Infection assays were done in quadruplicate and repeated three times. The Cell Titer 96 Aqueous Non-Radioactive Cell Proliferation Assay (Promega) was used according to the manufactor's protocol to test for toxicity after cell treatments and the absorption was measured at 450 nm.

Treatment of cells with neuraminidase

Prior to flow cytometry experiments, cells were detached non-enzymatically and washed twice with PBS (see above). Then, 10^6 cells were resuspended in 100 μl PBS and treated with 1 U/ml neuraminidase (*Clostridium perfringens* neuraminidase type V, Sigma-Aldrich) for 30 min at 37°C with agitation every 10 min. Prior to PsV infection or binding, cells were washed once with serum free media for infection studies or PBS for binding studies and incubated with 1 U/ml neuraminidase in PBS or in PBS (Mock) for 30 min at 37°C with agitation every 10 min. Cells were then washed with media and serum free media before PsV infection or twice with PBS before binding studies.

Treatment of cells with inhibitors of N- and O-linked glycosylations and inhibitor of glycosphingolipid synthesis

Cells were plated in 6-well plates (5 x 10^5 cells/per well) for the flow cytometry experiment and in a 96-well plate (10^4 cells/per well) for PsV infections. Glycosylation inhibitors were mixed in media containing 10% FBS at various concentrations to plated cells. Cells were preincubated with tunicamycin (Sigma-Aldrich) in methanol at 37°C for 16–18 h and with BenzylGalNAc (Calbiochem, USA) in DMSO for 48 h prior to the flow cytometry experiment. D,L-threo-phenyl-2-hexadecanoylamino-3-morpholino-propanol (PPMP) (Matreya LLC, USA) in ethanol was added to the media of cells for 6 days, and media supplemented with PPMP was changed after days 2 and 4. Vehicle controls with equivalent volumes of methanol, DMSO or ethanol for the highest concentration were used in each case. The Cell Titer 96 Aqueous Non-Radioactive Cell Proliferation Assay (Promega) was used according to the manufacturer's protocol to assess toxicity of the inhibitors (absorption at 450 nm).

Crystallization and data collection

TSPyV VP1 was concentrated to 4.8 mg/ml in 20 mM HEPES pH 7.5, 150 mM NaCl, 20 mM dithiothreitol (DTT) and crystallized at 20°C by sitting drop vapor diffusion against a reservoir containing 100 mM sodium malonate pH 5.0, 10% (w/v) PEG 3350. Drops were set up using 1 μl protein solution, 1 μl reservoir solution and 0.2 μl microseeding solution, which was prepared from previously obtained TSPyV VP1 microcrystals. For complex formation with the GM1 glycan and 6'SL, crystals were grown in a drop supplemented with 10 mM GM1 pentasaccharide (Elicityl, France) or 10 mM 6'SL (Carbosynth, UK). Crystals were harvested after

five days and transferred stepwise into reservoir solution complemented with the respective glycan and 30% (v/v) glycerol for cryoprotection before flash-freezing them in liquid nitrogen. For complex formation with 3'SL, preformed native TSPyV VP1 crystals were soaked for 15 min in the reservoir solution supplemented with 10 mM 3'SL (Carbosynth, UK) before flash freezing. Data sets were collected at beam lines X06DA and X06SA at the Swiss Light Source (Villigen, Switzerland).

Structure determination

Diffraction data was processed with XDS [60], and structures were solved by molecular replacement with Phaser in CCP4 [61,62]. The MCPyV VP1 core structure (pdb: 4FMG) was taken as a homology search model for the native TSPyV VP1 structure, which was then used for structure determination of glycan complex structures. Rigid body and simulated annealing refinement was carried out with Phenix [63]. Alternating rounds of model building in Coot [64] and restrained refinement including the translation-libration-screw (TLS) method [65] and 5-fold noncrystallographic symmetry (NCS) restrained refinement were done with Refmac5 [66]. The carbohydrate ligands were located in $2F_o$-F_c and F_o-F_c electron density maps. After incorporation into the model, ligands were refined using restraints from the CCP4 library and user defined restraints for the α2,3 and α2,3- glycosidic bonds. The native TSPyV VP1 structure and the structures complexed with 3'SL and 6'SL possess ten VP1 chains in the asymmetric unit, while the asymmetric unit of the VP1-GM1 glycan complex structure comprises five chains. The final VP1 structures contain amino acid residues 33–40, 44–101 and 110–303 for all chains. Residues of the short loop region between the A_{new}- and the B-strand (residues 40–44) and the CD-loop (residues 101–110) could be built into defined electron density for some VP1 chains. However, these residues generally display elevated temperature factors, indicating flexibility of these two regions at the bottom of the unassembled VP1 pentamer if not involved in crystal contacts. To assess amino acid conservations on the VP1 surface, an alignment was carried out with Clustal Omega and JalView [67] with the VP1 sequences from OraPyV (NCBI reference YP_003264533) and ApanPyV1 (NCBI reference YP_007195272). Assign values for the conservation of the chemical acid amino acid properties from 10 (conserved; grey) to 0 (non-conserved; red) in JalView were than highlighted on the TSPyV VP1-GM1 complex surface. Values of 2 to 0 are coloured in red.

Structure figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC).

Saturation transfer difference (STD) NMR

STD NMR spectra were recorded using 3 mm tubes on a Bruker AVIII-600 MHz spectrometer equipped with a room temperature probe head at 288 K. Data was processed with TOPSPIN 3.0 (Bruker). The sample contained 2 mM GM1 oligosaccharide (Elicityl, France) or 20 μM TSPyV VP1 and 2 mM GM1 oligosaccharide, respectively, in 20 mM K_2HPO_4/KH_2PO_4 pH 7.4, 150 mM NaCl, 99%D_2O. Off- and on-resonance frequencies were -30 ppm and 7.3 ppm, respectively. The irradiation power and length of the selective pulse train was 57 Hz and 2 s, respectively. In order to suppress residual protein resonances a continuous-wave spin-lock pulse with a strength of 3.2 kHz was employed. The relaxation delay was 3 s and a total of 5 k scans were recorded. Spectra were multiplied with a Gaussian window function prior to Fourier transform and referenced to the a-D-Glc anomeric proton as an internal standard [68].

Accession numbers

Coordinates and structure factor amplitudes were deposited in the RCSB Protein Data Bank (www.pdb.org) under accession codes 4U5Z (unliganded TSPyV VP1), 4U60 (TSPyV VP1-GM1 glycan), 4U61 (TSPyV VP1-6'SL) and 4U62 (TSPyV VP1-3'SL).

Supporting Information

S1 Fig. The binding of Alexa Fluor 488-labelled VP1 pentamers from TSPyV and BKPyV to Mock (PBS) or *Clostridium perfringens* neuraminidase type V pre-treated cells measured by flow cytometry. Non-standardized raw data is shown here for individual experiments. Data was standardized for signals of mock treated cells to obtain the histogram in Fig 1A showing the relative binding of TSPyV VP1 pentamers. The relative average fluorescence from three independent experiments is shown compared to binding to untreated cells. 30,000 gated events were measured for each sample.
(TIF)

S2 Fig. Screening cell lines for TSPsV infection. TSPyV pentamers bind to all tested cell lines in a sialic acid-dependent manner (see Fig 1 and S2 Fig panel B), but TSPsV transduction could only be detected for HEK293 and SVGA cells. (A) Common cell lines were infected with TSPsV to obtain information about the tropism. At 72 h post infection the luciferase expression was measured to determine transfection efficiency of TSPsV. Cell lines where TSPsV infection was undetectable are in general able to express the reporter plasmid [40]. The average relative luciferase units (RLUs) for one experiment performed in quadruplet are shown in log scale. Error bars represent standard deviations. Two different dilutions of particles were used and compared to Mock PsV control. The mock PsV control was generated by harvesting HEK293TT cells transfected with control plasmid instead of the capsid expression plasmids and then purifying according to the PsV purification protocol to measure background signal and non-specific transfer of luciferase to infected cells. (B) The binding of Alexa Fluor 488-labelled TSPyV VP1 pentamers to Mock (PBS) or *Clostridium perfringens* neuraminidase type V pre-treated Vero and CHO cells was measured by flow cytometry. BKPyV VP1 pentamers were used as a positive control. Non-standardized raw data is shown here for representative individual experiments. Three independent experiments were performed and 30,000 gated events were measured for each sample.
(TIF)

S3 Fig. Mapping the GM1 glycan binding epitope of TSPyV VP1 in solution. Saturation transfer difference (STD) NMR of TSPyV VP1 with the GM1 glycan. From top to bottom: STD-NMR difference spectrum of 50 μM TSPyV VP1 with 1 mM GM1 glycan; ^{1}H reference spectrum recorded with the same sample; STD spectrum of the GM1 glycan alone.
(TIF)

S4 Fig. Glycolipids are important for TSPsV infection in A549 cells. TSPsV transduction of A549 cells pre-treated with PPMP or the carrier control for 6 days was assayed 72 h post infection by quantification of the secreted luciferase due to transduction of the reporter plasmid phGluc. BKPsV were used as a positive control for a ganglioside-dependent infection. Mock PsV infections were done with a control sample obtained according the PsV purification protocol from cells only transfected with phGluc and control plasmid to assess the background signal of the luciferase assay. PsV experiments were done in triplicate or quintuplicate. The data from the quintuplicate experiment is shown as a representative example. Statistic analysis was

performed using the two-tailed unpaired t test.
(TIF)

Acknowledgments

We thank Markus Schrader for help with model building and members of the Stehle and Atwood laboratory for critical discussion. We thank the Swiss Light Source (Villigen, Switzerland) for beamtime and the staff at beamline X06DA and X06SA for assistance during data collection and Dr. Remco Sprangers (Max Planck Institute for Developmental Biology, Tübingen, Germany) for assistance in recording the NMR data.

Author Contributions

Conceived and designed the experiments: LJS GVG BSB WJA TS. Performed the experiments: LJS GVG BSB ASD. Analyzed the data: LJS GVG BSB WJA TS. Contributed reagents/materials/analysis tools: ASD MCWF. Wrote the paper: LJS GVG WJA TS.

References

1. van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, et al. (2010) Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLoS pathogens 6: e1001024. doi: 10.1371/journal.ppat.1001024 PMID: 20686659
2. Haycox CL, Kim S, Fleckman P, Smith LT, Piepkorn M, et al. (1999) Trichodysplasia spinulosa—a newly described folliculocentric viral infection in an immunocompromised host. The journal of investigative dermatology Symposium proceedings / the Society for Investigative Dermatology, Inc [and] European Society for Dermatological Research 4: 268–271.
3. Heaphy MR Jr., Shamma HN, Hickmann M, White MJ (2004) Cyclosporine-induced folliculodystrophy. Journal of the American Academy of Dermatology 50: 310–315. PMID: 14726894
4. Sperling LC, Tomaszewski MM, Thomas DA (2004) Viral-associated trichodysplasia in patients who are immunocompromised. Journal of the American Academy of Dermatology 50: 318–322. PMID: 14726896
5. Tan BH, Busam KJ (2011) Virus-associated Trichodysplasia spinulosa. Advances in anatomic pathology 18: 450–453. doi: 10.1097/PAP.0b013e318234aad2 PMID: 21993271
6. Kazem S, van der Meijden E, Kooijman S, Rosenberg AS, Hughey LC, et al. (2012) Trichodysplasia spinulosa is characterized by active polyomavirus infection. Journal of Clinical Virology 53: 225–230. doi: 10.1016/j.jcv.2011.11.007 PMID: 22196870
7. Kazem S, van der Meijden E, Wang RC, Rosenberg AS, Pope E, et al. (2014) Polyomavirus-associated Trichodysplasia spinulosa involves hyperproliferation, pRB phosphorylation and upregulation of p16 and p21. PLoS One 9: e108947. doi: 10.1371/journal.pone.0108947 PMID: 25291363
8. Osswald SS, Kulick KB, Tomaszewski MM, Sperling LC (2007) Viral-associated trichodysplasia in a patient with lymphoma: a case report and review. Journal of cutaneous pathology 34: 721–725. PMID: 17696921
9. Matthews MR, Wang RC, Reddick RL, Saldivar VA, Browning JC (2011) Viral-associated trichodysplasia spinulosa: a case with electron microscopic and molecular detection of the trichodysplasia spinulosa-associated human polyomavirus. Journal of cutaneous pathology 38: 420–431. doi: 10.1111/j.1600-0560.2010.01664.x PMID: 21251037
10. Kazem S, van der Meijden E, Feltkamp MC (2013) The trichodysplasia spinulosa-associated polyomavirus; virological background and clinical implications. APMIS.
11. van der Meijden E, Kazem S, Burgers MM, Janssens R, Bouwes Bavinck JN, et al. (2011) Seroprevalence of trichodysplasia spinulosa-associated polyomavirus. Emerging infectious diseases 17: 1355–1363. doi: 10.3201/eid1708.110114 PMID: 21801610
12. Nicol JT, Robinot R, Carpentier A, Carandina G, Mazzoni E, et al. (2013) Age-specific seroprevalences of merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. Clin Vaccine Immunol 20: 363–368. doi: 10.1128/CVI.00438-12 PMID: 23302741
13. van der Meijden E, Bialasiewicz S, Rockett RJ, Tozer SJ, Sloots TP, et al. (2013) Different serologic behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 polyomaviruses found on the skin. PLoS One 8: e81078. doi: 10.1371/journal.pone.0081078 PMID: 24278381

14. Sadeghi M, Aaltonen LM, Hedman L, Chen T, Soderlund-Venermo M, et al. (2014) Detection of TS polyomavirus DNA in tonsillar tissues of children and adults: Evidence for site of viral latency. J Clin Virol 59: 55–58. doi: 10.1016/j.jcv.2013.11.008 PMID: 24315796

15. Feng H, Shuda M, Chang Y, Moore PS (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 319: 1096–1100. doi: 10.1126/science.1152586 PMID: 18202256

16. DeCaprio JA, Garcea RL (2013) A cornucopia of human polyomaviruses. Nat Rev Microbiol 11: 264–276. doi: 10.1038/nrmicro2992 PMID: 23474680

17. Rennspiess D, Pujari S, Keijzers M, Abdul-Hamid MA, Hochstenbag M, et al. (2015) Detection of human polyomavirus 7 in human thymic epithelial tumors. J Thorac Oncol 10: 360–366. doi: 10.1097/JTO.0000000000000390 PMID: 25526237

18. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, et al. (1991) Structure of simian virus 40 at 3.8-A resolution. Nature 354: 278–284. PMID: 1659663

19. 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

20. 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

21. Neu U, Maginnis MS, Palma AS, Ströh LJ, Nelson CD, 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. doi: 10.1016/j.chom.2010.09.004 PMID: 20951965

22. 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. doi: 10.1128/JVI.00382-11 PMID: 21543504

23. 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 Pathog 8: e1002738. doi: 10.1371/journal.ppat.1002738 PMID: 22910713

24. 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 Pathog 9: e1003688. doi: 10.1371/journal.ppat.1003688 PMID: 24130487

25. 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 Pathog 9: e1003714. doi: 10.1371/journal.ppat.1003714 PMID: 24204265

26. Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, et al. (2014) Crystallographic and Glycan Microarray Analysis of Human Polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate. J Virol.

27. Ströh LJ, Neu U, Blaum BS, Buch MHC, Garcea RL, et al. (2014) Structure Analysis of the Major Capsid Proteins of Human Polyomaviruses 6 and 7 Reveals an Obstructed Sialic Acid Binding Site. Journal of Virology 88: 10831–10839. doi: 10.1128/JVI.01084-14 PMID: 25008942

28. Varki A, Schauer R (2009) Sialic Acids. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P et al., editors. Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY).

29. Stencel-Baerenwald JE, Reiss K, Reiter DM, Stehle T, Dermody TS (2014) The sweet spot: defining virus-sialic acid interactions. Nat Rev Microbiol 12: 739–749. doi: 10.1038/nrmicro3346 PMID: 25263223

30. Ströh LJ, Stehle T (2014) Glycan Engagement by Viruses: Receptor Switches and Specificity. Annual Review of Virology 1: 285–306.

31. Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A (1998) The molecular basis for the absence of N-glycolylneuraminic acid in humans. J Biol Chem 273: 15866–15871. PMID: 9624188

32. 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.

33. Samraj AN, Laubli H, Varki N, Varki A (2014) Involvement of a non-human sialic Acid in human cancer. Front Oncol 4: 33. doi: 10.3389/fonc.2014.00033 PMID: 24600589

34. Stehle T, Khan ZM (2014) Rules and Exceptions: Sialic Acid Variants and Their Role in Determining Viral Tropism. Journal of Virology 88: 7696–7699. doi: 10.1128/JVI.03683-13 PMID: 24807712

35. Magaldi TG, Buch MH, Murata H, Erickson KD, Neu U, et al. (2012) Mutations in the GM1 binding site of simian virus 40 VP1 alter receptor usage and cell tropism. J Virol 86: 7028–7042. doi: 10.1128/JVI.00371-12 PMID: 22514351

36. Maginnis MS, Ströh LJ, Gee GV, O'Hara BA, Derdowski A, et al. (2013) Progressive multifocal leukoencephalopathy-associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding. MBio 4: e00247–00213. doi: 10.1128/mBio.00247-13 PMID: 23760462

37. Ströh LJ, Maginnis MS, Blaum BS, Nelson CD, Neu U, et al. (2015) The Greater Affinity of JC Polyomavirus Capsid for alpha2,6-Linked Lactoseries Tetrasaccharide c than for Other Sialylated Glycans Is a Major Determinant of Infectivity. J Virol 89: 6364–6375. doi: 10.1128/JVI.00489-15 PMID: 25855729

38. Nakanishi A, Chapellier B, Maekawa N, Hiramoto M, Kuge T, et al. (2008) SV40 vectors carrying minimal sequence of viral origin with exchangeable capsids. Virology 379: 110–117. doi: 10.1016/j.virol.2008.06.032 PMID: 18667220

39. Schowalter RM, Pastrana DV, Buck CB (2011) Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry. PLoS Pathog 7: e1002161. doi: 10.1371/journal.ppat.1002161 PMID: 21829355

40. Gee GV, O'Hara BA, Derdowski A, Atwood WJ (2013) Pseudovirus mimics cell entry and trafficking of the human polyomavirus JCPyV. Virus Res 178: 281–286. doi: 10.1016/j.virusres.2013.09.030 PMID: 24100235

41. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797. PMID: 17681537

42. Mayer M, Meyer B (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angewandte Chemie-International Edition 38: 1784–1788.

43. Scuda N, Madinda NF, Akoua-Koffi C, Adjogoua EV, Wevers D, et al. (2013) Novel polyomaviruses of nonhuman primates: genetic and serological predictors for the existence of multiple unknown polyomaviruses within the human population. PLoS Pathog 9: e1003429. doi: 10.1371/journal.ppat.1003429 PMID: 23818846

44. Neu U, Bauer J, Stehle T (2011) Viruses and sialic acids: rules of engagement. Curr Opin Struct Biol 21: 610–618. doi: 10.1016/j.sbi.2011.08.009 PMID: 21917445

45. Schwieger-Briel A, Balma-Mena A, Ngan B, Dipchand A, Pope E (2010) Trichodysplasia Spinulosa-A Rare Complication in Immunosuppressed Patients. Pediatric Dermatology 27: 509–513. doi: 10.1111/j.1525-1470.2010.01278.x PMID: 20796236

46. Reano A, Faure M, Jacques Y, Reichert U, Schaefer H, et al. (1982) Lectins as Markers of Human Epidermal-Cell Differentiation. Differentiation 22: 205–210. PMID: 6816653

47. Schaumburg-Lever G, Alroy J, Ucci A, Lever WF (1984) Distribution of carbohydrate residues in normal skin. Arch Dermatol Res 276: 216–223. PMID: 6206805

48. Ohno J, Fukuyama K, Epstein WL (1990) Glycoconjugate expression of cells of human anagen hair follicles during keratinization. Anat Rec 228: 1–6. PMID: 1700646

49. Ishii M, Tsukise A, Meyer W (2001) Lectin histochemistry of glycoconjugates in the feline hair follicle and hair. Ann Anat 183: 449–458. PMID: 11677811

50. 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

51. Gorelik L, Reid C, Testa M, Brickelmaier M, Bossolasco S, et al. (2011) Progressive Multifocal Leukoencephalopathy (PML) Development Is Associated With Mutations in JC Virus Capsid Protein VP1 That Change Its Receptor Specificity. Journal of Infectious Diseases 204: 103–114. doi: 10.1093/infdis/jir198 PMID: 21628664

52. Reiter DM, Frierson JM, Halvorson EE, Kobayashi T, Dermody TS, et al. (2011) Crystal structure of reovirus attachment protein sigma1 in complex with sialylated oligosaccharides. PLoS Pathog 7: e1002166. doi: 10.1371/journal.ppat.1002166 PMID: 21829363

53. Reiss K, Stencel JE, Liu Y, Blaum BS, Reiter DM, et al. (2012) The GM2 glycan serves as a functional coreceptor for serotype 1 reovirus. PLoS Pathog 8: e1003078. doi: 10.1371/journal.ppat.1003078 PMID: 23236285

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. PMID: 15220447

55. Seiradake E, Henaff D, Wodrich H, Billet O, Perreau M, et al. (2009) The cell adhesion molecule "CAR" and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution. PLoS Pathog 5: e1000277. doi: 10.1371/journal.ppat.1000277 PMID: 19119424

56. Lenman A, Liaci AM, Liu Y, Ardahl C, Rajan A, et al. (2015) Human Adenovirus 52 Uses Sialic Acid-containing Glycoproteins and the Coxsackie and Adenovirus Receptor for Binding to Target Cells. PLoS Pathog 11: e1004657. doi: 10.1371/journal.ppat.1004657 PMID: 25674795

57. Nicol JT, Liais E, Potier R, Mazzoni E, Tognon M, et al. (2014) Serological cross-reactivity between Merkel cell polyomavirus and two closely related chimpanzee polyomaviruses. PLoS One 9: e97030. doi: 10.1371/journal.pone.0097030 PMID: 24816721

58. Major EO, Miller AE, Mourrain P, Traub RG, Dewidt E, et al. (1985) Establishment of a Line of Human-Fetal Glial-Cells That Supports Jc Virus Multiplication. Proceedings of the National Academy of Sciences of the United States of America 82: 1257–1261. PMID: 2983332

59. Nelson CD, Ströh LJ, Gee GV, O'Hara BA, Stehle T, et al. (2015) Modulation of a Pore in the Capsid of JC Polyomavirus Reduces Infectivity and Prevents Exposure of the Minor Capsid Proteins. J Virol 89: 3910–3921. doi: 10.1128/JVI.00089-15 PMID: 25609820

60. Kabsch W (2010) Xds. Acta Crystallographica Section D-Biological Crystallography 66: 125–132.

61. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. Journal of Applied Crystallography 40: 658–674. PMID: 19461840

62. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallographica Section D-Biological Crystallography 67: 235–242.

63. 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

64. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallographica Section D-Biological Crystallography 66: 486–501.

65. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallographica Section D-Biological Crystallography 62: 439–450.

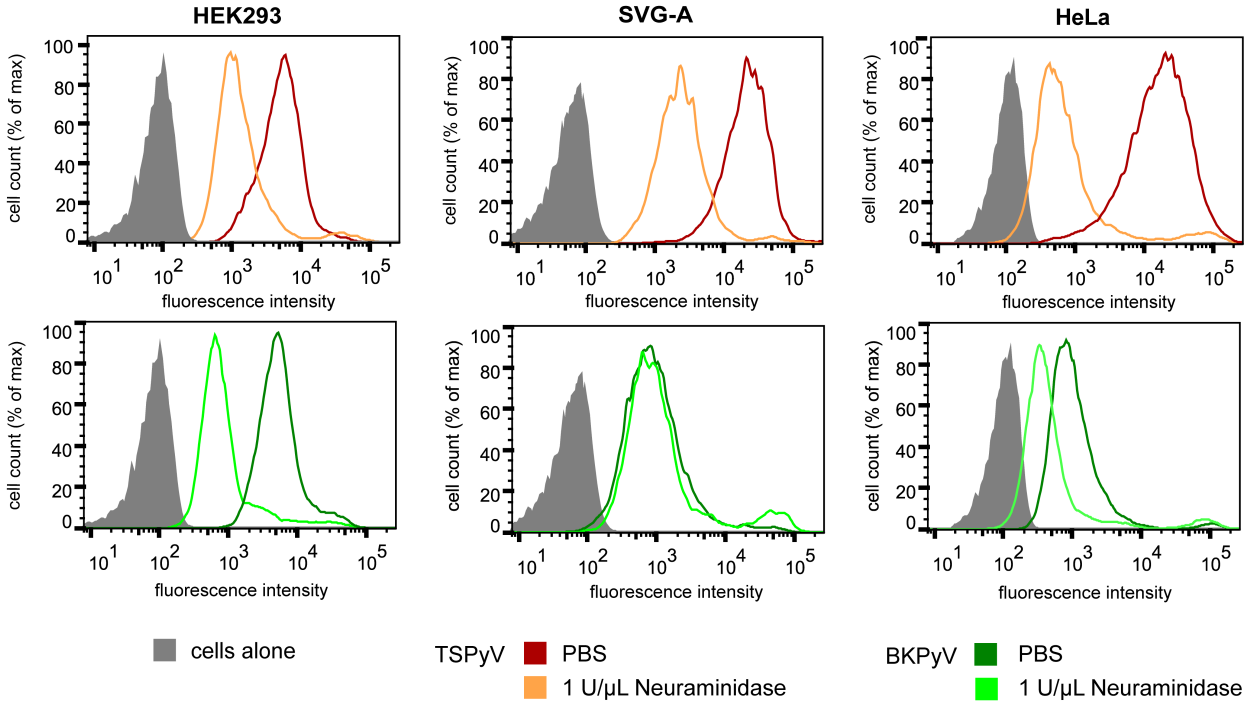
66. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallographica Section D-Biological Crystallography 53: 240–255.

67. Waterhouse AM, Procter JB, DMA Martin, Clamp M, GJ Barton (2009) Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 25: 1189–1191. doi: 10.1093/bioinformatics/btp033 PMID: 19151095

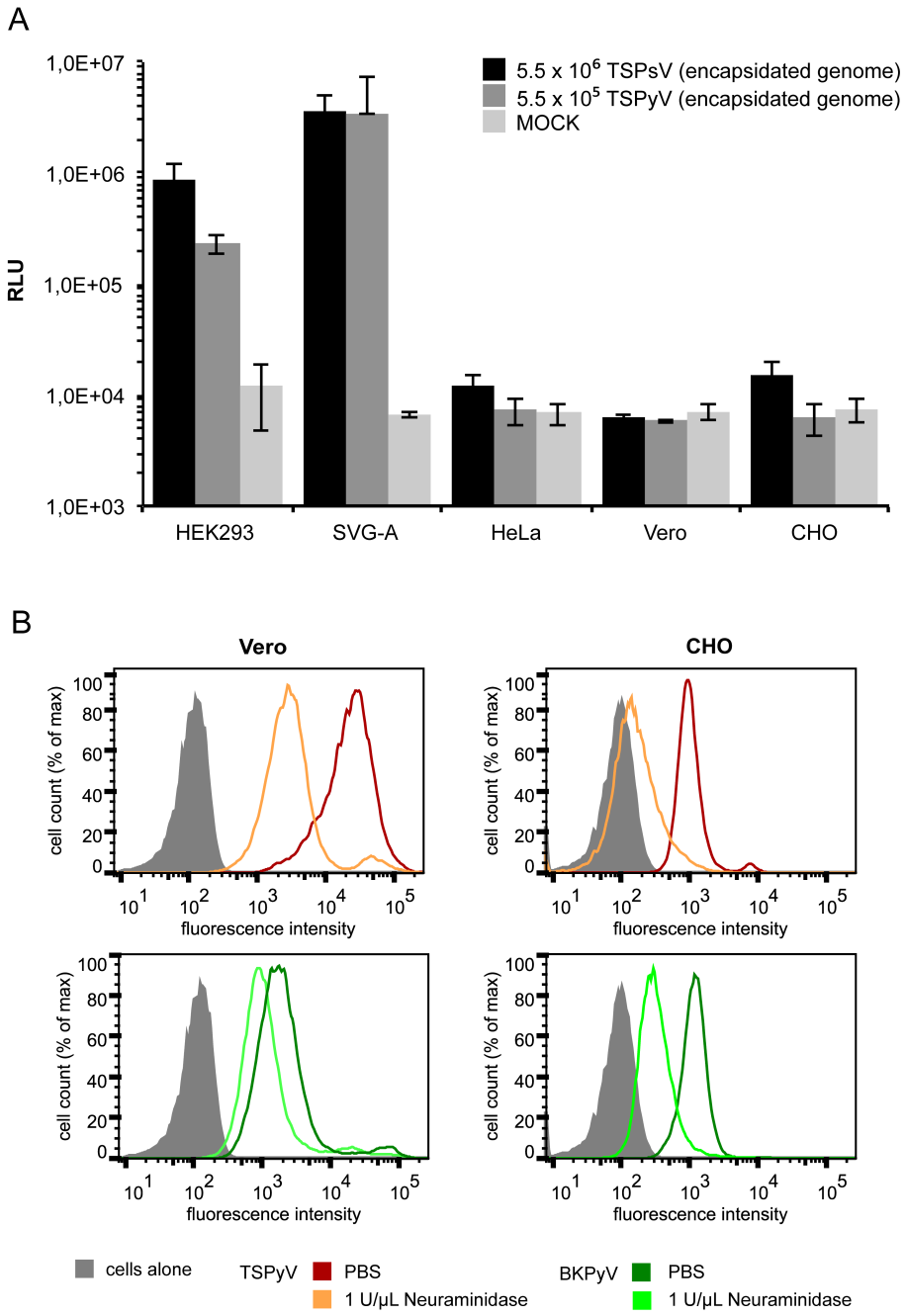
68. Nicholson JK, Foxall PJ, Spraul M, Farrant RD, Lindon JC (1995) 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Anal Chem 67: 793–811. PMID: 7762816

69. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D-Biological Crystallography 66: 12–21.

SUPPORTING INFORMATION



S1 Fig.: The binding of Alexa Fluor 488-labelled VP1 pentamers from TSPyV and BKPyV to Mock (PBS) or *Clostridium perfringens* neuraminidase type V pre-treated cells was measured by flow cytometry. Non-standardized raw data is shown here for individual experiments. To obtain the histogram in Fig. 1 B and C showing the relative binding of TSPyV VP1 pentamers data was standardized for signals of mock treated cells. The relative average fluorescence from three separated experiments is shown compared to binding to untreated cells. 30,000 gated events were measured for each sample.

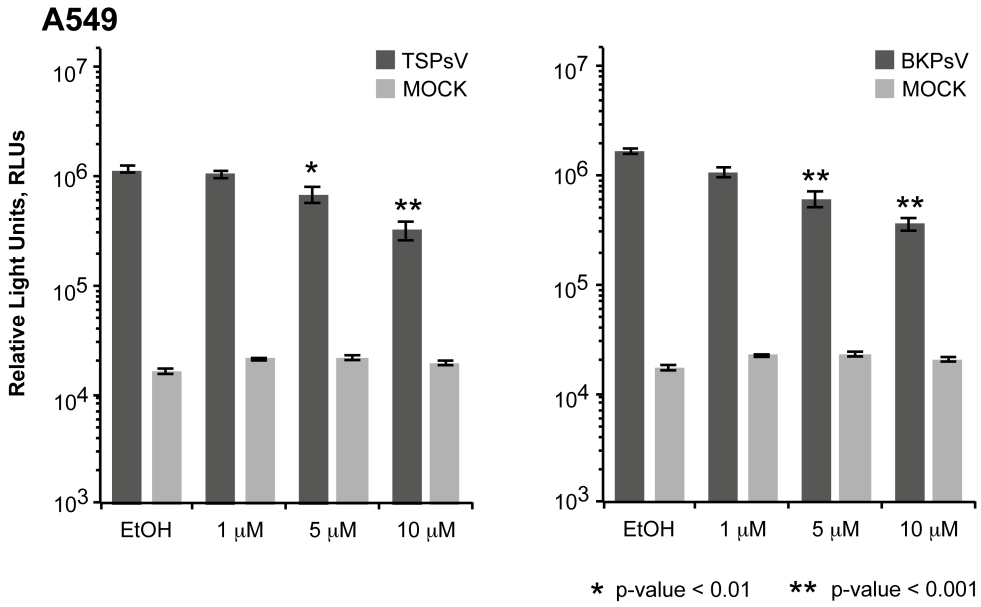


S2 Fig.: Screening cell lines for TSPsV infection. TSPyV pentamers bind to all tested cell lines in a sialic acid-dependent manner (see Fig. 1 and S1 Fig., and panel B), but TSPsV transduction could only be detected for HEK293 and SVG-A cells.

(A) Common cell lines were infected with TSPsV to obtain information about the tropism. At 72 h post infection the luciferase expression was measured to determine transfection efficiency of TSPsV. Cell lines where TSPsV infection was undetectable are in general able to express the reporter plasmid [39]. The average relative luciferase units (RLUs) for one experiment performed in quadruplet are shown in log scale. Error bars represent standard deviations. Two different dilutions of particles were used and compared to Mock PsV control. The mock PsV control was generated by harvesting HEK293TT cells transfected with control plasmid instead of the capsid expression plasmids and then purifying according to the PsV purification protocol to measure background signal and non-specific transfer of luciferase to infected cells.

(B) The binding of Alexa Fluor 488-labelled TSPyV VP1 pentamers to Mock (PBS) or *Clostridium perfringens* neuraminidase type V pre-treated Vero and CHO cells was measured

by flow cytometry. BKPyV VP1 pentamers were used as a positive control. Non-standardized raw data is shown here for representative individual experiments. Three independent experiments were performed and 30,000 gated events were measured for each sample.



S3 Fig.: Glycolipids are important for TSPsV infection in A549 cells.
TSPsV transduction of A549 cells pre-treated with PPMP or the carrier control for 6 days was assayed 72 h post infection by quantification of the secreted luciferase due to transduction of the reporter plasmid phGluc. BKPsV were used as a positive control for a ganglioside-dependent infection. Mock PsV infections were done with a control sample obtained according the PsV purification protocol from cells only transfected with phGluc and control plasmid to assess the background signal of the luciferase assay. PsV experiments were done in triplicate or quintuplicate. The data from the quintuplicate experiment is shown as a representative example. Statistic analysis was performed using the two-tailed unpaired t test.

JVI Journals.ASM.org

The Greater Affinity of JC Polyomavirus Capsid for α2,6-Linked Lactoseries Tetrasaccharide c than for Other Sialylated Glycans Is a Major Determinant of Infectivity

Luisa J. Ströh,[a] Melissa S. Maginnis,[b,c] Bärbel S. Blaum,[a] Christian D. S. Nelson,[b] Ursula Neu,[a*] Gretchen V. Gee,[b] Bethany A. O'Hara,[b] Nasim Motamedi,[d] Daniel DiMaio,[d] Walter J. Atwood,[b] Thilo Stehle[a,e]

Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany[a]; Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, USA[b]; Department of Molecular and Biomedical Sciences, University of Maine, Orono, Maine, USA[c]; Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA[d]; Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[e]

ABSTRACT

The human JC polyomavirus (JCPyV) establishes an asymptomatic, persistent infection in the kidneys of the majority of the population and is the causative agent of the fatal demyelinating disease progressive multifocal leukoencephalopathy (PML) in immunosuppressed individuals. The Mad-1 strain of JCPyV, a brain isolate, was shown earlier to require α2,6-linked sialic acid on the lactoseries tetrasaccharide c (LSTc) glycan for attachment to host cells. In contrast, a JCPyV kidney isolate type 3 strain, WT3, has been reported to interact with sialic acid-containing gangliosides, but the role of these glycans in JCPyV infection has remained unclear. To help rationalize these findings and probe the effects of strain-specific differences on receptor binding, we performed a comprehensive analysis of the glycan receptor specificities of these two representative JCPyV strains using high-resolution X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, and correlated these data with the results of infectivity assays. We show here that capsid proteins of Mad-1 and WT3 JCPyV can both engage LSTc as well as multiple sialylated gangliosides. However, the binding affinities exhibit subtle differences, with the highest affinity observed for LSTc. Engagement of LSTc is a prerequisite for functional receptor engagement, while the more weakly binding gangliosides are not required for productive infection. Our findings highlight the complexity of virus-carbohydrate interactions and demonstrate that subtle differences in binding affinities, rather than the binding event alone, help determine tissue tropism and viral pathogenesis.

IMPORTANCE

Viral infection is initiated by attachment to receptors on host cells, and this event plays an important role in viral disease. We investigated the receptor-binding properties of human JC polyomavirus (JCPyV), a virus that resides in the kidneys of the majority of the population and can cause the fatal demyelinating disease progressive multifocal leukoencephalopathy (PML) in the brains of immunosuppressed individuals. JCPyV has been reported to interact with multiple carbohydrate receptors, and we sought to clarify how the interactions between JCPyV and cellular carbohydrate receptors influenced infection. Here we demonstrate that JCPyV can engage numerous sialylated carbohydrate receptors. However, the virus displays preferential binding to LSTc, and only LSTc mediates a productive infection. Our findings demonstrate that subtle differences in binding affinity, rather than receptor engagement alone, are a key determinant of viral infection.

JC polyomavirus (JCPyV) infects ~50% of healthy individuals and causes an asymptomatic, lifelong persistent infection in the kidney (1, 2). The form of the virus that resides in the kidney is nonpathogenic and is excreted in the urine (3, 4). In immunosuppressed individuals, JCPyV can spread from the site of persistence to the central nervous system (CNS) (5, 6) and infect the glial cells astrocytes and oligodendroctyes (7, 8). Oligodendrocytes are myelin-producing cells, and astrocytes have been demonstrated to be critical to the process of CNS myelination (9–11). Infection of astrocytes and glial progenitor cells (GPCs), together with virus-induced cytolytic destruction of oligodendrocytes, causes the demyelinating disease progressive multifocal leukoencephalopathy (PML) (12–14). The disease, once considered fatal, is now managed with immune reconstitution therapy, but surviving patients remain severely debilitated (15). PML affects approximately 3 to 5% of HIV-1-positive individuals and patients receiving immunomodulatory therapies such as natalizumab and rituximab for immune-mediated diseases such as multiple sclerosis (MS), Crohn's disease, and rheumatoid arthritis (13, 16, 17). As of December 2014, there had been

Received 25 February 2015 Accepted 30 March 2015

Accepted manuscript posted online 8 April 2015

Citation Ströh LJ, Maginnis MS, Blaum BS, Nelson CDS, Neu U, Gee GV, O'Hara BA, Motamedi N, DiMaio D, Atwood WJ, Stehle T. 2015. The greater affinity of JC polyomavirus capsid for α2,6-linked lactoseries tetrasaccharide c than for other sialylated glycans is a major determinant of infectivity. J Virol 89:6364–6375. doi:10.1128/JVI.00489-15.

Editor: M. J. Imperiale

Address correspondence to Walter J. Atwood, walter_atwood@brown.edu, or Thilo Stehle, thilo.stehle@uni-tuebingen.de.

* Present address: Ursula Neu, The Francis Crick Institute, Mill Hill Laboratory, London, United Kingdom.

L.J.S. and M.S.M. contributed equally to this article.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.00489-15.

doi:10.1128/JVI.00489-15

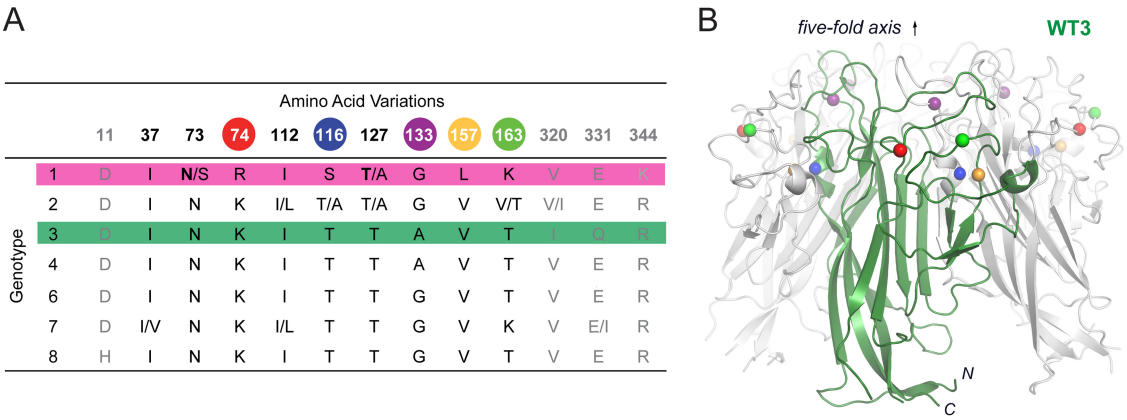


FIG 1 VP1 amino acid variations in JCPyV genotypes and crystal structure of WT3 VP1. (A) Classification of JCPyV strains into seven genotypes by phylogenetic analysis reveals VP1 amino acid variations (33). Colored spheres highlight residues differing between Mad-1 and WT3 VP1. Strain specific variations can be found within single genotypes. Bold letters indicate amino acids in Mad-1 VP1. N- and C-terminal residues (gray) are not included in the JCPyV Mad-1 VP1 crystal structure (24). (B) Crystal structure of WT3 VP1 pentamer with one VP1 monomer depicted in green. Spheres highlight residues (colored as in panel A) that differ between WT3 and Mad-1 VP1.

517 cases of natalizumab-induced PML, and 23% of these PML patients have died, since the introduction of the drug in 2004 (Biogen Idec, 2014). Newly marketed therapies such as dimethyl fumarate and other fumaric acid ester-containing drugs prescribed for MS and psoriasis have also led to PML (18, 19), indicating an urgent need to better understand the disease pathogenesis of JCPyV.

Polyomaviruses have an icosahedral, nonenveloped protein capsid that is comprised of VP1, VP2, and VP3. VP1 is a pentameric protein present as 360 copies on the surface of the capsid, which are connected to neighboring pentamers through C-terminal extensions (20). Each VP1 pentamer also interacts with a VP2 or a VP3 molecule in the interior of the capsid (21), and together they encase the double-stranded DNA (dsDNA) genome. Surface loops connecting the β-strand core structure of VP1 mediate interactions with sialic acid receptors (22–27). The Mad-1 strain of JCPyV utilizes the sialic acid receptor motif α2,6-linked lactoseries tetrasaccharide c (LSTc) for attachment (24) and then enters by clathrin-dependent endocytosis supported by the 5-hydroxytryptamine 2 (5-HT_2) family of serotonin receptors (28). Other members of the *Polyomaviridae* family, such as BK polyomavirus (BKPyV), Merkel cell polyomavirus (MCPyV), simian virus 40 (SV40), and mouse polyomavirus (MPyV), engage ganglioside receptors containing α2,3- and α2,8-linked sialic acids (23–25, 29, 30) and enter cells by caveola-dependent endocytosis (31). Gangliosides are a group of glycosphingolipids expressed on eukaryotic cells with a ceramide chain embedded in the membrane and an extracellular oligosaccharide (32).

JCPyV strains are classified into seven genotypes that possess amino acid differences at up to 13 positions within the VP1 capsid protein (33) (Fig. 1A). The JCPyV brain isolate Mad-1 strain, with a type 1 capsid, was found to specifically engage α2,6-linked sialic acid in the context of the LSTc pentasaccharide to mediate infection (24). However, other laboratories have demonstrated that virus-like particles (VLPs) of a strain referred to as WT3, with a type 3 capsid that differs from type 1 at 8 amino acid positions, are capable of engaging multiple sialic acid-containing ganglioside receptors (6). WT3 VLPs have been reported to bind to asialo-GM1, GD1a, GD1b, GD2, and GT1b (6) and GM1 and GM2 (Leonid Gorelik, Consortium for Functional Glycomics [CFG] [available at www.functionalglycomics.org]). Additionally, Mad-1 VLPs were demonstrated by virus overlay blotting assay to bind to a number of gangliosides, including GM3, GD2, GD3, GD1a, GD1b, GT1b, and GQ1b, but not GM1 or GM2 (34). However, the relevance of the interactions of JCPyV with multiple sialylated glycans in the context of JCPyV infection remained unclear.

Cellular entry of the JCPyV strain Mad-1 has been shown to be mediated by 5-HT_2 receptors (5-HT_2Rs) (28), and infection of 5-HT_2R-expressing HEK293A cells with Mad-1 and WT3 is equivalent (M. S. Maginnis and W. J. Atwood, unpublished results), suggesting that both strains utilize 5-HT_2Rs for internalization. Further, the expression of 5-HT_2Rs at the cell surface does not increase viral attachment to host cells (28). Therefore, JCPyV attachment and entry steps are likely related but distinct steps, and viral attachment to sialic acid receptors prior to entry plays a key regulatory role in infection and disease (24, 35). Viral isolates from the brains of individuals with PML contain polymorphisms in the capsid that are representative of both type 1 and type 3 strains (36), suggesting that if JCPyV utilizes multiple sialic acid-containing receptors, these polymorphisms might play a role in viral pathogenesis.

The goal of this study was to unravel the molecular details of carbohydrate receptor engagement by JCPyV in order to understand how recognition and specificity contribute to a pathogenesis that ultimately culminates in a devastating CNS disease. To clarify the role of sialic acid-containing receptors in JCPyV infection, we utilized a structure-function approach to define the role of these polymorphisms in JCPyV attachment to cellular receptors and infection of glial cells by generating WT3 VP1 in an infectious viral clone, in pseudoviruses, and in purified VP1 pentamers using the Mad-1 prototype PML strain as the backbone.

MATERIALS AND METHODS

Cells, viruses, and antibodies. SVGA cells (37) were grown in minimum essential medium (MEM) supplemented to contain 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (P-S) (Mediatech, Inc.) in a humidified incubator at 37°C with 5% CO_2. HEK293A cells (ATTC) were grown in Dulbecco's modified Eagle medium (DMEM) supplemented to

contain 10% FBS and 1% P-S. 293TT cells are derived from human embryonic kidney cells transformed with two copies of the SV40 large T antigen (a gift from Dan DiMaio's laboratory) and were grown in DMEM supplemented to contain 10% FBS, 0.1 mM nonessential amino acids (NEAA), and 250 μg/ml hygromycin B.

Generation and propagation of the virus strain Mad-1/SVEΔ was previously described (38, 39). Mad-1/WT3C virus (PubMed accession code AAQ88264) was generated by introducing the VP1 mutations R74K, S116T, G133A, L157V, and K163T into the genomic JCPyV DNA of strain JC12 with a Mad-1 VP1 (40) subcloned into pUC19 (41). Mutations R74K, S116T, G133A, and K163T were introduced by site-directed mutagenesis using the Agilent QuikChange II site-directed mutagenesis kit (Qiagen) according to the manufacturer's instructions and the following mutagenesis primers (5′ to 3′): R74K, ACATTTGAAAAGTGACTCCCC AAATAAGGACATGCTTCCTT and AAGGAAGCATGTCCTTATTTG GGGAGTCACTTTCAAATGT; S116T, AGGTTATAGGGGTGACAACT TTGATGAATGTGCACTC and GAGTGCACATTCATCAAAGTTGTC ACCCCTATAACCT; G133A, TGACAATGGTGCAGCGAAGCCAGTG CAGG and CCTGCACTGGCTTCGCTGCACCATTGTA; and K163T, GGGGTGCTTTTTAATTACAGAACAACGTACCCAGATGGAACAAT TTTT and AAAAATTGTTCCATCTGGGTACGTTGTTCTGTAATTAA AAAGCACCC. L157V was introduced by Phusion mutagenesis using the primers (5′ to 3′) ATTACAGGGGGTGGTCTTTAATTACAGAAC and TCTAAAGCCTCCCCCCCAACAGA. Sequencing was performed at Genewiz Inc.

Mad-1 and Mad-1/WT3C viruses were propagated by transfecting 2 to 4 T150 flasks of SVGA cells with 32 μg of the infectious JCPyV clone with either the Mad-1 or Mad-1/WT3C VP1 that had been digested from the pUC19 plasmid using BamHI using Fugene. At 48 h posttransfection, cells were split into roller bottles (1,650 cm^2) and incubated with complete MEM in a humidified roller cabinet at 37°C with 5% CO_2. Infected cells were incubated for 2 weeks, with the medium changed at 1 week postinfection with complete MEM. Cells were harvested by scraping, and viruses were purified as described previously (39).

GM3 siRNA treatment. siGENOME human ST3GAL5 small interfering RNAs (siRNAs) 1, 3, 4, and 17 were purchased from Thermo Scientific. The siRNA target sequences are as follows: 1, CAAUGGCGCUGUU AUUUGA; 3, GACCAUGCAUAAUGUGACA; 4, CGGAAGUUCUCCAG UAAAG; and 17, AGGAAUACUGCACGGAUUA. SignalSilence control siRNA (Cell Signaling Technology) was used as a control siRNA. SVGA cells were reverse transfected with siRNAs using RNAiMax transfection reagent (Invitrogen) by mixing transfection complexes in Opti-MEM in 24-well plates and incubating at room temperature for 20 min. Following incubation, 3×10^4 cells/well were added, and cells were incubated at 37°C for 72 h. Transfection efficiency was confirmed using BLOCK-iT Alexa Fluor Red control oligonucleotide (Life Technologies). At 72 h posttransfection, cells were either harvested for mRNA analysis or infected with JCPyV or SV40 as described below. For mRNA analysis, cells were harvested using CellStripper, washed in phosphate-buffered saline (PBS), and pelleted by centrifugation at 2,000 rpm at 4°C for 5 min. Cells were washed in PBS and pelleted, and RNA was extracted using the RNeasy minikit (Qiagen) with a DNase treatment (Qiagen). One microgram of RNA was reverse transcribed using Bio-Rad iScript cDNA synthesis according to the manufacturer's instructions, and 1 μl of the total reaction product was used for quantitative PCR (qPCR) analysis using Bio-Rad iQ SYBR green Supermix and Bio-Rad iQ5. Primers used for qPCR to detect ST3GAL5 (5′ to 3′) were CTGCCTTTGACATCCTTCAGT and CGATTGTGGGGACGTTCTTA, and those for the GAPDH (glyceraldehyde-3-phosphate dehydrogenase) housekeeping gene were AGTCAGCC GCATCTTCTTTTGC and CAATACGACCAAATCCGTTGACT. Relative ST3GAL5 gene expression was determined by calculating the change in threshold cycle (ΔC_T) relative to GAPDH (housekeeping) gene expression, and $\Delta\Delta C_T$ was calculated relative to scrambled control siRNA.

JCPyV infection. SVGA or HEK293A cells were washed following ganglioside treatment, infected with purified Mad-1 and Mad-1/WT3C JCPyV at a multiplicity of infection (MOI) of 5 focus-forming units (FFU)/cell in MEM containing 2% FBS in a volume of 100 to 200 μl, and incubated at 37°C for 1.5 h, and then complete MEM was added and cells were incubated at 37°C for 48 h (T antigen) or 72 h (VP1).

LSTc and GM1 pentasaccharide inhibition assay. Purified Mad-1 and Mad-1/WT3C were pretreated with 5 mM LSTc (V Labs, Inc.) or GM1 (Enzo Life Sciences) pentasaccharides (diluted in sterile distilled water) in medium containing 2% fetal calf serum (FCS) on ice for 1 h. SVGA cells in 96-well plates were prechilled at 4°C for 30 min. Virus-pentasaccharide complexes were added to cells and incubated at 4°C for 1 h. Cells were washed with PBS twice, complete medium was added, and cells were incubated at 37°C for 72 h. Cells were fixed and stained by indirect immunofluorescence as described below.

Indirect immunofluorescence. Cells were washed in PBS, fixed in cold methanol, and incubated at −20°C. Fixed cells were stained using PAB597, a hybridoma supernatant that produces a monoclonal antibody against JCPyV VP1 that was generously provided by Ed Harlow, as described previously (35) or using PAB962, a JCPyV-specific T antigen antibody that does not cross-react with SV40 T antigen and was provided by the Tevethia laboratory (Penn State University), as described previously (42). Cells were analyzed for nuclear VP1 or T antigen staining under a 10× or 20× objective using an Eclipse TE2000-U microscope (Nikon).

JCPyV pseudovirus production. Pseudoviruses were produced by transfection of the VP1, VP2, VP3, and phGluc plasmids into 293TT cells using FuGENE 6 transfection reagent (Promega) in a 5:1:1:1 ratio. Mock pseudovirus controls were generated by transfecting 293TT cells with pUC19 control plasmid and the phGluc reporter plasmid in a 7:1 ratio. Cells were split 1:3 at 48 h posttransfection and harvested at 7 days posttransfection, and titers for properly encapsidated genomes were determined (43).

Pseudovirus luciferase infectivity assay. Cells were plated in a 24 well or 96-well plate to 70% confluence and infected with equal particle equivalents (1×10^6 or 1×10^7 particles/well) of Mad-1 or Mad-1/WT3C pseudovirus in incomplete medium without phenol red. Infected cells were incubated at 37°C for 1.5 h and washed with PBS, and complete medium without phenol red was added and cells were incubated at 37°C for 72 h. Secreted luciferase was quantitated in 20 to 50 μl of cellular supernatants using the BioLux *Gaussia* luciferase assay (NE BioLabs) according to the manufacturer's instructions with an opaque 96-well microplate in a GloMax Multi detection system luminometer (Promega). Numbers of infected cells were also measured by quantifying green fluorescent protein (GFP)-positive cells by fluorescence microscopy using an Eclipse TE2000-U microscope (Nikon).

Protein expression, purification, and crystallization. DNA coding for amino acids 22 to 289 of VP1 from the WT3 strain (UniProtKB accession number AAQ88264), obtained by gene synthesis (Eurofins MWG Operons, Germany), was subcloned into the pET15b vector (EMD Millipore) in frame with an N-terminal hexahistidine tag (His tag) and a thrombin cleavage site using the NcoI and BamHI restriction sites (24). Mad-1 and WT3 VP1 were expressed, purified, and crystallized as described earlier (24, 44). The C-terminal truncation of VP1 prevents assembly of VP1 pentamers to capsids and other incorrectly multimeric particles, whereas the N-terminal truncation was shown to help with the crystallization in the case of MPyV VP1 (22). Prior to crystallization, the His tag was cleaved off with thrombin (GE Healthcare) in solution, resulting in a short nonnative sequence at the (GSHM) N termini of Mad-1 and WT3 VP1.

Crystals of Mad-1 and WT3 VP1 pentamers, grown in space group C2 with one VP1 pentamer in the asymmetric unit, are highly isomorphous (see Tables S1 to S7 in the supplemental material). For comparative soaking experiments, crystals grown for 2 days with similar sizes (~120 by 120 by 20 μm^3) were soaked for 30 min in the reservoir solution complemented with oligosaccharide LSTc (5, 2.5, 1.25, 0.62, or 0.31 mM; Dextra, United Kingdom), GD1a, GD1b, or GM1 (10, 5, 2.5, or 1.25 mM; Elicityl SA, France), or GM2 (20, 10, 5, 2.5, or 1.25 mM; Elicityl SA, France). For cryoprotection, crystals were transferred for 2 s into a reservoir solution

supplemented with 30% (vol/vol) glycerol and the respective ligand prior to freezing in liquid nitrogen.

Data collection and structure determination. Diffraction data sets were collected at beamlines X06DA and X06SA at SLS (Villigen, Switzerland) and processed with XDS (45). The JCPyV Mad-1 VP1 pentamer structure (RCSB Protein Data Bank [PDB] accession number 3NXD) was used for molecular replacement with Phaser MR (46), included in the CCP4 program suite (47). Rigid-body and simulated annealing refinement was carried out with Phenix (48). Alternating rounds of model building in Coot (49) and refinement with Refmac5 (50), including 5-fold noncrystallographic symmetry restraints, the translation-libration-screw method (51), and CCP4 library- and user-defined restraints for the α2,3- and α2,6-glycosidic bond were performed. Coordinates and structure factor amplitudes have been deposited in PDB for the native and complex structures of WT3 VP1 pentamer with 5 mM LSTc, 10 mM GM1, 10 mM GD1a, 10 mM GD1b, and 20 mM GM2. Deposited Mad-1 VP1 pentamer complex structures were obtained with either 10 mM GM1, 10 mM GD1a, 10 mM GD1b, or 20 mM GM2. Structure figures were prepared with PyMOL (PyMOL Molecular Graphics System, version 1.3; Schrödinger, LLC).

Calculation of ligand electron density. Unit cell parameters of all soaked crystals were treated as isomorphous. The unliganded WT3 VP1 data set was taken as a reference data set during data processing in XDS. CAD of CCP4 (47) was used to combine structure factor amplitudes and structure factor sigmas from all data sets of a concentration-dependent soaking experiment into one single file including free R flags and unit cell parameters from the reference data set (native WT3 VP1). Data sets were scaled (low- and high-resolution cutoff, 12.0 to 2.4 Å) to the reference data set with SCALEIT (52). Refinement with Refmac5 was done to obtain complex structure models in the referenced unit cell. Simulated annealing omit maps were calculated 5 Å around Neu5Ac or marker residues (Y80, Y160, and W200) with Phenix and FFT (CCP4 program suite) under each soaking condition. Masks 1 Å around the Neu5Ac moiety of the respective glycan ligand and around marker residues were generated with the CCP4 program NCSMASK (47). MAPMAN (53, 54) was used to estimate the Neu5Ac and marker residue electron densities by summation of the values for the grid points within the respective mask using the mstats command. Maps and masks were calculated with a grid spacing of $\sim d_{max}/5$, with maximum resolution d_{max}. Integrated electron densities were plotted against the ligand concentration. Data were analyzed and fitted assuming a simple equilibrium of the VP1-glycan interaction using GraphPad Prism 6 (GraphPad Software, San Diego, CA, USA). Assuming a simple equilibrium, a dissociation constants $K_{d\ crystal}$ for the interactions within the crystal can be estimated by applying a least-squares fit to the experimental data points and the equation electron density$_{observed}$ = [C_{ligand}/(C_{ligand} + $K_{d\ crystal}$)](electron density$_{max}$ − electron density$_{nat}$) + electron density$_{nat}$, where C_{ligand} is the concentration of the ligand used for the soaking experiment. Electron density$_{observed}$ is the output from the MAPMAN electron density integration, whereas electron density$_{nat}$ is the electron density at the respective position within the native crystal and electron density$_{max}$ is that observed for the glycan ligand in a fully occupancy binding site. Electron density$_{max}$ was not set to a defined value for the fit.

Saturation transfer difference (STD) NMR measurements. All nuclear magnetic resonance (NMR) spectra were recorded using 3-mm tubes on a Bruker AVIII 600-MHz spectrometer fitted with a 5-mm cryogenic probe at 285 K and processed with TOPSPIN 3.0 (Bruker). Samples were prepared in D_2O buffer containing 20 mM potassium phosphate (pH 7.4) and 150 mM NaCl. Protein concentrations were 20 μM VP1, and the GM1 glycan was added to 1 mM. A protein-free GM1 sample was used as a reference. The off- and on-resonance frequencies were set to −30 ppm and 6.9 ppm, respectively. The total relaxation delay was 3 s. A cascade of 40 Gaussian-shaped pulses with 50-ms duration each, corresponding to a strength of 65 Hz, and a saturation time of 2 s were used for selective excitation. A 10-ms continuous-wave spin lock filter with a strength of 3.7 kHz was employed. A total of 32,000 points were collected, and zero filling to 64,000 data points was employed. Spectra were multiplied with an exponential line broadening factor of 1 Hz prior to Fourier transformation.

Protein structure accession numbers. Coordinates and structure factor amplitudes were deposited in PDB for the native (accession number 4X0Y) and the complex structures of WT3 VP1 pentamer with LSTc (4X13), GM1 (4X0Z), GD1a (4X11), GD1b (4X12), and GM2 (4X10). The Mad-1 VP1 pentamer complex structures obtained with GM1 (4X14), GD1a (4X16), GD1b (4X17), and GM2 (4X15) were deposited in PDB with the indicated accession numbers.

RESULTS

VP1 amino acid variations in JCPyV genotypes and glycan receptor specificity. JCPyV strains are classified into seven genotypes that possess amino acid differences at up to 13 positions within VP1 (33) (Fig. 1A). Full-length VP1 from the prototype Mad-1 strain, a genotype 1 strain, differs at eight positions (residues 74, 116, 133, 157, 163, 320, 331, and 344) from VP1 of the urine isolate WT3 (genotype 3). Five of those residues are included in the crystal structure of the truncated, assembly-incompetent JCPyV Mad-1 VP1 pentamer comprising residues 22 to 289 (24). The remaining three residues are located in the C-terminal arms of VP1, where they help mediate assembly of the virion particle but are distant from the receptor-binding site (20). In order to assess strain- and genotype-specific glycan receptor engagement on the molecular level, we introduced WT3 polymorphisms into our VP1 construct and solved its X-ray structure (Fig. 1B; see Table S1 in the supplemental material). Structures of Mad-1 and WT3 VP1 are very similar and superimpose with root mean square deviation (RMSD) values of 0.2 Å and 0.5 Å for Cα atoms of monomers and pentamers, respectively. VP1 adopts the iconic jelly-roll fold of virus capsid proteins with two four-stranded β-sheets (strands B, I, D, and G and C, H, E and F) packed against each other, which assemble around a 5-fold axis and are linked on the outside of the virion by extended loops (24). Amino acid differences between Mad-1 and WT3 occur predominantly within or in close proximity to these surface loops (Fig. 1B).

WT3 VP1 engages sialic acids in diverse linkages. In order to analyze glycan binding specificities of WT3, we solved structures of WT3 VP1 in complex with the glycan portions of a-series gangliosides GM1, GM2, and GD1a, b-series ganglioside GD1b, and the α2,6-linked Mad-1 receptor motif LSTc (Fig. 2; see Table S1 in the supplemental material). WT3 pentamers bound to all glycans, specifically interacting with terminal sialic acids in an almost identical manner independent of the presence of an α2,3-, α2,6-, or α2,8-linkage (Fig. 2; see Fig. S1 in the supplemental material). The conserved sialic acid-binding site is located in a groove at the junction of the extended BC1, BC2, DE, and HI loops of a VP1 monomer and the BC2 and DE loops of the counterclockwise (CCW) and clockwise (CW) neighboring VP1 monomers, respectively. The majority of interactions with the five different glycans involves functional groups of Neu5Ac (described in detail in reference 24). Glycan residues of LSTc, GM1, GM2, and GD1b contribute additional interactions. VP1 residue N123 forms a hydrogen bond with the carbonyl group of the GlcNAc *N*-acetyl group and fixes LSTc in a unique L-shaped conformation on top of the pentamer (Fig. 2A). The Glc and two Gal residues contribute to the buried surface area upon ligand engagement and are well ordered in the structure due to van der Waals interactions with VP1. In contrast, the Gal-(β1,3)-GalNAc arm of the GM1 glycan points in a different direction, resting in an elongated groove on the

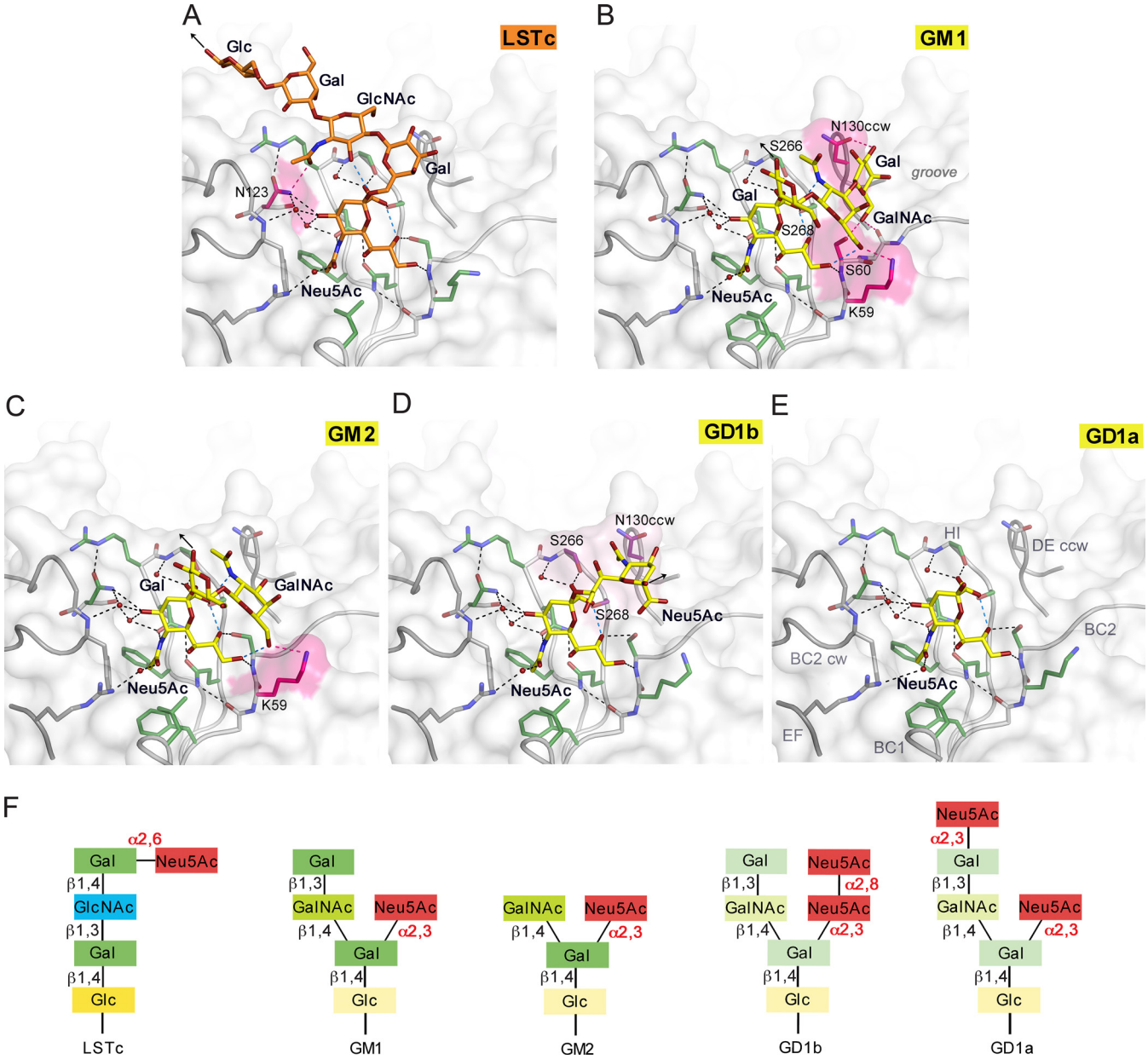


FIG 2 JCPyV WT3 interacts with α2,3-, α2,6-linked Neu5Ac and α2,8-, α2,3-di-Neu5Ac in a highly plastic binding site. (A to E) Structures of WT3 VP1 bound to LSTc (A), GM1 (B), GM2 (C), GD1b (D), and GD1a (E) glycans. VP1 is shown in surface and cartoon representation. Residues involved in polar or van der Waals interactions in only one of the glycan complex structures are highlighted in pink or purple, respectively. Interactions are depicted using dashed lines, with direct and water-mediated contacts with the glycans in black, intramolecular interactions of glycans in blue, and interaction with either LSTc, GM1, or GM2 in pink. Glycans are shown in stick representation and colored according to atom type, with nitrogens in blue, oxygens in red, and carbons in orange for LSTc and in yellow for ganglioside-derived glycans, respectively. Water molecules are shown as red spheres. Black arrows indicate the part of LSTc that is linked to a lipid or protein (A), the direction of Glc that is further linked to the lipid in the context of the GM1 and GM2 gangliosides (B and C), and the O2 of the second Neu5Ac, to which the Gal residue within the branched GD1b ganglioside motif is attached (D). (F) Glycans used for comparative binding studies. Additionally, the GT1b ganglioside was used in cell supplementation assays. GT1b features, compared to GD1b, an additional α2,3-linked Neu5Ac at the Gal in the left arm of the branched motif.

protein surface (Fig. 2B). The terminal Gal residue is recognized by hydrogen bonds to the side chains of N130 of the CCW VP1 monomer, S60, and the backbone of residue 61, as well as van der Waals interactions with N130CCW and S268 (Fig. 2B). However, this interaction is not seen in all binding sites, indicating structural flexibility of the terminal Gal. In contrast, the branched "ganglioside core" Neu5Ac-(α2,3)-(GalNAc-(β1,3))-Gal is conformationally more restrained through intramolecular interactions, and it is recognized by a hydrogen bond to K59 and van der Waals interactions with N130CW, S266, and S268. The Glc residue, which is linked to the membrane-anchoring ceramide in the context of the GM1 ganglioside, is pointing away from the protein and is not involved in contacts with VP1.

Similar interactions hold GalNAc-(β1,4)-Gal in place in the

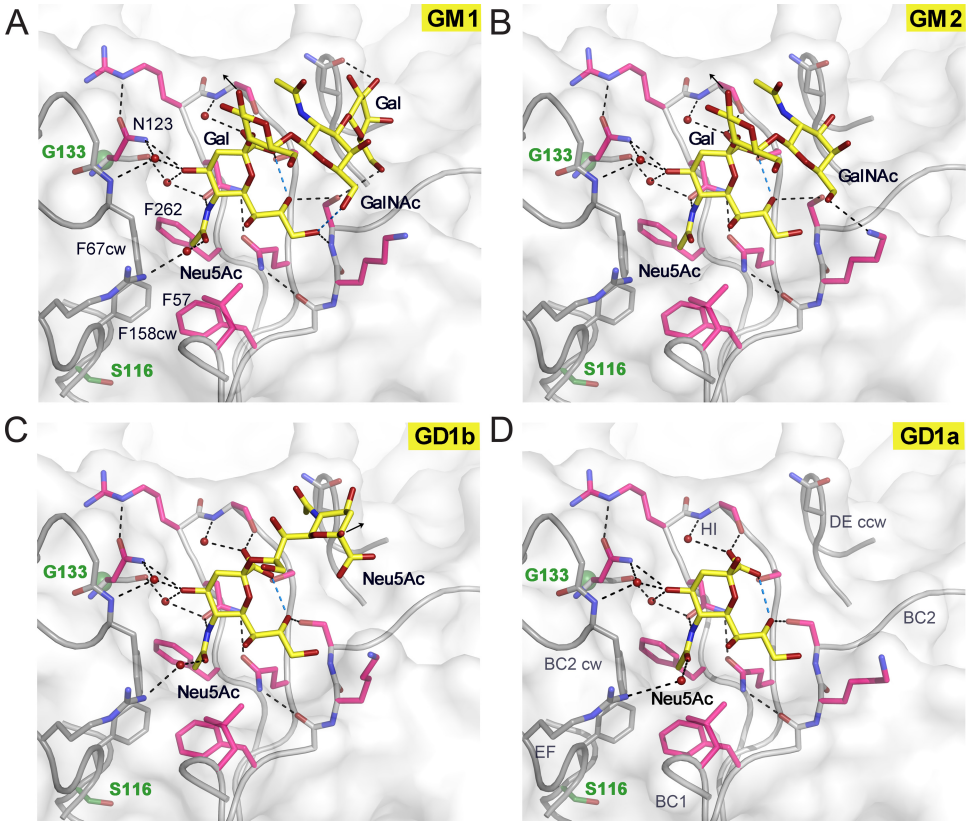


FIG 3 JCPyV Mad-1 engages α2,3-, α2,6-linked Neu5Ac and α2,8-, α2,3-di-Neu5Ac highly similarly to WT3. Complex structures of Mad-1 VP1 with GM1 (A), GM2 (B), GD1b (C), and GD1a (D) ganglioside glycans are shown. VP1 is shown in surface and cartoon representation. Interactions are depicted by dashed lines, with direct and water-mediated contacts with the glycans in black and intramolecular interactions of glycans in blue. VP1 amino acid differences between Mad-1 and WT3 are located predominantly within or close to the surface loops. Residue glycine 133 (green), which is an alanine in WT3, is located beneath N123, which undergoes an induced-fit movement to accommodate the terminal Neu5Ac during glycan engagement (21). Mad-1 residue serine 116, a threonine in WT3, is located below the hydrophobic cavity of the binding site that is formed by F67CW, F158CW, F57, and F262 and encloses the Neu5Ac *N*-acetyl group. The three additional mutations in WT3 VP1, R74K, L157V, and K163T, are distant from the binding site and are not shown in the close-up views.

GM2 complex structure, which shows interactions primarily with Neu5Ac (Fig. 2C). In the GD1b complex structure, the second Neu5Ac of the α2,8-α2,3-diNeu5Ac motif adds nonpolar interactions to the protein-glycan interaction network (Fig. 2D). The methyl group of the *N*-acetyl group is pointing against a hydrophobic platform formed by parts of the N130CW, S266, and S268 side chains. Interactions with the second arm of the branched motif GD1b are not observed. In the case of GD1a, defined electron density was observed only for a single Neu5Ac (Fig. 2E). We also solved equivalent complex structures of JCPyV Mad-1 VP1 (Fig. 3; see Table S2 and Fig. S1 in the supplemental material). Mad-1 VP1 engages GM1, GM2, GD1a, and GD1b glycans using the same contacts described above (Fig. 3 and 4).

Binding affinities for LSTc and ganglioside glycans. We hypothesized that binding affinities could modulate differences in glycan specificities between the two JCPyV strains. A comparative crystallographic approach, similar to strategies previously used for adenovirus fibers (55) and MPyV VP1 (56), was applied to obtain estimates for binding affinities. Isomorphous crystals of Mad-1 and WT3 VP1 were soaked in parallel in solutions supplemented with GM1, GM2, GD1a, GD1b, and LSTc glycans in a concentration-dependent manner, and structures were solved. With increasing glycan concentrations, simulated annealing $F_{obs} - F_{calc}$

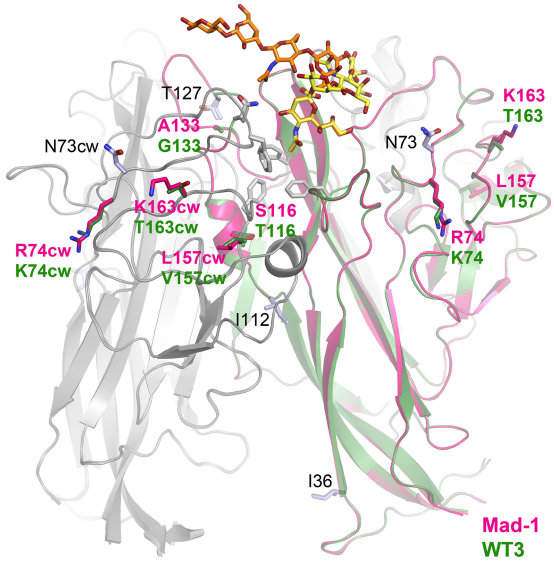


FIG 4 Strain-specific amino acid differences and engagement of GM1 and LSTc. Two VP1 monomers from JCPyV Mad-1 and WT3 are shown in cartoon representations with glycans depicted in orange (LSTc) or yellow (GM1). VP1 residues characteristic for either one of the two strains are colored in pink (Mad-1) or green (WT3). Other genotype-specific amino acid differences are shown in light blue and are labeled in black. Key residues of the glycan binding pocket, the hydrophobic cavity, and N123 are shown in gray sticks.

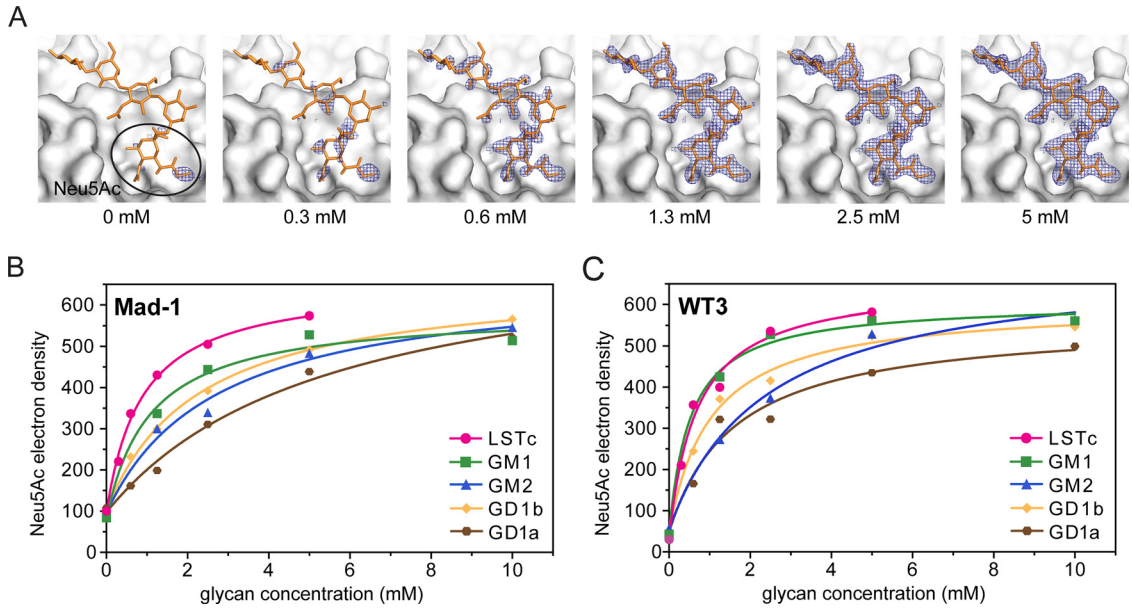


FIG 5 Mad-1 and WT3 bind to LSTc with higher affinity than gangliosides. (A) Example for the concentration-dependent crystal soaking experiment. The simulated annealing omit map contoured at 3.0 σ around 2.0 Å of LSTc at 2-4 Å resolution for JCPyV WT3 shows better defined features of the ligand with higher concentration. LSTc (yellow) is shown as a reference in all panels. Data from soaked crystals were scaled to a reference data set, and density around the Neu5Ac moiety was used for electron density integration. (B and C) Concentration-dependent crystal soaking experiments and electron density integration of $F_{obs} - F_{calc}$ simulated annealing omit maps within a 1.0-Å mask around Neu5Ac moieties of LSTc, GM1, GM2, GD1a, or GD1b within the conserved binding site of Mad-1 and WT3 VP1. Values for the electron density obtained by MAPMAN (53, 54) are plotted against the ligand concentration. Water molecules in the binding site contribute to the overall electron density of Neu5Ac when the occupancy is <1.0.

omit electron density maps reveal better-defined features of the bound ligand until saturation is reached (Fig. 5A). The electron density within a 1.0-Å mask around the terminal Neu5Ac moieties was integrated and plotted against the respective glycan concentration (Fig. 5B and C). In the outcome of this, electron density binding curves could be fitted assuming a simple equilibrium of the VP1-glycan interaction. In the nonliganded VP1, water molecules are present in the binding site, contributing to the overall electron density at the position of Neu5Ac, when the occupancy is <1.0. Dissociation constants of about 0.5 mM and 6 mM for the strongest interaction with LSTc and the weakest interaction with GD1a, respectively, could be determined. Both strains engage LSTc with the highest affinity but possess similar, lower binding affinities for GD1a, GD1b, and GM2 (Fig. 5B and C). GD1a is engaged with the lowest affinity, consistent with the observation that only a single terminal Neu5Ac is engaged in all complexes (Fig. 2E). A reduced binding affinity is also observed for GD1b, in which only the second Neu5Ac contributes additional hydrophobic contacts with VP1, and for GM2 (Fig. 2D). Of the ganglioside glycans tested, GM1 binds best to JCPyV VP1, although the binding curves differ slightly between the two strains. While the estimated dissociation constant is about 1.2 mM for Mad-1 VP, WT3 VP1 engages the GM1 glycan with a binding affinity very similar to that seen for LSTc, with a dissociation constant of about 0.7 mM.

In the crystal, only one binding site in the asymmetric unit is completely open toward the solvent, allowing engagement of all five glycans without steric hindrance or favorable interaction for a particular ligand. Therefore, a cross-validation between binding sites to estimate errors of the measurements is not possible. Electron densities for marker residues on the surface and in the hydrophobic core were instead integrated to test for the nonisomorphism of crystals and resulting differences of the electron density (see Fig. S2 in the supplemental material). The electron density of all marker atoms is clearly independent of the ligand concentrations. Therefore, measured differences in electron density values and binding affinities for the different ligands could be treated as significant.

WT3, Mad-1, and SV40 engage the GM1 glycan in a similar manner. GM1 is the functional receptor of the closely related SV40 (23, 57). We therefore compared interactions of JCPyV and SV40 with GM1 in solution using saturation transfer difference (STD) NMR spectroscopy (58). This technique can be used to map protein-bound parts of the ligand in order to analyze and compare binding epitopes. STD NMR spectra clearly show that WT3 and Mad-1 VP1 interactions with GM1 are highly similar to the functional SV40 VP1-GM1 receptor interaction (Fig. 6A). Moreover, the superposition of VP1-GM1 complex structures illustrates a conserved binding mode, with distinct differences limited to the GalNAc-(β1,3)-Gal arm of GM1 (Fig. 6B).

LSTc and GM1 glycans block JCPyV infection. For functional characterization of the strain-related VP1 polymorphisms, we generated a Mad-1/WT3C virus by introducing WT3 sequences at positions 74, 116, 133, 157, and 163 into VP1 in the genomic Mad-1 DNA. To determine if the GM1 glycan could inhibit JCPyV infection, the purified JCPyV Mad-1 and Mad-1/WT3C viruses were pretreated with GM1 or LSTc pentasaccharide prior to addition to cells. Both glycans reduced JCPyV infection of SVGA cells (Fig. 6C). Interestingly, the infection was reduced to a greater extent with LSTc than with GM1, which likely reflects the increased affinity for LSTc. Furthermore, LSTc and GM1 binding sites are overlapping, and therefore, binding of either glycan to VP1 could block binding to either LSTc or GM1.

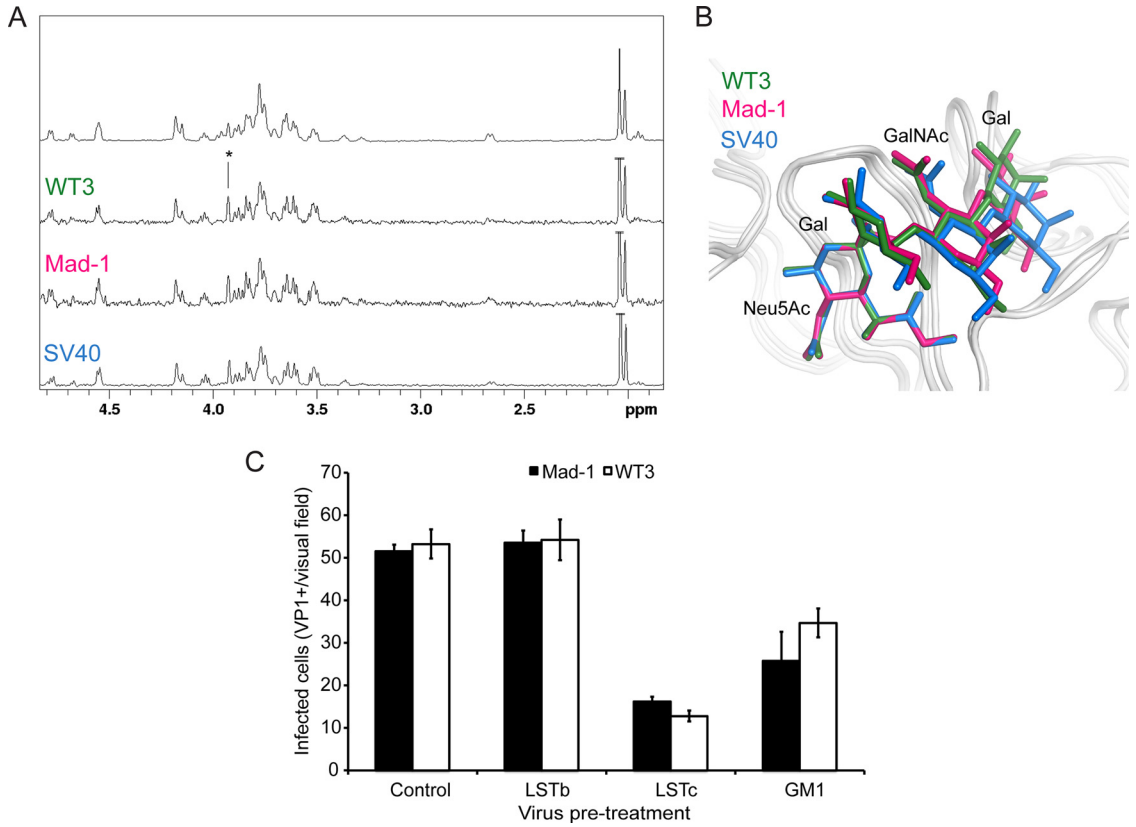


FIG 6 Interactions of JCPyV VP1 with the GM1 glycan reduces virus infection. (A) VP1 engages the GM1 glycan in solution in a manner similar to that for SV40 VP1. From top to bottom: NMR spectrum of the GM1 glycan and STD NMR difference spectra of WT3 VP1, Mad-1 VP1, and SV40 VP1 with GM1 glycan (50-fold molecular excess). (B) GM1 glycan engagement by SV40, Mad-1, and WT3 VP1. GM1 glycans are colored blue, pink, or green for SV40 (PDB 3BWR), Mad-1, and WT3, respectively. The Glc residue of GM1, which is not in contact with VP1, is not shown for clarity. (C) Pretreatment of Mad-1 or Mad-1/WT3C virus with LSTc or GM1 glycan reduces infection. Data represent the average number of infected cells per visual field (10×) for 4 fields of view of triplicate samples. Error bars indicate standard deviations (SD).

Exogenous addition of gangliosides does not enhance JCPyV infection. JCPyV Mad-1 and WT3 demonstrate affinity for gangliosides GM1, GM2, GD1a, and GD1b as determined by structural analysis of the pentamers in complex with the respective glycans. To determine whether JCPyV is able to utilize gangliosides as functional receptors for infection, SVGA cells were supplemented with a panel of gangliosides prior to infection with JCPyV Mad-1, Mad-1/WT3C, BKPyV, and SV40. BKPyV utilizes the gangliosides GD1b and GT1b as functional receptors, while SV40 utilizes GM1, and exogenous ganglioside addition has been previously demonstrated to enhance infection (25, 59). Interestingly, supplementation of cells with gangliosides GM1, GM2, GD1a, GD1b, and GT1b did not alter JCPyV infection compared to that for the dimethyl sulfoxide (DMSO)-treated control. In contrast, addition of GD1b and GT1b enhanced BKPyV infection, and GM1 supplementation enhanced SV40 infection (Fig. 7A). Given the subtle differences in affinity of JCPyV Mad-1 and Mad-1/WT3C for ganglioside-derived glycans compared to LSTc, we hypothesized that these small differences might not be appreciable in a permissive cell type that has been demonstrated to require LSTc for infection (24). Therefore, we also tested whether ganglioside supplementation had an effect in a kidney cell line that is poorly permissive for JCPyV infection (28, 42). HEK293A cells were supplemented with gangliosides overnight and then infected with JCPyV Mad-1, Mad-1/WT3C, BKPyV, and SV40. While supplementation of cells with GD1b and GT1b enhanced BKPyV infection and exogenous GM1 addition resulted in increased SV40 infection, ganglioside supplementation had no effect on JCPyV infection (Fig. 7B). Additionally, supplementation of cells with gangliosides was not toxic to cells as measured by a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophe-nyl)-2H-tetrazolium (MTS) assay (data not shown).

To extend these findings, we used the JCPyV pseudovirus system, which is more robust and has greater sensitivity. JCPyV pseudoviruses contain only the capsid components and a reporter genome, allowing for examination of only the initial steps in the virus life cycle from viral attachment through trafficking and transcription. We tested whether addition of gangliosides to SVGA (Fig. 7C) and HEK293A (data not shown) cells could enhance infection by JCPyV Mad-1, Mad-1/WT3C, and BKPyV pseudovirus infection. Addition of GD1b and GT1b enhanced BKPyV infection, while ganglioside supplementation had no effect on JCPyV infection (Fig. 7C).

Silencing of the GM3 synthase has no effect on JCPyV infection. JCPyV Mad-1 and Mad-1/WT3C are both capable of binding to a panel of ganglioside structures. However, supplementation of cells with these gangliosides neither enhances nor reduces JCPyV infection. To test whether ganglioside expression contrib-

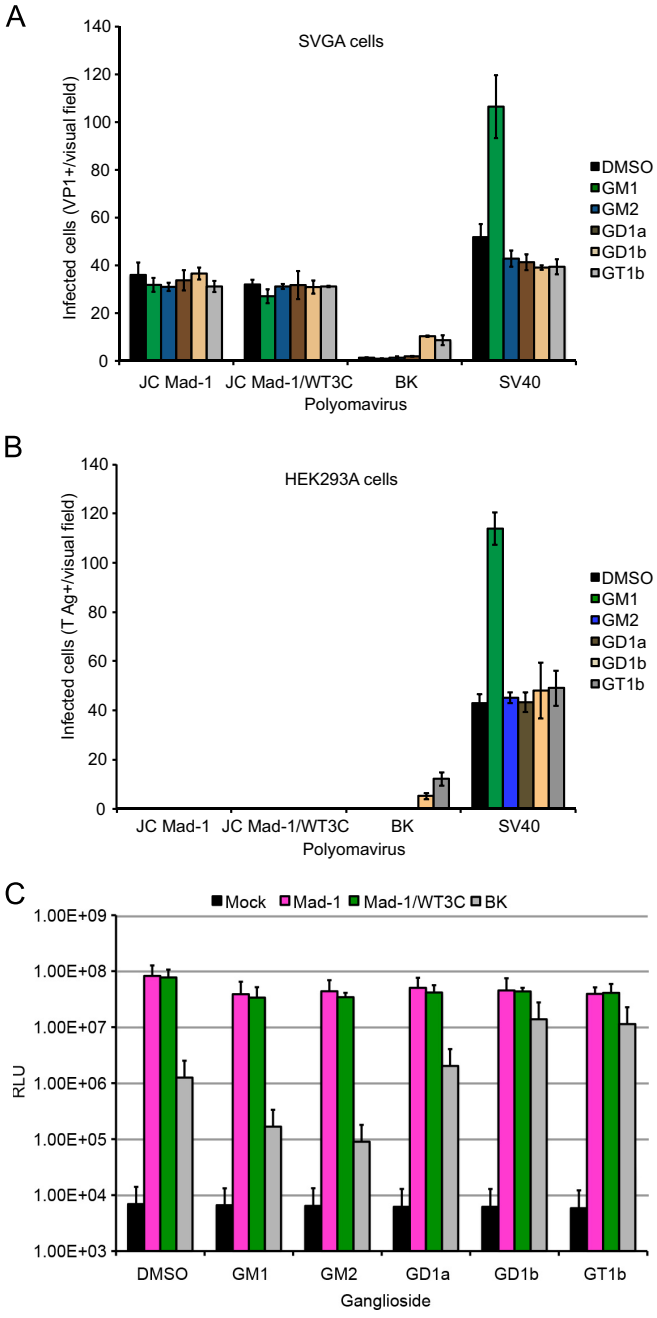


FIG 7 Exogenous addition of gangliosides does not alter JCPyV infection. (A) Permissive SVGA cells in a 24-well plate were supplemented with a 30 μM concentration of the indicated gangliosides or DMSO (vehicle control) overnight and then washed with medium and infected with JCPyV Mad-1, Mad-1/WT3C, or BKPyV and SV40 as controls. Infected cells were scored by indirect immunofluorescence for nuclear VP1 at 72 h. (B) Poorly permissive HEK293A cells were supplemented and infected as for panel A, and infected cells were scored by indirect immunofluorescence for nuclear T antigen at 48 h postinfection. Data represent the average number of infected cells per visual field for five 10× fields of view for triplicate samples. Error bars indicate SD. (C) SVGA cells were supplemented as for panel A and then infected with Mad-1, Mad-1/WT3C, or BK pseudovirus (control). Supernatant from infected cells was collected and analyzed for secreted luciferase. The average relative luciferase units (RLU) for triplicate samples are shown in log scale. Error bars represent SD. These data are representative of 3 experiments.

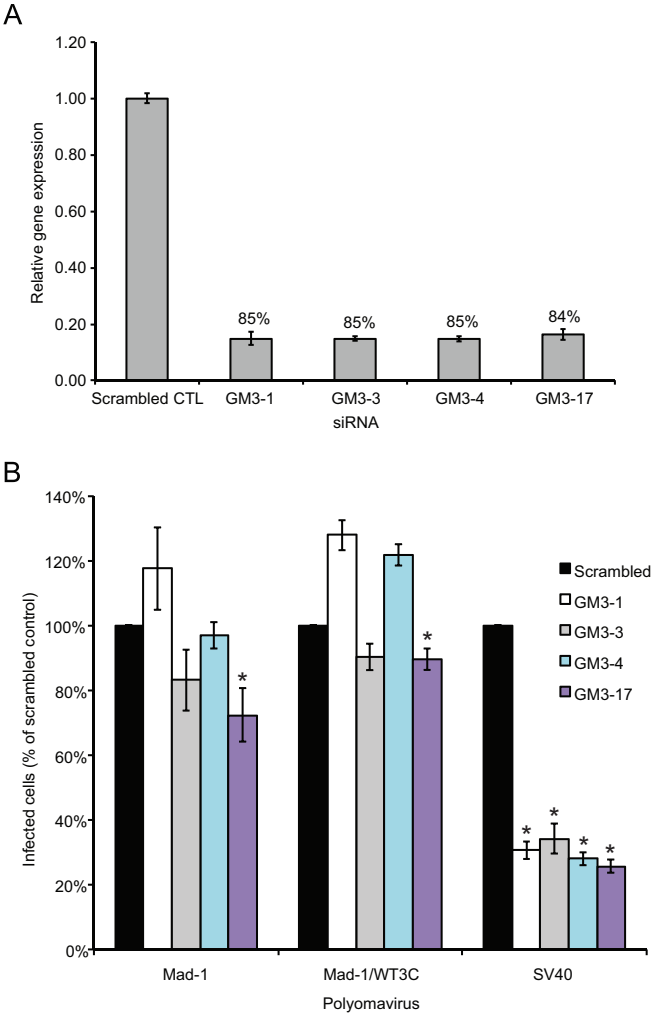


FIG 8 Gene silencing of GM3 synthase does not affect JCPyV infection in SVGA cells. SVGA cells were transfected with siRNAs targeted to the ST3GAL5 or control (scrambled) siRNA. At 72 h posttransfection, cells were either harvested for analysis of GM3 synthase mRNA expression (A) or infected (B). (A) RNA was extracted from 2 wells and reverse transcribed by reverse transcription-PCR (RT-PCR), and real-time PCR was performed with 1 μg cDNA using primers specific for GM3 synthase or GAPDH (housekeeping gene). Results are represented as the relative gene expression for triplicate real-time PCR samples relative to the GAPDH siRNA control. The percent reduction in gene expression relative to the scrambled control is expressed for each siRNA. (B) SVGA cells transfected with ST3GAL5 siRNAs or controls were infected with JCPyV or SV40. Infected cells were scored by indirect immunofluorescence. Data represent the average number of infected cells per visual field for four 10× fields of view for triplicate samples. Error bars indicate SD.

utes to JCPyV infection, GM3 synthase, the enzyme responsible for ganglioside biosynthesis, was silenced by siRNA. SVGA cells were transfected with RNA oligonucleotides specific for human ST3 β-galactoside α-2,3-sialyltransferase 5 (ST3GAL5), and the expression of GM3 synthase mRNA was determined by qPCR (Fig. 8A). The siRNA treatment of SVGA cells resulted in an ~85% reduction in GM3 mRNA expression (Fig. 8A). SVGA cells treated with siRNA to GM3 synthase were infected with JCPyV Mad-1, Mad-1/WT3C, and SV40 (Fig. 8B). Only the GM3-17 siRNA resulted in a modest reduction in infection of Mad-1 and Mad-1/WT3C, while the other siRNAs did not reduce JCPyV in-

fection. Treatment of SVGA cells with GM3 synthase siRNAs reduces SV40 infection by ~75%, indicating that SV40 infection is dependent on ganglioside expression while JCPyV infection is not.

DISCUSSION

It was previously shown that JCPyV requires the α2,6-linked LSTc glycan for attachment and that engagement of LSTc is critical for infection and spread in culture (24). In addition to the linear LSTc motif, different sialylated glycans, as well as a number of a- and b-series ganglioside structures, have been implicated as receptor candidates for JCPyV (6, 34). However, it has remained unclear whether JCPyV can utilize all of these glycans to mount an infection, particularly as different JCPyV strains and genotypes were used in different studies. To help reconcile the existing findings about JCPyV glycan receptors and strain- or genotype-specific differences, we performed a comprehensive analysis of the binding specificities and affinities of the two representative JCPyV strains Mad-1 and WT3 for a range of glycan receptors using X-ray crystallography and NMR spectroscopy, and we correlated these data with infectivity assays. Our analysis unambiguously shows that both JCPyV strains are able to bind a range of sialylated receptors, including LSTc and several ganglioside motifs. The shallow, surface-exposed JCPyV VP1 binding site allows the engagement of various sialylated ligands without steric clashes, independent from the linkage of the terminal Neu5Ac residue. However, there are critical differences in binding affinities for the analyzed glycans, with the linear α2,6-linked LSTc glycan binding best to both strains. Distinct differences in affinities correlate with infectivity assays, which show that LSTc is utilized by both strains as a functional receptor whereas the other ganglioside-type receptors are not.

We estimate the dissociation constants ($K_{d\ \mathrm{crystal}}$) for LSTc and ganglioside glycans to be in the millimolar range, i.e., 0.5 mM for the strongest interaction with LSTc and 6 mM for the weakest interaction with GD1a. Of the ganglioside glycans tested here, GM1 binds best to Mad-1 and WT3 VP1, with dissociation constants of about 1.2 and 0.7 mM, respectively. These estimated affinities are in line with affinities observed for other virus-glycan interactions (23, 55). Glycans with dissociation constants higher than 0.5 mM seem not to bind tightly enough to promote JCPyV infection. Thus, one could argue that a putative threshold for receptor binding might lie between 0.5 mM and 0.7 mM for the interaction of one glycan in one binding site. While differences in binding affinities are small, they will be amplified by a virus particle engaging a cell surface due to the availability of multiple low-affinity binding sites (60). For example, a relationship between the typically millimolar affinities for single protein-glycan interactions and femtomolar avidities observed for virus binding has been established for influenza virus, explaining how a relatively modest decrease in the affinity of one binding site can abolish virus binding (61).

As the SV40 GM1 and BKPyV GD1b binding is still conserved in JCPyV, we aimed to retarget JCPyV toward engagement of GM1 gangliosides rather than LSTc via site-directed mutagenesis at residue N123 of JCPyV. This amino acid, which is a glycine in SV40 and BKPyV (25), is the key residue that makes contacts with both arms of the L-shaped LSTc, yet potential Neu5Ac contacts would be preserved. Evaluation by crystal structure analysis revealed abolished binding to LSTc in the L-shape conformation and conserved engagement of GM1, but supplementation of cells with GM1 ganglioside did not alter cell binding or infection (data not shown). Thus, it seems that there are additional evolutionary constraints on JCPyV toward utilizing LSTc, leading JCPyV VP1 to favor this interaction over GM1 (or GD1b) ganglioside engagement. In contrast to the case for SV40 and BKPyV, JCPyV entry is achieved by clathrin-mediated rather than cholesterol-dependent endocytosis (28, 62, 63). It is tempting to speculate that the increased affinity for LSTc has coevolved with changes in the viral entry pathway of JCPyV. Only a few amino acid changes in the VP1 binding site are sufficient for switching glycan binding specificity, but the adaption of alternative attachment and entry mechanisms may require additional genetic changes. JCPyV infects only a narrow range of cell types and entry, and subsequent infection is supported by 5-HT_2 family of serotonin receptors (28). In fact, infection of 5-HT_2R-expressing HEK293A cells with viruses containing either a Mad-1 or WT3 capsid is equivalent (28; Maginnis and Atwood, unpublished results), yet exogenous addition of gangliosides in HEK293A cells, which express low levels of 5-HT_2Rs, does not lead to a productive infection (Fig. 7B), indicating that both strains of JCPyV require 5-HT_2Rs for productive infection. In contrast, BKPyV and SV40 can still switch their glycan receptor specificity through a single VP1 point mutation, but these viruses both utilize lipid-linked ganglioside receptors and share a cholesterol-dependent entry mechanism (25, 59). Binding of JCPyV to LSTc may induce organizational and/or structural changes at the plasma membrane that favor interactions with the entry receptor 5-HT_2R.

Sialic acid receptors commonly govern outcomes in viral infection and pathogenesis due to evolution and switches of receptor specificity (64). For instance, avian influenza A virus hemagglutinin (HA) attachment protein engages α2,3-linked sialic acids, which are abundant in the gastrointestinal tracts of birds, but it cannot infect humans due to the scarcity of these receptors on cells of the human respiratory tract (65). However, when a switch in sialic acid utilization from α2,3-Neu5Ac to α2,6-Neu5Ac occurs in avian or porcine reservoirs, influenza viruses from nonhuman species can be transmitted to human hosts, as α2,6-linked sialic acids are highly expressed on cells of the human respiratory tract. Influenza A virus strains with HA preference for human α2,6-linked receptors can lead to further human-to-human transmissibility and cause widespread infection in humans and viral pandemics (65). It was recently shown that certain BKPyV serotypes engage a distinct spectrum of cell surface attachment receptors, resulting in different cellular attachment mechanisms and cell tropism (66). Consequently, these findings raise questions whether different JCPyV genotypes use distinct glycan receptors besides LSTc and therefore exhibit altered tissue tropism resulting in different levels of virulence *in vivo* and are a possible risk factor for PML as proposed earlier (67). The VP1 amino acid differences in the two strains investigated here comprise the majority of genotype-specific variability (Fig. 1A). Due to the location of remaining VP1 amino acid changes not investigated in this study (Fig. 4), it is unlikely that other genotypes possess altered LSTc specificities not detected here. Consequently, genotype- and strain-related variations differ from PML-associated VP1 mutations, which are located within the central Neu5Ac binding site and render JCPyV variants noninfectious in cell culture (35). PML-associated mutations are commonly found in patients with PML in the back-

ground of different genotypes, but mechanisms by which they arise and their effects on cellular tropism remain undefined.

Our findings help explain conflicting reports on JCPyV glycan specificity. Biochemical methods such as glycan microarray screening, virus overlay blotting, and enzyme-linked immunosorbent assays to investigate glycan binding *in vitro* have variable detection limits and drawbacks. Binding alone does not indicate whether a glycan is a receptor or a pseudoreceptor. Due to the relatively low affinity of virus-glycan interactions and avidity effects, the virus and the presented glycan have a significant impact on the experimental outcome and its meaning for the physiological tissue tropism (68). Critical differences in binding affinities, rather than the binding events alone, can therefore determine receptor usage. Although a varying susceptibility of certain genotypes for PML development cannot be excluded, our study shows that it is unlikely that genotype-specific VP1 variations alter JCPyV attachment mechanisms or influence the risk for PML development. Our findings resolve a long-standing question in the field and also provide a proper framework for interpreting how interactions with viral receptors must be analyzed for functionality in infectious assays to define the outcomes on viral pathogenicity.

ACKNOWLEDGMENTS

We thank members of the Stehle, Atwood, and DiMaio laboratories for critical discussion. We also thank the staff at beamlines X06SA and X06DA of the Swiss Light Source (Villigen, Switzerland) for assistance with X-ray data collection and Remco Sprangers, MPI for Developmental Biology in Tübingen, Germany, for assistance with NMR data collection.

This work was funded by NIH grant PPG P01NS065719-16 (W.J.A., D.D., and T.S.) and Ruth L. Kirchstein National Research Service Awards F32NS064870 (M.S.M.) and F32NS070687 (C.D.S.N.) from the National Institute of Neurological Disorders and Stroke. Research reported in this project was also supported by Institutional Development Award (IDeA) NIH P20GM103423 (M.S.M.) from the National Institute of General Medical Sciences and P01CA16038 (D.D.) from the National Cancer Institute. Core facilities that support research in the Atwood laboratory are funded by P30GM103410 (W.J.A.).

REFERENCES

1. **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.
2. **Egli A, Infanti L, Dumoulin A, Buser A, Samaridis J, Stebler C, Gosert R, Hirsch HH.** 2009. Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. J Infect Dis **199:**837–846. http://dx.doi.org/10.1086/597126.
3. **Yogo Y, Kitamura T, Sugimoto C, Ueki T, Aso Y, Hara K, Taguchi F.** 1990. Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J Virol **64:**3139–3143.
4. **Daniel AM, Swenson JJ, Mayreddy RP, Khalili K, Frisque RJ.** 1996. Sequences within the early and late promoters of archetype JC virus restrict viral DNA replication and infectivity. Virology **216:**90–101. http://dx.doi.org/10.1006/viro.1996.0037.
5. **Dubois V, Dutronc H, Lafon ME, Poinsot V, Pellegrin JL, Ragnaud JM, Ferrer AM, Fleury HJ.** 1997. Latency and reactivation of JC virus in peripheral blood of human immunodeficiency virus type 1-infected patients. J Clin Microbiol **35:**2288–2292.
6. **Gorelik L, Reid C, Testa M, Brickelmaier M, Bossolasco S, Pazzi A, Bestetti A, Carmillo P, Wilson E, McAuliffe M, Tonkin C, Carulli JP, Lugovskoy A, Lazzarin A, Sunyaev S, Simon K, Cinque P.** 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. http://dx.doi.org/10.1093/infdis/jir198.
7. **Silverman L, Rubinstein LJ.** 1965. Electron microscopic observations on a case of progressive multifocal leukoencephalopathy. Acta Neuropathol **5:**215–224. http://dx.doi.org/10.1007/BF00686519.
8. **Zurhein G, Chou SM.** 1965. Particles resembling papova viruses in human cerebral demyelinating disease. Science **148:**1477–1479. http://dx.doi.org/10.1126/science.148.3676.1477.
9. **Liedtke W, Edelmann W, Bieri PL, Chiu FC, Cowan NJ, Kucherlapati R, Raine CS.** 1996. GFAP is necessary for the integrity of CNS white matter architecture and long-term maintenance of myelination. Neuron **17:**607–615. http://dx.doi.org/10.1016/S0896-6273(00)80194-4.
10. **Spiegel I, Peles E.** 2006. A new player in CNS myelination. Neuron **49:**777–778. http://dx.doi.org/10.1016/j.neuron.2006.03.001.
11. **Bradl M, Lassmann H.** 2010. Oligodendrocytes: biology and pathology. Acta Neuropathol **119:**37–53. http://dx.doi.org/10.1007/s00401-009-0601-5.
12. **Astrom KE, Mancall EL, Richardson EP, Jr.** 1958. Progressive multifocal leuko-encephalopathy; a hitherto unrecognized complication of chronic lymphatic leukaemia and Hodgkin's disease. Brain **81:**93–111. http://dx.doi.org/10.1093/brain/81.1.93.
13. **Ferenczy MW, Marshall LJ, Nelson CD, Atwood WJ, Nath A, Khalili K, Major EO.** 2012. Molecular biology, epidemiology, and pathogenesis of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. Clin Microbiol Rev **25:**471–506. http://dx.doi.org/10.1128/CMR.05031-11.
14. **Kondo Y, Windrem MS, Zou L, Chandler-Militello D, Schanz SJ, Auvergne RM, Betstadt SJ, Harrington AR, Johnson M, Kazarov A, Gorelik L, Goldman SA.** 2014. Human glial chimeric mice reveal astrocytic dependence of JC virus infection. J Clin Invest **124:**5323–5336. http://dx.doi.org/10.1172/JCI76629.
15. **Hirsch HH, Kardas P, Kranz D, Leboeuf C.** 2013. The human JC polyomavirus (JCPyV): virological background and clinical implications. APMIS **121:**685–727. http://dx.doi.org/10.1111/apm.12128.
16. **Carson KR, Focosi D, Major EO, Petrini M, Richey EA, West DP, Bennett CL.** 2009. Monoclonal antibody-associated progressive multifocal leucoencephalopathy in patients treated with rituximab, natalizumab, and efalizumab: a review from the Research on Adverse Drug Events and Reports (RADAR) Project. Lancet Oncol **10:**816–824. http://dx.doi.org/10.1016/S1470-2045(09)70161-5.
17. **Bloomgren G, Richman S, Hotermans C, Subramanyam M, Goelz S, Natarajan A, Lee S, Plavina T, Scanlon JV, Sandrock A, Bozic C.** 2012. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N Engl J Med **366:**1870–1880. http://dx.doi.org/10.1056/NEJMoa1107829.
18. **Ermis U, Weis J, Schulz JB.** 2013. Case reports of PML in patients treated for psoriasis. N Engl J Med **369:**1081. http://dx.doi.org/10.1056/NEJMc1307680.
19. **Ermis U, Weis J, Schulz JB.** 2013. PML in a patient treated with fumaric acid. N Engl J Med **368:**1657–1658. http://dx.doi.org/10.1056/NEJMc1211805.
20. **Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC.** 1991. Structure of simian virus 40 at 3.8-A resolution. Nature **354:**278–284. http://dx.doi.org/10.1038/354278a0.
21. **Chen XS, Stehle T, Harrison SC.** 1998. Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry. EMBO J **17:**3233–3240. http://dx.doi.org/10.1093/emboj/17.12.3233.
22. **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.
23. **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.
24. **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.
25. **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.
26. **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.
27. **Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, Feizi T, Stehle T.** 2014. Crystallographic and glycan microarray analysis of human polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate. J Virol **88:**6100–6111. http://dx.doi.org/10.1128/JVI.03455-13.
28. **Assetta B, Maginnis MS, Gracia Ahufinger I, Haley SA, Gee GV, Nelson CD, O'Hara BA, Allen Ramdial SA, Atwood WJ.** 2013. 5-Ht2 receptors

facilitate Jc polyomavirus entry. J Virol **87:**13490–13498. http://dx.doi.org/10.1128/JVI.02252-13.
29. **Erickson KD, Garcea RL, Tsai B.** 2009. Ganglioside GT1b is a putative host cell receptor for the Merkel cell polyomavirus. J Virol **83:**10275–10279. http://dx.doi.org/10.1128/JVI.00949-09.
30. **Qian M, Tsai B.** 2010. Lipids and proteins act in opposing manners to regulate polyomavirus infection. J Virol **84:**9840–9852. http://dx.doi.org/10.1128/JVI.01093-10.
31. **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. http://dx.doi.org/10.1016/j.virol.2008.12.021.
32. **Maccioni HJ, Quiroga R, Ferrari ML.** 2011. Cellular and molecular biology of glycosphingolipid glycosylation. J Neurochem **117:**589–602. http://dx.doi.org/10.1111/j.1471-4159.2011.07232.x.
33. **Cubitt CL, Cui X, Agostini HT, Nerurkar VR, Scheirich I, Yanagihara R, Ryschkewitsch CF, Stoner GL.** 2001. Predicted amino acid sequences for 100 JCV strains. J Neurovirol **7:**339–344. http://dx.doi.org/10.1080/13550280152537201.
34. **Komagome R, Sawa H, Suzuki T, Suzuki Y, Tanaka S, Atwood WJ, Nagashima K.** 2002. Oligosaccharides as receptors for JC virus. J Virol **76:**12992–13000. http://dx.doi.org/10.1128/JVI.76.24.12992-13000.2002.
35. **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**(3):e00247–00213. http://dx.doi.org/10.1128/mBio.00247-13.
36. **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. http://dx.doi.org/10.1371/journal.pgen.1000368.
37. **Major EO, Miller AE, Mourrain P, Traub RG, de Widt E, Sever J.** 1985. Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc Natl Acad Sci U S A **82:**1257–1261. http://dx.doi.org/10.1073/pnas.82.4.1257.
38. **Vacante DA, Traub R, Major EO.** 1989. Extension of JC virus host range to monkey cells by insertion of a simian virus 40 enhancer into the JC virus regulatory region. Virology **170:**353–361. http://dx.doi.org/10.1016/0042-6822(89)90425-X.
39. **Nelson C, Carney D, Derdowski A, Lipovsky A, Gee G, O'Hara B, Williard P, DiMaio D, Sello J, Atwood W.** 2013. A retrograde trafficking inhibitor of ricin and Shiga-like toxins inhibits infection of cells by human and monkey polyomaviruses. mBio **4**(6):00729–13. http://dx.doi.org/10.1128/mBio.00729-13.
40. **Chen BJ, Atwood WJ.** 2002. Construction of a novel JCV/SV40 hybrid virus (JCSV) reveals a role for the JCV capsid in viral tropism. Virology **300:**282–290. http://dx.doi.org/10.1006/viro.2002.1522.
41. **Gee GV, Tsomaia N, Mierke DF, Atwood WJ.** 2004. Modeling a sialic acid binding pocket in the external loops of JC virus VP1. J Biol Chem **279:**49172–49176. http://dx.doi.org/10.1074/jbc.M409326200.
42. **Maginnis MS, Haley SA, Gee GV, Atwood WJ.** 2010. Role of N-linked glycosylation of the 5-HT2A receptor in JC virus infection. J Virol **84:**9677–9684. http://dx.doi.org/10.1128/JVI.00978-10.
43. **Gee GV, O'Hara BA, Derdowski A, Atwood WJ.** 2013. Pseudovirus mimics cell entry and trafficking of the human polyomavirus JCPyV. Virus Res **178:**281–286. http://dx.doi.org/10.1016/j.virusres.2013.09.030.
44. **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.
45. **Kabsch W.** 2010. Xds. Acta Crystallogr D Biol Crystallogr **66:**125–132. http://dx.doi.org/10.1107/S0907444909047337.
46. **McCoy AJ.** 2007. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr **63:**32–41. http://dx.doi.org/10.1107/S0907444906045975.
47. **Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, 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.
48. **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.
49. **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.
50. **Murshudov GN, Vagin AA, Dodson EJ.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr **53:**240–255. http://dx.doi.org/10.1107/S0907444996012255.
51. **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.
52. **Howell PL, Smith GD.** 1992. Identification of heavy-atom derivatives by normal probability methods. J Appl Crystallogr **25:**81–86. http://dx.doi.org/10.1107/S0021889891010385.
53. **Kleywegt GJ, Jones TA.** 1996. xdlMAPMAN and xdlDATAMAN—programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. Acta Crystallogr D Biol Crystallogr **52:**826–828. http://dx.doi.org/10.1107/S0907444995014983.
54. **Kleywegt GJ.** 1999. Experimental assessment of differences between related protein crystal structures. Acta Crystallogr D Biol Crystallogr **55:**1878–1884. http://dx.doi.org/10.1107/S0907444999010495.
55. **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.
56. **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. http://dx.doi.org/10.1016/S0969-2126(96)00021-4.
57. **Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, Rapoport TA.** 2003. Gangliosides are receptors for murine polyoma virus and SV40. EMBO J **22:**4346–4355. http://dx.doi.org/10.1093/emboj/cdg439.
58. **Mayer M, Meyer B.** 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew Chem Int Ed Engl **38:**1784–1788.
59. **Magaldi TG, Buch MH, 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 VP1 alter receptor usage and cell tropism. J Virol **86:**7028–7042. http://dx.doi.org/10.1128/JVI.00371-12.
60. **Szklarczyk OM, Gonzalez-Segredo N, Kukura P, Oppenheim A, Choquet D, Sandoghdar V, Helenius A, Sbalzarini IF, Ewers H.** 2013. Receptor concentration and diffusivity control multivalent binding of Sv40 to membrane bilayers. PLoS Comput Biol **9:**e1003310. http://dx.doi.org/10.1371/journal.pcbi.1003310.
61. **Xiong X, Coombs PJ, Martin SR, Liu J, Xiao H, McCauley JW, Locher K, Walker PA, Collins PJ, Kawaoka Y, Skehel JJ, Gamblin SJ.** 2013. Receptor binding by a ferret-transmissible H5 avian influenza virus. Nature **497:**392–396. http://dx.doi.org/10.1038/nature12144.
62. **Anderson HA, Chen Y, Norkin LC.** 1996. Bound simian virus 40 translocates to caveolin-enriched membrane domains, and its entry is inhibited by drugs that selectively disrupt caveolae. Mol Biol Cell **7:**1825–1834. http://dx.doi.org/10.1091/mbc.7.11.1825.
63. **Eash S, Querbes W, Atwood WJ.** 2004. Infection of vero cells by BK virus is dependent on caveolae. J Virol **78:**11583–11590. http://dx.doi.org/10.1128/JVI.78.21.11583-11590.2004.
64. **Stroh LJ, Stehle T.** 2014. Glycan Engagement by viruses: receptor switches and specificity. Annu Rev Virol **1:**285–306. http://dx.doi.org/10.1146/annurev-virology-031413-085417.
65. **Imai M, Kawaoka Y.** 2012. The role of receptor binding specificity in interspecies transmission of influenza viruses. Curr Opin Virol **2:**160–167. http://dx.doi.org/10.1016/j.coviro.2012.03.003.
66. **Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, Buck CB.** 2013. BK polyomavirus genotypes represent distinct serotypes with distinct entry tropism. J Virol **87:**10105–10113. http://dx.doi.org/10.1128/JVI.01189-13.
67. **Agostini HT, Ryschkewitsch CF, Mory R, Singer EJ, Stoner GL.** 1997. JC virus (JCV) genotypes in brain tissue from patients with progressive multifocal leukoencephalopathy (PML) and in urine from controls without PML: increased frequency of JCV type 2 in PML. J Infect Dis **176:**1–8. http://dx.doi.org/10.1086/514010.
68. **Demarco ML, Woods RJ, Prestegard JH, Tian F.** 2010. Presentation of membrane-anchored glycosphingolipids determined from molecular dynamics simulations and NMR paramagnetic relaxation rate enhancement. J Am Chem Soc **132:**1334–1338. http://dx.doi.org/10.1021/ja907518x.

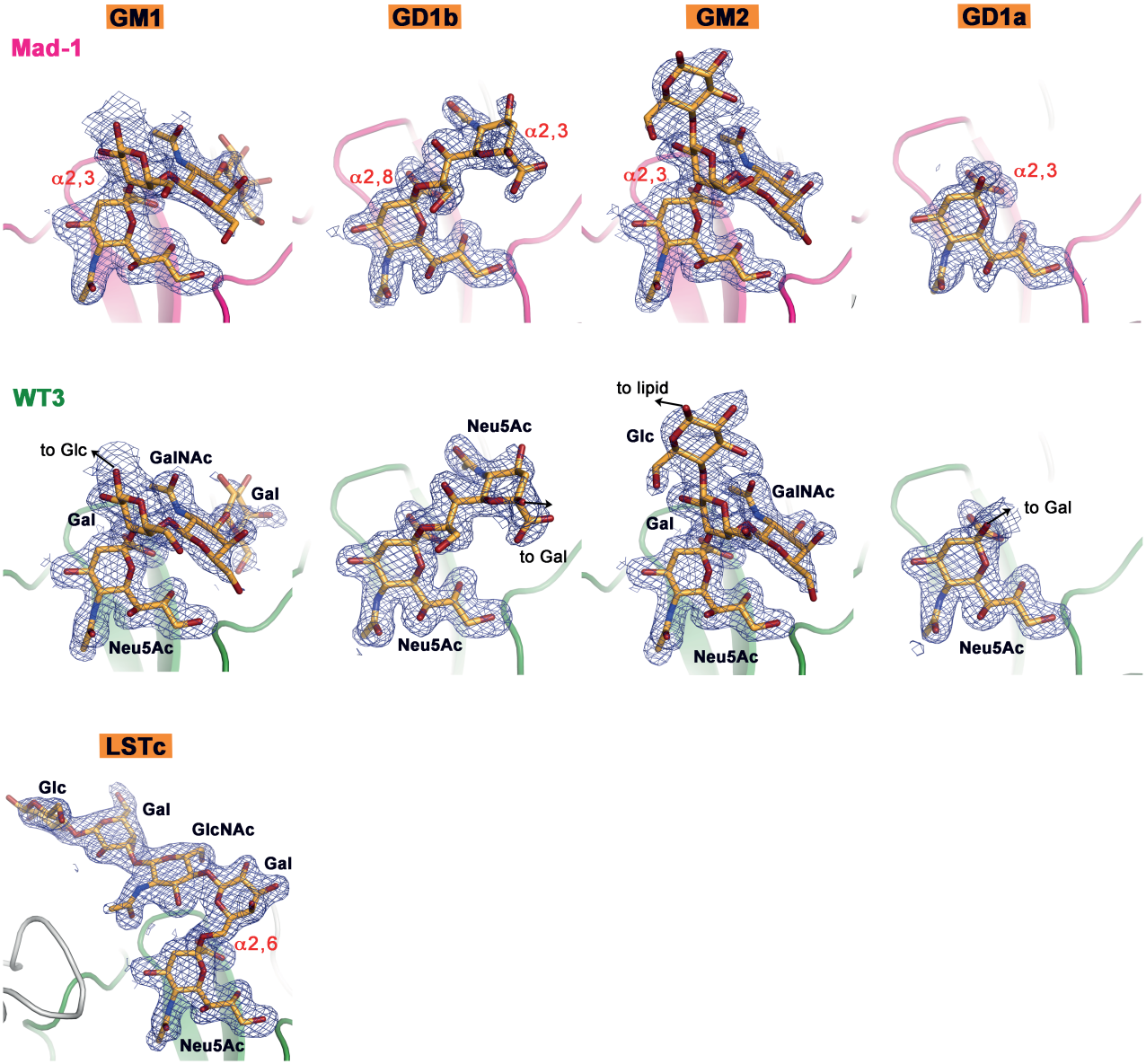


FIG. S1. Simulated annealing omit difference electron density maps for glycan complex structures of Mad-1 and WT3 VP1. Proteins are shown in cartoon representations and glycans as sticks colored according to the atom type (nitrogens in blue, oxygens in red and carbons in yellow). Fobs-Fcalc simulated annealing omit maps are shown at 3.0 sigma around 2.0 Å of the respective ligand.

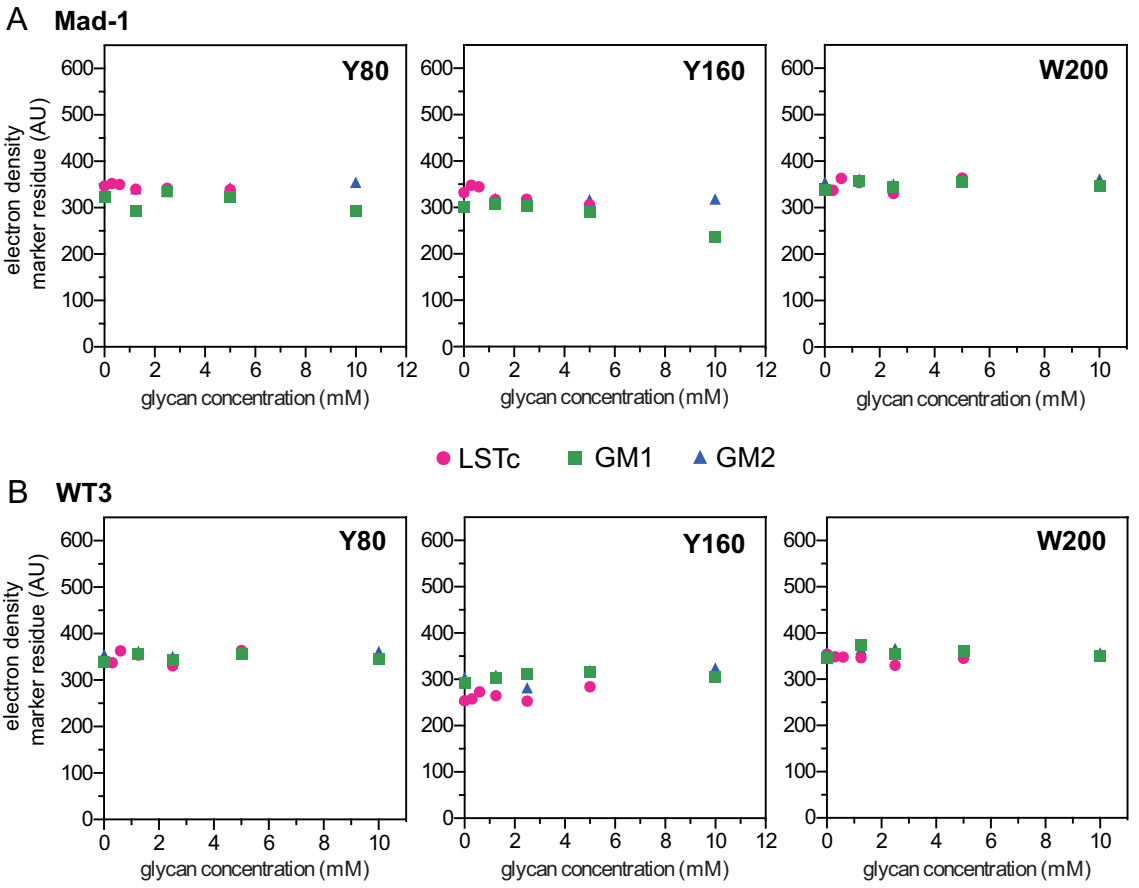


FIG. S2. Electron density integration within a 1.0 Å mask around marker residues of Mad-1 and WT3 VP1. In the JCPyV VP1 crystal environment only one binding site in the asymmetric unit is completely open towards the solvent, allowing engagement of all five ligands without steric hindrance or favorable interactions with a particular ligand. Therefore, statistic data analysis, calculation of the standard deviation, and cross-validation between binding sites was not possible for the electron density integration. Instead, Fobs-Fcalc simulated annealing omit electron densities maps for marker residues on the surface and in the hydrophobic core were integrated to test for the non-isomorphism of crystals and background differences of the electron density. Values for the electron density obtained by MAPMAN (53, 54) are plotted against the ligand concentrations, which were used in the concentration-dependent crystal soaking experiments. Representative examples are shown for experiments with LSTc, GM1, and GM2 glycans. The integrated electron density of marker atoms is independent of the ligand concentrations with standard deviations of about 1.5% to maximal 7% in rare cases.

Supplementary Tables

Table S1: Structures of native JCPyV WT3 VP1 and of JCPyV WT3 VP1 in complex with LSTc, GM1, GM2, GD1a and GD1a glycans.

Data collection and structural refinement statistics. The space group is C2 in all cases.

	WT3 VP1 native	GM1 (10 mM)	GM2 (20 mM)	GD1a (10 mM)	GD1b (10 mM)	LSTc (5 mM)
PDB accession code	4X0Y	4X0Z	4X10	4X11	4X12	4X13
Data Collection						
	X06SA	X06SA	X06SA	X06DA	X06DA	X06DA
a, *b*, *c* [Å]	150.5, 96.8, 128.2	150.3, 95.9, 128.9	150.4, 96.1, 128.8	150.1, 96.1, 128.0	150.2, 96.3, 128.33	150.1, 96.3, 128.1
β [°]	110.6	110.2	110.2	110.2	110.3	109.8
Resolution [Å]	30-1.70 (1.75-1.70)	30-1.85 (1.90-1.85)	30-1.90 (1.95-1.90)	30-2.25 (2.31-2.25)	30-1.90 (1.95-1.90)	30-2.00 (2.05-2.00)
Unique reflections	187,337 (13,674)	143,308 (10,293)	131,009 (9,366)	81,125 (5,921)	134,142 (9,504)	115,162 (8,278)
Total reflections	782,632 (57,482)	609,757 (41,002)	563,422 (38,910)	433,299 (27,562)	764,331 (46,179)	443,640 (32,140)
$CC_{1/2}$	99.2 (80.9)	99.6 (76.5)	99.4 (71.8)	99.5 (89.4)	99.6 (68.3)	99.6 (77.8)
R_{meas} [%]	12.5 (58.9)	11.2 (73.7)	13.8 (85.5)	12.3 (80.8)	13.0 (84.7)	11.5 (65.8)
I/σI	9.1 (2.6)	11.5 (2.1)	9.9 (2.0)	16.1 (2.5)	13.6 (2.1)	15.0 (2.5)
Completeness [%]	99.6 (98.4)	98.3 (96.0)	97.3 (94.3)	98.8 (89.1)	99.6 (95.9)	99.6 (97.3)
Wilson B-factor [$Å^2$]	22.1	23.3	24.8	24.9	21.8	22.6
Refinement						
Resolution [Å]	30-1.70 (1.75-1.70)	30-1.85 (1.90-1.85)	30-1.90 (1.95-1.90)	30-2.3 (2.36-2.30)	30.0-1.90 (1.95-1.90)	30-2.00 (2.05-2.00)
R_{work}/R_{free} [%]	15.7/17.8 (25.8/28.0)	16.3/19.1 (24.4/27.2)	16.4/19.6 (31.3/35.0)	16.1/20.2 (22.6/28.5)	15.8/18.7 (29.1/33.3)	15.6/19.2 (22.5/25.4)
No. of atoms						
protein / carbohydrate	10,108/-	10,026/182	10,079/160	9,866/63	10,029/123	10,022/193
water	1,274	1,284	1,181	809	1,034	1,063
Average B-Factor [$Å^2$]						
protein / carbohydrate	17.7/-	19.6/30.6	22.4/34.9	19.9/22.5	18.4/29.6	20.5/29.0
water	28.7	27.5	31.4	23.7	27.1	28.1
R.m.s. deviations						
bond length [Å]	0.009	0.010	0.008	0.009	0.008	0.009
bond angles [°]	1.408	1.437	1.350	1.320	1.336	1.356

Table S2: JCPyV Mad-1 VP1 pentamers in complex with LSTc, GM1, GM2, GD1a and GD1b glycans

Data collection and structural refinement statistics. The space group is C2 in all cases.

	GM1 (10 mM)	GM2 (20 mM)	GD1a (10 mM)	GD1b (10 mM)
PDB accession code	4X14	4X15	4X16	4X17
Data Collection				
	X06SA	X06SA	X06DA	X06DA
a, *b*, *c* [Å]	150.4, 95.6, 29.7	150.5, 95.7, 129.7	149.8, 95.6, 128.8	149.7, 95.4, 128.9
β [°]	110.2	110.3	110.3	110.4
Resolution [Å]	30-2.30 (2.36-2.30)	30-2.11 (2.16-2.11)	30-1.80 (1.85-1.80)	30-1.75 (1.80-1.75)
Unique reflections	75,740 (5,143)	99,141 (6,988)	157,148 (11,412)	170,305 (12,129)
Total reflections	318,028 (21,120)	413,312 (26,434)	892,253 (61,049)	970,603 (63,839)
$CC_{1/2}$	99.4 (79.9)	99.3 (77.7)	99.8 (73.5)	99.8 (79.9)
R_{meas} [%]	13.2 (76.9)	14.3 (75.7)	10.8 (86.9)	9.5 (71.0)
I/σI	11.2 (2.0)	9.1 (2.3)	15.4 (2.0)	15.5 (2.4)
Completeness [%]	98.8 (91.1)	99.3 (94.8)	99.7 (98.2)	99.6 (96.2)
Wilson B-factor [$Å^2$]	31.0	30.4	22.6	23.1
Refinement				
Resolution [Å]	30-2.30 (2.36-2.30)	30-2.11 (2.16-2.11)	30.0-1.80 (1.85-1.80)	30-1.75 (1.80-1.75)
R_{work}/R_{free} [%]	18.8/21.1 (35.1/36.3)	17.2/20.5 (28.7/31.0)	16.6/19.1 (29.4/32.1)	16.0/18.2 (26.9/29.4)
No. of atoms				
protein / carbohydrate	9,957/193	9,953/171	10,173/63	10,117/103
water	669	891	1,070	1,163
Average B-Factor [$Å^2$]				
protein / carbohydrate	27.2/41.8	25.4/39.1	19.4/27.9	18.7/30.8
water	29.1	32.2	28.9	29.3
R.m.s. deviations				
bond length [Å]	0.009	0.008	0.010	0.008
bond angles [°]	1.361	1.324	1.443	1.373

Table S3: Data collection statistics of data sets for the concentration-dependent density integration of LSTc. Crystals have the space group C2.

VP1	**WT3**	**WT3**	**WT3**	**WT3**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**
LSTc (mM)	0.31	0.63	1.25	2.5	-	0.31	0.63	1.25	2.5	5.0
beamline	X06DA	X06DA	X06DA	X06DA	X06SA	X06DA	X06DA	X06DA	X06DA	X06DA
Unit cell										
a, *b*, *c* [Å]	150.3, 96.8, 128.0	150.0, 96.5, 127.8	150.0, 96.9, 127.9	150.2, 96.7, 127.8	150.2, 96.7, 128.8	149.5, 97.3, 129.1	150.2, 96.5, 128.7	150.0, 96.4, 128.2	150.0, 96.5, 128.4	150.0, 96.4, 128.6
β [°]	110.5	110.3	110.2	110.1	110.5	110.2	110.4	110.4	110.4	110.3
Resolution [Å]	30-1.80 (1.85-1.80)	30-1.80 (1.85-1.80)	30-1.80 (1.85-1.80)	30-2.10 (2.15-2.10)	30-2.00 (2.06-2.00)	30-2.00 (2.05-2.00)	30-1.85 (1.90-185)	30-2.00 (2.05-2.00)	30-2.00 (2.05-2.00)	30-2.00 (2.05-2.00)
Unique reflections	158,177 (11,321)	157,331 (11,411)	157,639 (11,353)	99,844 (7,182)	114,410 (8,216)	115,891 (7,552)	145,564 (10,393)	115,158 (8,442)	113,914 (7,948)	115,576 (8,494)
Total reflections	662,748 (46,154)	662,524 (43,426)	667,378 (46,437)	424,630 (30,167)	482,593 (34,735)	481,090 (28,957)	616,941 (43,853)	486,667 (32,835)	488,740 (31,561)	488,619 (33,443)
$CC_{1/2}$	99.7 (70.4)	99.6 (77.0)	99.7 (79.7)	99.2 (75.2)	99.2 (75.4)	99.5 (80.3)	99.7 (72.6)	99.4 (72.2)	99.5 (74.2)	99.4 (64.3)
R_{meas} [%]	11.3 (76.4)	10.9 (59.8)	9.8 (62.1)	15.5 (68.9)	15.6 (74.2)	13.5 (77.9)	11.4 (77.5)	15.7 (78.5)	14.1 (71.5)	14.3 (80.7)
I/σI	12.6 (2.1)	11.7 (2.6)	14.0 (2.6)	11.7 (2.5)	8.5 (2.2)	9.5 (2.1)	11.2 (2.2)	9.7 (2.2)	10.7 (2.3)	10.5 (2.00)
Completeness [%]	99.6 (96.8)	99.7 (98.0)	99.6 (97.5)	99.6 (97.4)	98.7 (95.9)	98.9 (87.6)	99.5 (96.2)	99.6 (98.7)	98.1 (92.3)	99.7 (98.8)
Wilson B-factor [$Å^2$]	21.9	21.0	21.7	21.8	26.3	27.2	25.6	25.8	25.4	25.3

Table S4: Data collection statistics of data sets used for the concentration-dependent density integration of the GM1 glycan. Crystals have the space group C2.

VP1	**WT3**	**WT3**	**WT3**	**Mad-1**	**Mad-1**	**Mad-1**
GM1 (mM)	1.25	2.5	5	1.25	2.5	5
beamline	X06SA	X06SA	X06SA	X06SA	X06SA	X06SA
Unit cell						
a, *b*, *c* [Å]	150.4, 96.8, 128.5	150.5, 96.2, 129.0	150.6, 96.6, 129.3	150.4, 96.3, 128.9	150.4, 95.9, 129.5	150.0, 95.5,129.2
β [°]	110.3	110.3	110.3	110.5	110.4	110.2
Resolution [Å]	30-1.80 (1.85-1.80)	30-1.80 (1.85-1.80)	30-1.80 (1.85-1.81)	30-1.80 (1.85-1.80)	30-2.00 (2.06-2.00)	30-1.80 (1.85-1.80)
Unique reflections	158,084 (11,327)	158,032 (11,485)	158,601 (11,373)	155,805 (11,186)	114,858 (8,177)	154,151 (11,336)
Total reflections	659,532 (45,540)	660,134 (47,014)	662,871 (46,212)	661,296 (46,700)	481,966 (33,639)	639,240 (46,812)
$CC_{1/2}$	98.5 74.5	99.1 (71.0)	99.4 (80.8)	98.6 (81.6)	99.4 (78.8)	99.6 (82.1)
R_{meas} [%]	16.0 (77.5)	14.2 (76.3)	11.9 (64.1)	14.3 (55.0)	14.4 (77.2)	10.6 (65.3)
I/σI	7.0 (1.9)	7.9 (2.0)	9.4 (2.4)	7.7 (2.8)	8.9 (2.2)	11.0 (2.5)
Completeness [%]	99.5 (97.0)	99.7 (98.0)	99.5 (96.9)	98.1 (95.3)	99.3 (96.0)	97.9 (97.5)
Wilson B-factor [Å^2]	24.0	24.3	23.8	26.6	27.9	24.7

Table S5: Data collection statistics of the additional data sets used for the concentration dependent density integration for the GM2 glycan. Crystals have the space group C2.

VP1	**WT3**	**WT3**	**WT3**	**WT3**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**
GM2 (mM)	1.25	2.5	5	10	1.25	2.5	5	10
beamline	X06SA	X06SA	X06SA	X06SA	X06SA	X06SA	X06SA	X06SA
Unit cell								
a, *b*, *c* [Å]	150.3, 96.4, 128.2	151.3, 97.2, 128.8	150.8, 96.3, 128.6	150.9, 96.5, 128.7	150.1, 96.0, 128.9	150.2, 96.1, 129.0	150.3, 96.0, 129.3	150.1, 95.6, 129.3
β [°]	110.4	110.3	110.3	110.4	110.4	110.4	110.4	110.4
Resolution [Å]	30-1.80 (1.85-1.80)	30-2.11 (2.16-2.11)	30-2.00 (2.06-2.01)	30-1.70 (1.75-1.70	30-1.80 (1.85-1.80)	30-1.90 (1.95-1.90)	30-1.80 (1.85-1.80)	30-1.80 (1.85-1.80)
Unique reflections	152,248 (10,494)	99,824 (6,994)	115,134 (8,255)	188,271 (13,716)	154,257 (11,103)	134,315 (9,581)	157,972 (1,581)	157,529 (11,343)
Total reflections	637,365 (43,358)	414,098 (27,757)	481,083 (34,136)	788,337 (56,634)	663,743 (48,147)	563,375 (39,404)	664,477 (48,821)	662,083 (47,766)
$CC_{1/2}$	99.3 (82.9)	99.3 (78.5)	99.5 (79.7)	97.2 (69.1)	99.6 (75.7)	99.7 (78.8)	99.2 (79.7)	98.4 (86.2)
R_{meas} [%]	13.0 (61.4)	13.1 (63.6)	14.0 (73.1)	19.8 (76.7)	10.3 (76.3)	10.6 (73.3)	12.7 (60.8)	15.4 (51.9)
I/σI	8.8 (2.5)	9.7 (2.6)	11.48 (2.66)	5.4 (1.9)	10.7 (2.2)	11.4 (2.2)	8.4 (2.4)	7.1 (2.7)
Completeness [%]	96.0 (89.8)	99.4 (94.7)	99.5 (96.5)	99.7 (98.6)	97.6 (95.0)	99.5 (96.7)	99.7 (98.7)	99.6 (97.2)
Wilson B-factor [$Å^2$]	22.3	28.2	24.2	24.7	26.7	26.1	25.2	25.1

Table S6: Data collection statistics of the additional data sets used for the concentration dependent density integration for the GD1a glycan. Crystals have the space group C2.

VP1	**WT3**	**WT3**	**WT3**	**WT3**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**
GD1a (mM)	0.63	1.25	2.5	5	0.63	1.25	2.5	5
beamline	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA
Unit cell								
a, *b*, *c* [Å]	150.2, 96.0, 128.1	149.9, 95.7, 128.1	150.5, 96.1, 128.5	150.3, 96.2, 128.2	150.1, 96.2, 128.5	149.8, 95.5, 128.5	149.8, 95.8, 128.7	149.9, 95.8, 128.8
β [°]	110.3	110.3	110.4	110.2	110.4	110.3	110.4	110.4
Resolution [Å]	30-2.05 (2.10-2.05)	30-1.85 (1.90-1.85)	30.0-1.90 (1.95-1.90)	30-2.40 (2.46-2.40)	30-1.60 (1.64-1.60)	30-1.65 (1.69-1.65)	30-1.90 (1.95-1.90)	30-1.75 (1.80-1.75)
Unique reflections	106,508 (7,661)	144,017 (10,301)	134,218 (9,235)	66,652 (4,604)	224,814 (16,388)	200,663 (14,146)	133,940 (9,734)	171,089 (12,331)
Total reflections	608,355 (40,545)	826,671 (52,526)	767,567 (49,522)	376,400 (23,692)	1,269,968 (86,946)	1,052,395 (68,886)	764,945 (48,815)	953,738 (62,373)
$CC_{1/2}$	99.3 (79.8)	99.7 (81.8)	99.7 (76.2)	99.2 (78.8)	99.9 (79.5)	99.9 (85.2)	99.7 (76.7)	99.8 (80.4)
R_{meas} [Å]	16.5 (79.7)	12.3 (12.3)	13.3 (90.7)	15.3 (74.0)	7.0 (75.7)	6.8 (64.6)	12.1 (81.5)	9.7 (77.7)
I/σI	13.0 (2.3)	14.04 (2.5)	13.5 (2.0)	14.4 (2.6)	19.1 (2.5)	18.9 (2.6)	14.7 (2.2)	15.1 (2.3)
Completeness [%]	99.6 (97.6)	99.6 (96.4)	99.3 (92.9)	99.2 (92.6)	99.7 (98.5)	98.5 (94.6)	99.7 (98.6)	99.7 (97.6)
Wilson B-factor [$Å^2$]	21.1	21.1	22.7	21.4	22.4	22.5	22.4	22.5

Table S7: Data collection statistics of the additional data sets used for the concentration dependent density integration for the GD1b glycan. Crystals have the space group C2.

VP1	**WT3**	**WT3**	**WT3**	**WT3**	**Mad-1**	**Mad-1**	**Mad-1**	**Mad-1**
GD1b (mM)	0.63	1.25	2.5	5	0.63	1.25	2.5	5
Data Collection	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA	X06DA
Unit cell								
a, *b*, *c* [Å]	145.0, 96.2, 128.0	150.4, 96.3, 128.4	150.2, 96.3, 128.2	150.1, 96.4, 128.2	149.8, 95.8, 128.4	149.8, 95.8, 128.4	149.8, 95.6, 128.8	149.9, 95.4, 128.9
β [°]	110.4	110.3	110.6	110.3	110.4	110.4	110.3	110.3
Resolution [Å]	30-1.85 (1.90-1.85)	30-1.95 (2.00- 1.95)	30-1.85 (1.90-1.85)	30-1.90 (1.95-1.90)	30-1.75 (1.80-1.75)	30-1.80 (1.85- 1.80)	30-1.90 (1.95-1.90)	30-2.10 (2.15-2.10)
Unique reflections	144,880 (10,544)	123,742 (8,470)	144,803 (10,370)	134,448 (9,647)	166,353 (11,910)	157,123 (11,513)	132,768 (9,444)	98,777 (7,210)
Total reflections	823,303 (54,911)	707,337 (45,494)	832,599 (53,184)	765,085 (47,900)	932,942 (61,102)	896,684 (58,683)	711,947 (43,508)	536,551 (35,871)
$CC_{1/2}$	99.5 (78.7)	99.6 (79.9)	99.7 (75.4)	99.6 (71.3)	99.8 (80.6)	99.8 (82.9)	99.5 (75.3)	99.5 (83.0)
R_{meas} (%)	15.3 (84.8)	15.2 (90.4)	12.2 (78.1)	13.1 (79.2)	9.6 (67.3)	10.3 (77.0)	13.5 (80.0)	13.3 (69.4)
I/σI	10.5 (2.3)	13.7 (2.8)	14.2 (2.3)	13.6 (2.2)	16.0 (2.6)	16.2 (2.4)	12.3 (2.1)	14.1 (2.6)
Completeness [%]	99.8 (98.5)	99.3 (92.3)	99.7 (96.9)	99.7 (97.4)	97.1 (94.4)	99.8 (98.8)	99.0 (95.5)	99.2 (97.8)
Wilson B-factor [$Å^2$]	21.2	23.9	22.5	21.7	22.6	22.1	22.0	23.6

Annotation for Tables S1-S7:

Values for the highest resolution bin are given in parentheses.

$CC_{1/2}$: correlation between intensities from random half-dataset

$R_{meas} = \frac{\sum_{hkl}\sqrt{\frac{n}{n-1}}\ \sum_{j=1}^{n}\left|I_{hkl,j}-\langle I_{hkl,j}\rangle\right|}{\sum_{hkl}\sum_{j}^{n} I_{hkl,j}}$, with n is the number of observation of the reflection and $\langle I_{h,k,l}\rangle$ the intensity of symmetry- (or Friedel-) related observations.

$R_{work} = \frac{\sum_{hkl}|F_{obs}(hkl)-F_{calc}(hkl)|}{\sum_{hkl}F_{obs}(hkl)}$, F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. 5% of reflections were not used during structure refinement to calculate R_{free}.

JVI Journals.ASM.org

Modulation of a Pore in the Capsid of JC Polyomavirus Reduces Infectivity and Prevents Exposure of the Minor Capsid Proteins

Christian D. S. Nelson,[a] Luisa J. Ströh,[b] Gretchen V. Gee,[a] Bethany A. O'Hara,[a] Thilo Stehle,[b,c] Walter J. Atwood[a]

Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, USA[a]; Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany[b]; Department of Pediatrics, Vanderbilt University, School of Medicine, Nashville, Tennessee, USA[c]

ABSTRACT

JC polyomavirus (JCPyV) infection of immunocompromised individuals results in the fatal demyelinating disease progressive multifocal leukoencephalopathy (PML). The viral capsid of JCPyV is composed primarily of the major capsid protein virus protein 1 (VP1), and pentameric arrangement of VP1 monomers results in the formation of a pore at the 5-fold axis of symmetry. While the presence of this pore is conserved among polyomaviruses, its functional role in infection or assembly is unknown. Here, we investigate the role of the 5-fold pore in assembly and infection of JCPyV by generating a panel of mutant viruses containing amino acid substitutions of the residues lining this pore. Multicycle growth assays demonstrated that the fitness of all mutants was reduced compared to that of the wild-type virus. Bacterial expression of VP1 pentamers containing substitutions to residues lining the 5-fold pore did not affect pentamer assembly or prevent association with the VP2 minor capsid protein. The X-ray crystal structures of selected pore mutants contained subtle changes to the 5-fold pore, and no other changes to VP1 were observed. Pore mutant pseudoviruses were not deficient in assembly, packaging of the minor capsid proteins, or binding to cells or in transport to the host cell endoplasmic reticulum. Instead, these mutant viruses were unable to expose VP2 upon arrival to the endoplasmic reticulum, a step that is critical for infection. This study demonstrated that the 5-fold pore is an important structural feature of JCPyV and that minor modifications to this structure have significant impacts on infectious entry.

IMPORTANCE

JCPyV is an important human pathogen that causes a severe neurological disease in immunocompromised individuals. While the high-resolution X-ray structure of the major capsid protein of JCPyV has been solved, the importance of a major structural feature of the capsid, the 5-fold pore, remains poorly understood. This pore is conserved across polyomaviruses and suggests either that these viruses have limited structural plasticity in this region or that this pore is important in infection or assembly. Using a structure-guided mutational approach, we showed that modulation of this pore severely inhibits JCPyV infection. These mutants do not appear deficient in assembly or early steps in infectious entry and are instead reduced in their ability to expose a minor capsid protein in the host cell endoplasmic reticulum. Our work demonstrates that the 5-fold pore is an important structural feature for JCPyV.

The JC polyomavirus (JCPyV) is a common human pathogen that causes severe disease in immunocompromised individuals. Serological studies have demonstrated that primary infection occurs early in childhood, and it is estimated that over 50% of the adult population is seropositive for JCPyV (1). In healthy individuals, JCPyV has been reported to establish a persistent infection in tissues of the kidney, bone marrow, and B cells. However, conditions of immunosuppression result in dissemination of JCPyV to the central nervous system, leading to a lytic infection of astrocytes and oligodendrocytes (reviewed in reference 2). Destruction of oligodendrocytes results in demyelination of neurons and causes the fatal demyelinating disease progressive multifocal leukoencephalopathy (PML) (3). There are currently no effective treatments for PML.

JCPyV is a small nonenveloped double-stranded DNA virus with a capsid diameter of approximately 50 nm. The viral capsid is composed of the major capsid protein viral protein 1 (VP1) and minor capsid proteins VP2 and VP3 (4). The exterior surface of the virus is formed by 360 copies of VP1 that are arrayed in T=7d pentameric symmetry to form 72 pentamers (5). VP1 also contains residues that are necessary for binding to the sialic acid-containing motif lactoseries tetrasaccharide c (LSTc) (6). A prominent feature of the JCPyV pentamer is a cavity located at the 5-fold axis of symmetry (6). One copy of either VP2 or VP3 is incorporated into each VP1 pentamer and is sequestered within the viral capsid. In the X-ray costructure of mouse polyomavirus (MPyV) VP1 and a VP2 fragment, electron density corresponding to the C terminus of VP2 was observed bound to the wall of the interior axial cavity of VP1 (7). Thus, it is likely that VP2 occupies much of the space of the interior axial cavity of the VP1 pentamer. The axial cavity connects to the exterior of the viral capsid through a 12.5-Å opening, referred to in this paper as the 5-fold pore. The importance of this pore in entry or assembly of JCPyV is unknown.

Received 13 January 2015 **Accepted** 15 January 2015

Accepted manuscript posted online 21 January 2015

Citation Nelson CDS, Ströh LJ, Gee GV, O'Hara BA, Stehle T, Atwood WJ. 2015. Modulation of a pore in the capsid of JC polyomavirus reduces infectivity and prevents exposure of the minor capsid proteins. J Virol 89:3910–3921. doi:10.1128/JVI.00089-15.

Editor: M. J. Imperiale

Address correspondence to Walter J. Atwood, Walter_Atwood@brown.edu.

doi:10.1128/JVI.00089-15

JCPyV utilizes complex intracellular transport pathways to gain access to the host cell nucleus for viral replication. JCPyV initiates infection by binding to LSTc on the surface of cells. This virus-receptor interaction is highly specific, as closely related molecules such as LSTb do not bind to JCPyV (6). JCPyV then enters cells by clathrin-mediated endocytosis, and the human serotonin receptor 2A family has been reported to enhance uptake of virus (8–11). After endocytosis, JCPyV enters the endocytic system and associates with Rab5-positive endosomes prior to retrograde transport to the endoplasmic reticulum (ER) (12, 13). Upon delivery to the ER, it is likely that JCPyV depends on isomerization of intrapentameric disulfide bonds to allow partial uncoating of the virus, as pharmacological treatments that reduce ER function or knockdown of the ER chaperones PDI and ERp57 decrease infectivity (13). The VP2 minor capsid protein of JCPyV is also exposed upon delivery to the ER (14). From there, JCPyV is likely retrotranslocated into the cytoplasm prior to nuclear transport and the initiation of viral replication.

Exposure of VP2 is a critical aspect of viral entry, and JCPyV mutants that lack VP2 or VP3 are noninfectious (15). Studies in MPyV or simian virus 40 (SV40) have demonstrated that VP2 and VP3 bind to cellular membranes or liposomes following exposure (16, 17). Additionally, VP2 is able to perforate membranes, likely by the insertion of additional predicted transmembrane regions or the myristolyated N terminus of VP2 (18). Antibodies raised to the N terminus of VP2 were able to prevent association of virions with liposomes, suggesting that the N terminus of VP2 mediates association of SV40 with membranes (16). Despite this importance, the mechanism by which VP2 is exposed from within the viral capsid remains unclear.

In this study, we sought to define the role of the 5-fold pore in JCPyV infection. JCPyV mutants were generated that contained amino acid substitutions to residues lining this pore and were expected to reduce the diameter of this pore. We demonstrated that alteration of this pore significantly reduces viral infectivity but does not interfere with viral assembly or packaging of VP2 and VP3. Using JCPyV pseudoviruses (PSVs) that contain these mutations, we show that these mutant PSVs are defective in transducing cells, despite showing no morphological changes compared to wild-type PSVs. Finally, these mutant PSVs are defective in exposing VP2 in the ER of cells, despite being transported to these compartments. These data suggest that amino acid substitutions to residues lining the 5-fold pore prevent exposure of the minor capsid proteins.

MATERIALS AND METHODS

Viruses, cells, and plasmids. The Mad-1 strain of JCPyV and the human fetal glial (SVGA) cells used in these experiments have been previously described (19, 20). SVGA cells were maintained in complete media (minimum essential media [MEM] supplemented with 10% fetal bovine serum, 0.5% penicillin, and 0.5% streptomycin). 293TT cells were maintained in Dulbecco's modified Eagle medium that was supplemented with 10% fetal bovine serum and MEM nonessential amino acids (Cellgro).

Western blotting. Protein samples were loaded onto 12% SDS-PAGE gels (Mini-Protean TGX; Bio-Rad) and transferred to nitrocellulose membranes (Trans-Blot SD semidry transfer cell; Bio-Rad). Membranes were blocked with 1% (wt/vol) bovine serum albumin–phosphate-buffered saline (PBS). VP1 was then detected using a mouse monoclonal antibody to VP1 (PAB597; a kind gift from Ed Harlow), and VP2 and VP3 were detected using a rabbit polyclonal antibody to SV40 VP2 that is cross-reactive with VP3 (ab53983; Abcam). Both antibodies were diluted in PBS containing 0.05% Tween 20 (PBS-T) and incubated overnight at 4°C. Membranes were washed in PBS-T and then incubated with secondary antibodies to mouse (goat anti-mouse Alexa Fluor 800; Life Technologies) or rabbit (goat anti-rabbit Alexa Fluor 680; Life Technologies). Antibodies were diluted 1:5,000 in PBS-T and incubated for 1 h at 21°C. Membranes were washed three times in PBS-T and once in PBS, and fluorescence was visualized and analyzed using an Odyssey infrared imaging system (Li-COR).

PSV production. Viral genes were codon optimized for optimal expression in the 293TT human-derived cell line using the National Cancer Institute Center for Cancer Research Laboratory of Cellular Oncology Technical Files (http://home.ccr.cancer.gov/LCO/production.asp). Codon-optimized sequences were based on the Mad-1 *Polyomaviridae* orthopolyomavirus strain (NC_001699.1) and commercially synthesized (Blue Heron Biotech). These genes were then subcloned into pwP expression vector in place of the MPyV VP1 gene (22519; Addgene). PSVs were produced by transfection of the codon-optimized VP1, VP2, VP3, and phGluc plasmids into 293TT cells using FuGENE 6 transfection reagent (Promega) in a 3:1:1:2 ratio. Cells were split 1:3 at 48 h posttransfection (hpi) and harvested at 7 d posttransfection by scraping and then pelleted and resuspended in buffer A with EDTA-free protease inhibitors (Roche Applied Science). Cells were lysed by three rounds of freezing and thawing, sonicated and treated with 0.25% deoxycholic acid at 37°C for 30 min, and then sonicated three times on ice (Branson) (50% amplitude, 50% duty cycle, power 4, 1 min). The pH was lowered to 6.0, and the lysates were treated with type V neuraminidase (Sigma) at 37°C for 1 h to release JCPyV still bound to cells. The pH was then raised to 7.5, cellular debris was pelleted by centrifugation, and the viral supernatant was pelleted through a 20% sucrose cushion in a SW40ti rotor (Beckman Coulter) at 150,000 × *g* at 4°C for 3 h. The viral pellet was resuspended into buffer A (10 mM Tris-HCl [pH 7.4], 50 mM NaCl, 0.1 mM $CaCl_2$) and sonicated 3 times (30% amplitude, 50% duty cycle, power 3, 1 min). The resuspended pellet was loaded onto a CsCl step gradient (1.29 to 1.35 g/ml) and spun at 115,000 × *g* at 4°C for 18 h in a SW55ti rotor (Beckman Coulter). The band corresponding to DNA-containing virions was isolated and dialyzed against buffer A.

PSVs were quantified by absolute genome copy numbers. Purified virions were treated with DNase I, and encapsidated phGluc plasmid was extracted using a DNeasy blood and tissue kit (Qiagen). Absolute quantification was performed using TaqMan quantitative PCR (Applied Biosystems), based on a standard curve of phGluc plasmid. The number of copies for the known plasmid was plotted in a scatter plot against the threshold cycle (C_T) value determined for each dilution. A best-fit line was generated, and the trend line equation from regression analysis was used to calculate the absolute number of genomes in each virus sample.

PSV luciferase infectivity assay. SVGA cells were seeded in 96-well plates at 70% confluence and infected with 1×10^7 genomes/well of wild-type and mutant PSV in 35 μl of serum-free medium. Cells were incubated at 37°C for 1 h and washed with PBS, complete medium without phenol red was added, and cells were incubated at 37°C for 72 h.

Secreted luciferase was quantitated in 50 μl of cellular supernatants using the BioLux Gaussia luciferase assay (New England BioLabs) according to the manufacturer's instructions using an opaque 96-well microplate in a GloMax multidetection system luminometer (Promega) equipped with an autoinjector.

Mutagenesis. An infectious clone of JCPyV (Mad-1 strain, pUC19 backbone) was mutagenized by Phusion site-directed mutagenesis according to the protocol of the manufacturer (New England BioLabs). The mutagenic primers used in this study are listed in Table 1. The entire genome of each mutant was sequenced to ensure that no erroneous point mutations were generated during the PCR amplification. Mutagenesis of VP1 pentamers was performed using a pET15b plasmid encoding JCPyV VP1 (residues 22 to 289) from the Mad-1 strain and an N-terminal hexahistidine tag (6). The same primers were used to mutagenize this pentamer expression plasmid as were used for the infectious clone. Mutants

TABLE 1 Primers used to generate pore mutations

Mutation	Orientation	Primer sequence (5′–3′)
P223V	F	GGAGAAAATGTTGUACCAGTTCTTC
	R	TCCTGTTAGTGTCCCAAAATATCTTG
P223I	F	GGAGAAAATGTTATCCCAGTTCTTC
	R	TCCTGTTAGTGTCCCAAAATATCTTG
P223L	F	GGAGAAAATGTTCTACCAGTTCTTC
	R	TCCTGTTAGTGTCCCAAAATATCTTG
P223M	F	GGAGAAAATGTTATGCCAGTTCTTC
	R	TCCTGTTAGTGTCCCAAAATATCTTG
N221Q	F	GGAGGAGAACAAGTTCCTCCAG
	R	TGTTAGTGTCCCAAAATATCTTGTG
N221Y	F	GGAGGAGAATACGTTCCTCCAG
	R	TGTTAGTGTCCCAAAATATCTTGTG
N221W	F	GGAGGAGAATGGGTTCCTCCAG
	R	TGTTAGTGTCCCAAAATATCTTGTG
Q137Y	F	GCCAGTGTACGGCACCAG
	R	TTCCCTGCACCATTGTCATGAG
137W	F	GCCAGTGTGGGGCACCAG
	R	TTCCCTGCACCATTGTCATGAG

were maintained in the JM109 cells, and *Escherichia coli* BL21(DE3) cells were used for pentamer expression (Invitrogen).

Pentamer expression and purification. Purification of JCPyV pentamers has previously been described (6). Briefly, 2 liters of Luria broth was inoculated with *E. coli* BL21(DE3) cells containing a pET15b-derived plasmid expressing wild-type or mutant VP1. These cultures were grown at 37°C to an optical density of 0.8. These cultures were then induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside and grown at 21°C for 18 h. Cultures were pelleted by centrifugation, resuspended in 20 ml of loading buffer (50 mM Tris-HCl [pH 7.5], 250 mM NaCl, 10 mM imidazole, 5% [vol/vol] glycerol), and frozen in liquid nitrogen. After thawing, the pellet was sonicated for 1 min on ice using a 50% duty cycle and 50% amplitude followed by a 5-min cooldown. After five rounds of sonication, bacterial debris was separated from the supernatant centrifugation at 17,640 × *g* for 20 min at 4°C. The supernatant was filtered through a 0.22-μm-pore-size filter and purified by immobilized metal ion affinity chromatography (IMAC) using a HisTrap 5-ml column and an Akta Purifier system (GE Healthcare). Pentamers were eluted from the column with a linear gradient of 10 to 500 mM imidazole, and fractions corresponding to pentamers were collected and dialyzed against a reaction mixture consisting of 20 mM Tris base (pH 7.5), 5% (vol/vol) glycerol, 250 mM NaCl, and 10 mM dithiothreitol (DTT). The sample was then concentrated, and nonaggregated VP1 pentamers were purified by size exclusion chromatography (SEC) using a Superdex S200 column and eluted in 20 mM HEPES (pH 7.5)–150 mM NaCl–50 mM DTT (GE Healthcare). All samples were concentrated to 1 mg/ml using a molar extinction coefficient of 28,030 M^{-1} cm^{-1} at 280 nm. For crystallization, the N-terminal hexahistidine tag was cleaved off in solution by incubation with 10 U thrombin (GE Healthcare) per mg VP1 for 20 h at 20°C prior to size exclusion chromatography (Superdex S200 column).

Coexpression and purification of VP1 and VP2 pentamers. Full-length JCPyV VP2 lacking affinity tags was cloned into the pACYC-duet plasmid. *E. coli* BL21(DE3) cells were then cotransformed with full-length VP2 and wild-type or mutant VP1. VP1 and VP2 pentamers were then purified as described for the VP1-only pentamers.

Reconstitution of mutants and multicycle growth assay. The infectious clone of each mutant was digested with EcoRI to liberate the viral genome from the plasmid backbone. SVGA cells were seeded onto 18-mm-diameter coverslips, and 1 μg of this digestion product was transfected using Fugene 6, according to the instructions of the manufacturer (Roche). Wild-type JCPyV was included as a positive control, and pUC19 alone was included as a negative control. Replicate samples were then assayed every 3 days, and the number of infected cells was scored to determine viral spread. At the time points specified, replicates for each mutant were fixed with methanol and stained with a monoclonal antibody against JCPyV VP1 (PAb 597). A goat anti-mouse Alexa 488 antibody was used as a secondary antibody, and nuclei were counterstained with DAPI (4′,6-diamidino-2-phenylindole; Life Technologies). Positive cells were counted on a T2000E inverted fluorescence microscope (Nikon), using the DAPI filter block to find a field of representative cells. Five representative fields for each replicate were counted, and samples were processed in a blind manner to minimize bias.

Crystallization. JCPyV VP1 pentamers with an N221W, N221Q, or P223M mutation were concentrated to 4.5 mg/ml in 20 mM HEPES (pH 7.5)–150 mM NaCl. Crystals were set up at 20°C using the sitting-drop vapor diffusion technique and a reservoir solution containing 100 mM HEPES (pH 7.5), 200 mM KSCN, and 12% (wt/vol) polyethylene glycol (PEG) 3350. Drops (1 μl protein solution and 1 μl reservoir solution) were cross-seeded by adding microseeds obtained from previously grown JCPyV Mad-1 VP1 crystals. Crystals were flash-cooled in liquid nitrogen using a harvesting solution supplemented with 30% (vol/vol) glycerol.

Data collection and structure determination. Diffraction data were collected at beamline X06DA at the Swiss Light Source (Villigen, Switzerland) and processed with XDS (21). Structures were solved by molecular replacement with Phaser (22) in CCP4 (23) using the JCPyV Mad-1 VP1 pentamer structure (PDB: 3NXG) lacking a solvent molecule(s) as a search model. In order to remove model bias, rigid-body and simulated annealing refinement was carried out with Phenix (24). Crystals of mutant VP1 pentamers have the same space group and very similar unit cell parameters. Mutations were introduced into the respective structural models, water molecules were added, and alternating rounds of model building in Coot (25) were performed. Refmac5 (26) was used for restrained refinement, including 5-fold NCS restraints and the translation-libration-screw (TLS) method (27). Simulated annealing F_{obs}-F_{calc} omit electron density maps were calculated after refinement with Phenix by omitting the respective mutated amino acid residue for the validation of the final structural models. Coordinates and structure factor amplitudes have been deposited with the RCSB Protein Data Bank (www.pdb.org). Structural characteristics and radii of the 5-fold VP1 pore tunnels were visualized and analyzed with CAVER 3.0 (28). Structure figures were prepared with PyMOL (The PyMOL Molecular Graphics System, version 1.3; Schrödinger LLC). All amino acids are numbered without the inclusion of the N-terminal methionine, consistent with the previously described JCPyV VP1 structures (6, 29).

For the P223L mutant, a homology model was manually prepared in the Coot program by selecting possible rotamer conformations of P223L. Additionally, homology modeling was carried out with the program MODELLER (30). MODELLER implements comparative protein structure modeling via energy minimization and spatial restraints. The two models are highly similar and feature a diameter more similar to those seen with the wild type, N221W, and N221Q. However, due to the surrounding of the P223L, only a more uncommon side chain rotamer of L223 is allowed in the context of VP1 without steric hindrance.

Differential scanning fluorometry. In order to assess the stability of wild-type and mutant pentamers, differential scanning fluorometry was performed using Sypro Orange, as previously described (31). Briefly, 4 μM VP1 monomers (0.8 μM VP1 pentamers) and 20 mM HEPES (pH 7.5)–150 mM NaCl–50 mM dithiothreitol (DTT) were mixed with a 10× solution of Sypro Orange in a final volume of 50 μl (Sypro Orange is provided as a 5,000× solution). Fluorescence was then monitored at the

indicated temperatures in a Bio-Rad IQ5 real-time PCR machine, using the excitation filter for 6-carboxyfluorescein (FAM) and emission filter for ROX (Bio-Rad). A blank sample containing buffer but not pentamers was used as a negative control. The fluorescence from the negative control was subtracted from that measured for each sample. Six replicates were performed for each sample.

Negative-staining electron microscopy. PSV virion formation was determined using negative-staining transmission electron microscopy (TEM). A 5-μl sample of each purified PSV was adsorbed onto carbon-coated Formvar copper grids for 2 min. Grids were washed with water, and 5 μl of nanoW, a tungsten-based pH neutral stain, was added to each grid for 1 min. Excess stain was blotted, and samples were visualized on a Phillips 410 transmission electron microscope. Micrographs were acquired at 80 kV and ×112,000 magnification.

Flow cytometry. SVGA cells were seeded into 6-well plates at a density of 5×10^5 cells per well and incubated overnight. The following day, cells were detached using a nonenzymatic dissociation agent (Cellstripper; Cellgro) and incubated on ice for 30 min. For each PSV, an equivalent number of genomes was then added to each sample and allowed to incubate on ice for 2 h. Samples were then washed in cold PBS and immunostained with a fluorescence-conjugated monoclonal antibody to VP1 (PAB597-488). After incubation for 1 h on ice, excess antibody was washed away with PBS and samples were read on a BD FACSCalibur II flow cytometer (Becton Dickinson).

PLA. Colocalization between JCPyV and the ER was performed as previously described (14). The primary antibodies used for the proximity ligation assay (PLA) were rabbit anti-VP1 and (in the case of JCPyV) mouse anti-PDI. Cells were either left untreated or pretreated for 30 min with 0.1 mM Retro-2^{cycl} (Chembridge) and inoculated with 2×10^5 genome equivalents of JCPyV per cell. Virus was removed by aspiration, and any unbound virus was removed by washing with media. Fresh media containing 0.1 mM Retro-2^{cycl} or media alone were added. Cells were then incubated at 37°C for 8 h prior to fixation with 4% paraformaldehyde (PFA). Cells were permeabilized with PBS containing 0.5% Triton X-100 for 0.5 h, washed three times in PBS, and then blocked in 5% donkey serum for 1 h at 37°C and stained for VP1 (ab53977; Abcam) (1:1,000 dilution) and PDI (ab2792; Abcam) (1:100 dilution) by overnight incubation at 4°C. Colocalization between VP1 and PDI was then assessed using the proximity ligation assay (PLA), following the manufacturer's instructions (Bethyl Labs). Cells were washed, and the cell nuclei were counterstained using DAPI. Fluorescence micrographs were collected by confocal microscopy, and maximal *z*-projections were displayed. Fluorescent foci were quantified using the Blobfinder program (http://www.cb.uu.se/~amin/BlobFinder/). Threshold levels were set using a PLA-processed uninfected sample, and the same threshold levels were used across all samples. Blobfinder analysis was performed using a 7-by-7-pixel blob size, and at least 50 cells per sample were analyzed. The average number of blobs per cell was normalized to the wild-type PSV sample. Experiments were performed in triplicate, and error bars denote standard deviations.

VP2 exposure assays. SVGA cells were seeded in 12-well plates at a concentration of 5×10^4 cells per well and incubated overnight. Cells were either left untreated or pretreated with 0.1 mM Retro-2^{cycl} for 30 min. Samples were then inoculated with equivalent numbers of genome copies of JCPyV PSVs for 1 h at 37°C. Excess virus was washed off, and samples were incubated in either complete media or complete media with 0.1 mM Retro-2^{cycl}. The cells were then stained with a polyclonal antibody to VP2 (ab53983; Abcam) and counterstained with DAPI. Randomized fields of view were imaged, and the number of cells with exposed VP2 was determined. Results were normalized to the wild-type PSV sample. Experiments were performed in triplicate, and error bars denote standard deviations.

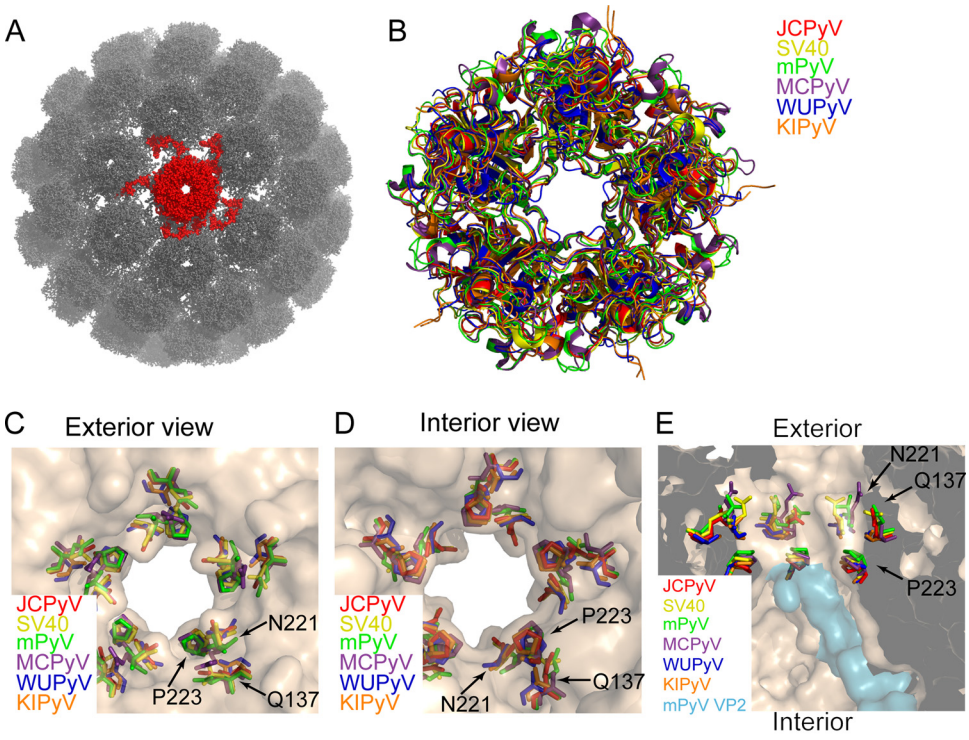
Protein structure accession numbers. Coordinates and structure factor amplitudes have been deposited with the RCSB Protein Data Bank (www.pdb.org) under accession codes 4WDY (N221Q), 4WDZ (N221W), and 4WE0 (P223M).

RESULTS

The 5-fold pore is conserved among all polyomaviruses. The polyomavirus 5-fold pore is formed by pentameric arrangement of VP1 monomers around the 5-fold axis of symmetry, as shown in the crystal structure of SV40 (Fig. 1A). Alignment of six divergent polyomavirus VP1 crystal structures demonstrates that the architecture of the 5-fold pore is conserved across polyomaviruses (Fig. 1B). In JCPyV, the N221 and P223 residues line the surface of the pore at its narrowest point, and Q137 is in close proximity to the 5-fold pore (Fig. 1C and D). This narrow constriction is in close proximity to VP2, as seen in the crystal structure of the VP1-VP2 complex in MPyV (Fig. 1E). The diameter of this pore is well conserved across structurally characterized polyomaviruses, with diameters ranging from 7.6 Å (WUPyV) to 8.6 Å (MPyV) at the respective bottlenecks describing the accessibility of the pore. The diameter of the pore was calculated by employing the Caver 3.0 algorithm (28). The sequence identity of amino acids lining the VP1 pore, including the highly conserved proline at position 223 in JCPyV VP1, which is present at corresponding positions in all analyzed polyomaviruses (Fig. 1F), is rather high among polyomaviruses.

Amino acid substitutions to residues lining the pore restrict growth. Using mutational analysis, we first sought to determine whether the 5-fold pore played a role in viral assembly or infectious cellular entry. A panel of mutant viruses were generated that contained amino acid substitutions to residues lining or near the 5-fold pore at its narrowest point. These residues, P223, N221, and Q137, were replaced with bulky amino acids that were predicted to reduce the diameter of the 5-fold pore. In addition, one mutation, N221W, was predicted to increase the hydrophobicity of this pore. Infectious clones of wild-type or mutant JCPyV were transfected into SVGA cells, and viral replication was assessed by multicycle growth assays. Transfection efficiencies were equivalent, and all samples expressed VP1, with an average of 4 VP1-positive cells per ×20 field (data not shown). This result indicated that these mutants were not defective in expressing the viral capsid protein. Whereas wild-type JCPyV spread throughout the tissue culture monolayer over a time period of 22 days, the 5-fold pore mutants were severely restricted in viral spread (Fig. 2). Only N221Q and N221Y, with 30% and 17% of wild-type levels, respectively, demonstrated limited viral spread. In order to verify that reversion mutations had not occurred in N221Q and N221Y, PCR and sequencing were performed on samples from the day 22 time point. VP1 sequencing demonstrated that these viruses had not reverted to a wild-type sequence. Thus, substitution of residues lining the 5-fold pore significantly reduced viral spread.

Substitutions to the 5-fold pore do not inhibit assembly of the VP1 pentamer. Since Q137, N221, and P223 are spatially close to the 5-fold axis, it is likely that any steric clashes resulting from amino acid substitutions would be amplified due to the close proximity of symmetry-related VP1 monomers also containing these substitutions. We therefore sought to determine whether amino acid substitutions to residues lining the 5-fold pore would interfere with assembly of the VP1 pentamer. Five-fold pore mutations were introduced into a bacterial expression plasmid encoding a truncated VP1 pentamer that has been used extensively to characterize the structures of polyomaviruses (6). Recombinant VP1 pentamers containing these 5-fold pore mutations were expressed in bacteria and purified by affinity and size exclusion



F

JCPyV	127	THDNGAGKPVQGTSFHFFSVG	147	212	YFGTLTGGENVPPVLHITNTAT	233
BKPyV	135	VHEHGGGKPIQGSNFHFFAVG	155	220	YFGTFTGGENVPPVLHVTNTAT	241
SV40	137	THENGAGKPIQGSNFHFFAVG	157	222	YFGTYTGGENVPPVLHITNTAT	243
MPyV	148	VNTKGISTPVEGSQYHVFAVG	168	239	YFGNYTGGTTPPVLQFTNTLT	260
MCPyV	148	VHDYGAGIPVSGVNYHMFAIG	168	242	YYGSIQTGSQTPTVLQFSNTLT	263
KIPyV	146	-NTNGVVHGLAGTQLYFWAVG	164	230	YFGRVVGGAATPPVVSYGNNST	251
WUPyV	136	-LTTGGVGGIAGPQLYFWAVG	155	232	YFGRMVGGAATPPVVSFSNNST	243
		* * *			*** * *** * *	

FIG 1 The VP1 5-fold pore is conserved across polyomaviruses. (A) The X-ray crystal structure of SV40 (PDB: 1SVA), represented by a ball and stick rendering. A single VP1 pentamer, containing the 5-fold pore, is colored red. (B) Alignment of the VP1 structures from six divergent polyomaviruses: JCPyV (PDB: 3NXG), SV40 (PDB: 3BWQ), MPyV (PDB: 1VPN), Merkel cell polyomavirus (MCPyV) (PDB: 4FMG), unassembled KI polyomavirus (KIPyV) (PDB: 3S7V), and unassembled Washington University polyomavirus (WUPyV) (PDB: 3S7X). (C) Exterior view of the 5-fold pore, illustrating that residues Q137, N221, and P223 line the pore at its narrowest point. (D) Interior view of the 5-fold pore. (E) Cross-sectional view of the interior axial cavity, showing the relationship of the indicated residues to VP2 in the MPyV VP1-VP2 complex structure (PDB: 1CN3). (F) Alignment of VP1 amino acid residues from seven different polyomaviruses. Red letters indicate residues that were substituted in this study; asterisks indicate residues that are conserved in all seven polyomaviruses.

chromatography to determine if these substitutions either inhibited protein expression or resulted in pentamers of altered size. Size exclusion chromatography demonstrated that the elution times of all pentamers were similar, indicating that substitution of the 5-fold pore does not result in larger structural defects in the VP1 pentamer (Fig. 3A). A secondary peak is seen in some chromatographs and is due to the formation of interpentameric disulfide bonds.

Pentamers containing pore mutations package the minor capsid proteins. Previous studies have shown that a segment of VP2 is close to the 5-fold pore in MPyV (7). To address the possibility that substitutions to residues lining the 5-fold pore inhibit association with VP2, wild-type or mutant VP1 pentamers were coexpressed with full-length VP2 lacking any affinity tags. Therefore, VP2 should copurify with VP1 only when these proteins can still interact. After coexpression and purification, immunoblotting was used to assay for the presence of VP2 (Fig. 3B). The results demonstrate that wild-type VP1 and each mutant VP1 are able to bind to VP2 (Fig. 3B).

Pore mutant pentamers are not reduced in thermal stability. We next sought to determine whether substitutions to residues lining the 5-fold pore reduced the stability of recombinant pentamers. To test this, differential scanning fluorometry was employed. This method involves the use of Sypro Orange, a fluorophore that fluorescences brightly when bound to nonpolar and, usually, buried residues. Assaying Sypro Orange fluorescence at various temperatures should therefore give an indication of the relative stability of each mutant pentamer. The results show an increase in Sypro Orange binding at between 40 and 50°C and no subsequent fluorescence at higher temperatures (Fig. 3C). Increased fluorescence was also seen for some mutants below 35°C. The cause of this increased fluorescence at low temperatures is not readily apparent but may be due to an increase in the levels of surface-accessible nonpolar residues in some mutants. The results were similar when VP1 and VP2 pentamers were used (data not shown). These results show that the wild-type VP1 pentamer and most 5-fold pore mutants are very similar with respect to thermal stability, with the excep-

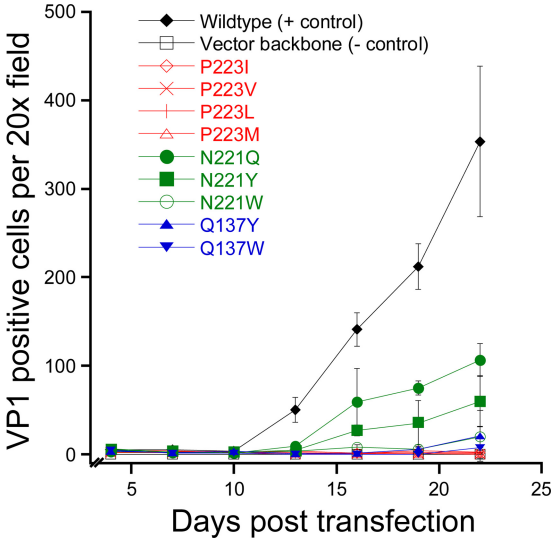


FIG 2 Amino acid substitutions to residues lining the 5-fold pore reduce infectivity. Infectious clones of wild-type JCPyV or pore mutants were transfected into SVGA cells, and viral spread was monitored over 22 days. Data represent the means of the results of three independent experiments, and the error bars indicate standard deviations.

tion of the Q137W mutant, which exhibits increased resistance to thermal denaturation.

VP1 mutations alter the pore morphology. Next, in order to assess structural effects of the pore mutations on the atomic level, we solved X-ray structures of VP1 pentamers containing N221Q, N221W, and P223M substitutions at resolutions of 1.9, 1.8, and 2.1 Å, respectively (Table 2). Well-defined electron density was observed in all cases in the simulated annealing F_{obs}-F_{calc} omit maps for the respective mutations. VP1 pentamer structures superpose with equivalent Cα atoms with root mean square deviation (RMSD) values of about 0.2 Å and are, as expected, very similar. A detailed analysis of the 5-fold pore shows very minor local structural changes due to the introduced mutations (Fig. 4). The diameter of the accessible pore of wild-type JCPyV VP1 was calculated to be about 8.2 Å at the narrowest point formed by residue N221 (28). Very minor local changes of the VP1 backbone were seen for N221Q and N221W (Fig. 4A and B). The glutamine alters the pore morphology and slightly narrows the accessible bottleneck diameter to about 7.1 Å, which is the tightest constriction of the pore, whereas the rather drastic N221W mutation does not reduce the pore diameter (the bottleneck diameter at W221 was calculated to be 8.6 Å) (Fig. 4D). Additional bulky electron density was observed within the pore of the N221W VP1 pentamer around the 5-fold axis. This density is unique for the N221W VP1 structure. It source could not be conclusively identified, but it is likely a hydrophobic solvent molecule that has no physiological relevance. The 5-fold cluster of hydrophobic tryptophans within the narrow pore favors hydrophobic interactions, including small hydrophobic solvent molecules also. The P223M substitution constrains the pore to a reduced diameter of about 5.2 Å at the bottleneck, where methionines face toward the 5-fold axis (Fig. 4C and D). Furthermore, additional minor structural rearrangements of the Cα backbone within the FG1-loop, such that N221 moves by about 2.3 Å toward the center of the pore (Fig. 4C), are seen.

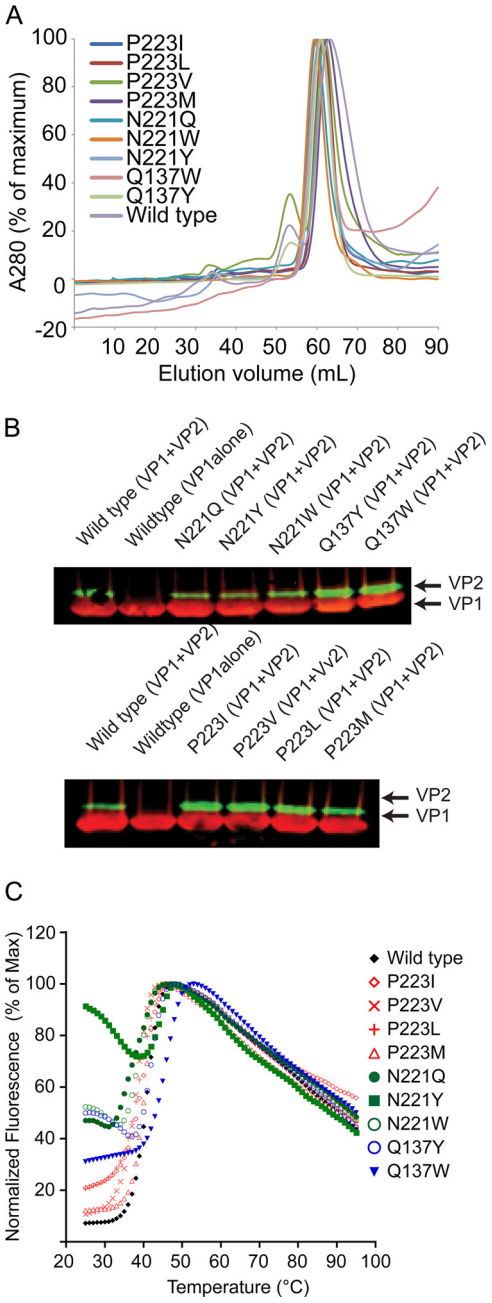


FIG 3 VP1 pentamers with altered 5-fold pores are not deficient in assembly. (A) A truncated VP1 containing a hexahistidine tag was expressed in bacteria and purified in order to assess the ability of pore mutants to form pentamers. (B) Truncated VP1 containing a hexahistidine tag and full-length VP2 lacking affinity tags were expressed in bacteria and purified by IMAC and SEC. Western blotting was used to determine whether the indicated pore mutants were able to copurify VP2. (C) Differential scanning fluorometry of VP1 pore mutants. The indicated VP1 was mixed with Sypro Orange, and fluorescence was monitored over the indicated range of temperatures. Max, maximum.

Because the P223L mutation was to be used in later experiments, a homology model of this mutant was generated. The diameter of the pore in the P223L homology model is much less restricted than in the P223M model and contains an accessible diameter of about 8.3 Å that is more similar to those seen with the wild type and the N221W and N221Q mutants (Fig. 4D).

TABLE 2 Data collection and structural refinement statistics[a]

Statistic	Result		
	VP1 N221Q	VP1 N221W	VP1 P223M
Space group	C2	C2	C2
Unit cell dimensions			
a, b, c (Å)	149.71, 95.79, 128.51	149.34, 95.98, 128.5	149.72, 95.63, 128.51
α, β, γ (°)	90.0, 110.4, 90.0	90.0, 110.5, 90.0	90.0, 110.40, 90.0
Resolution (Å)	50–1.90 (1.95–1.90)	50–1.8 (1.85–1.80)	50–2.10 (2.15–2.10)
No. of unique reflections	131,987 (8,272)	155,280 (10,329)	98,969 (7,258)
No. of observed reflections	538,018 (24,462)	633,087 (29,258)	378,369 (27,027)
R_{meas} (%)	11.9 (91.0)	9.2 (84.9)	17.3 (96.6)
I/σI	9.43 (1.7)	11.7 (1.5)	9.2 (1.7)
$CC_{1/2}$ (%)	99.7 (82.1)	99.8 (59.1)	99.2 (65.2)
Completeness (%)	98.1 (83.5)	98.8 (88.5)	99.8 (99.3)
Wilson B-factor (Å^2)	31.9	30.7	29.8
Refinement			
R_{work}, R_{free} (%)	19.1, 22.9	16.0, 18.6	18.0, 21.8
No. of atoms			
Protein	10,061	10,096	10,058
Water	724	873	530
Avg B-factor (Å^2)			
Protein	31.2	26.7	29.1
Water	35.8	33.7	30.8
RMS deviations			
Bond length (Å)	0.008	0.007	0.008
Bond angle (°)	1.315	1.269	1.308

[a] Values for the highest resolution bins are given in parentheses. meas, measured; RMS, root mean square.

JCPyV PSVs containing 5-fold pore substitutions are not morphologically altered or deficient in packaging VP2 or VP3. In order to study the effect of these pore mutations on infectious entry of JCPyV, we next generated wild-type or mutant JCPyV PSVs. PSVs have been extensively used to characterize events in infectious entry of polyomaviruses and additionally allow the study of mutants that could not otherwise be propagated. We selected two pore mutants, N221W and P223L, for further study in order to directly assess the effect of amino acid substitutions of residues lining the surface of the 5-fold pore. PSVs were generated for Q137W, but large variations in the number of infectious particles recovered between preparations were observed, suggesting that amino acid substitutions at this residue may alter virion stability. Additionally, since the Q138W pentamer appeared to have altered stability in our thermal denaturation experiments compared to wild-type or other mutant pentamers, Q137W was excluded from further analysis.

After purification, these PSVs were examined by negative-staining transmission electron microscopy to determine the effect of these substitutions on virion assembly. Wild-type and mutant PSVs appeared to be similar in morphology, demonstrating that PSVs containing these substitutions to the 5-fold pore are not grossly distorted or deficient in assembly (Fig. 5A). Immunoblotting also demonstrated that N221W and P223L PSVs packaged the minor capsid protein, further suggesting that substitution of residues lining the 5-fold pore does not interfere with minor capsid protein packaging (Fig. 5B).

Pore mutant PSVs are deficient in transduction. To determine whether modulation of residues surrounding the 5-fold pore affected viral entry, we performed single-cycle growth assays using wild-type or mutant PSVs that package a reporter plasmid expressing luciferase. Previous studies have demonstrated that expression of this reporter is directly proportional to the number of cells infected (32). Equivalent numbers of genome copies of each PSV were added to SVGA cells, and transduction efficiency was determined by determinations of relative luciferase expression levels. The results of these studies demonstrated that both N221W and P223L PSVs were deficient in transducing cells, suggesting a defect in viral entry (Fig. 5C).

Mutants are not defective in binding to host cells. We next sought to determine what step in the infectious life cycle is inhibited as a result of amino acid substitutions in residues lining the 5-fold pore. Binding assays were used to determine whether mutant PSVs were deficient in binding to cells. Equivalent numbers of genome copies of each PSV were added to SVGA cells, and binding was determined using a fluorescently labeled monoclonal antibody to VP1. This analysis demonstrated that wild-type, N22Q, and P223L PSVs all bound to cells equivalently and suggests that these amino acid substitutions do not reduce the ability of these mutant virions to bind to cells (Fig. 5D).

Pore mutant PSVs are transported to the ER but are deficient in exposing VP2. Transport of JCPyV to the ER is a critical step in infectious entry of JCPyV. Using proximity ligation assays, a method for reporting specific colocalization, we sought to deter-

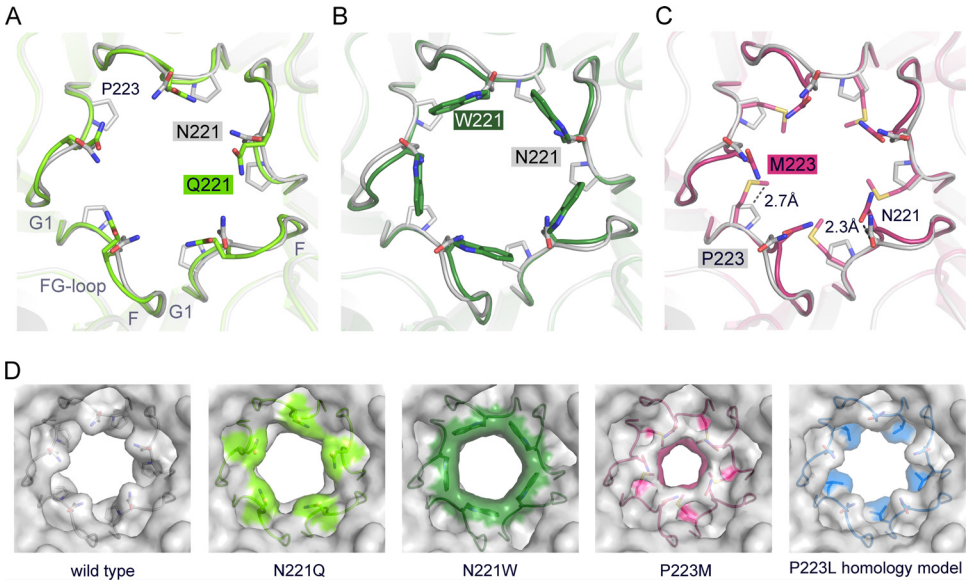


FIG 4 Crystal structures of VP1 pentamers with N221Q, N221W, and P223M pore mutations. (A to C) Structural superposition of pore mutants N221Q (A), N221W (B), and P221M (C) with wild-type Mad-1 VP1. VP1 pentamers are shown in cartoon representations with side chains of key amino acid residues highlighted in stick representations and colored according to atom type (oxygens are shown in red, nitrogens in blue, and carbons in the colors assigned for the respective mutants). Wild-type VP1 is shown in gray. The accessible diameter at the narrowest constriction of the 5-fold pore was determined (28). The diameter was calculated to be 8.2 Å for wild-type VP1 and to be 7.1 Å and 8.6 Å for N221Q and N222W, respectively. The P223M mutation results in a constrained accessible pore diameter of 5.2 Å. (D) Surface representations of the 5-fold pore. Views are equivalent in all cases, and mutated amino acids are colored on the surface. A VP1 homology model for P223L was generated, and the most favorable rotamer conformation of Leu is shown in a surface representation using sticks.

mine whether these pore mutants were defective in ER transport. Cells were inoculated with equivalent numbers of genome copies of each virus, and samples were fixed and stained at 8 h postinfection, a time point that has been shown to correspond to the arrival of JCPyV at the ER (14). PLA using antibodies to VP1 and PDI demonstrated that the pore mutants were transported to the ER as efficiently as wild-type viruses. Conversely, treatment of cells with Retro-2, a recently described inhibitor of JCPyV retrograde transport, prevented colocalization of JCPyV and PDI (Fig. 6).

Upon delivery to the ER, JCPyV interacts with ER resident chaperones and exposes the VP2 minor capsid protein. Since these pore mutations are in close proximity to the packaging site of VP2, we next determined whether these mutants are able to expose VP2 in the ER. Cells were inoculated with equivalent numbers of genome copies of each virus, and samples were fixed and immunostained for VP2 at 8 h hpi. This analysis revealed that VP2 foci could be observed for wild-type PSVs. In contrast, Retro-2 treatment of cells reduced the level of VP2 exposure. Surprisingly, both of the pore mutants failed to expose VP2, suggesting that these mutants were deficient in undergoing productive conformational changes in the ER needed to expose VP2 (Fig. 7).

DISCUSSION

Our report demonstrates that relatively subtle changes in residues lining the 5-fold pore of JCPyV result in dramatic decreases in infectivity. We show that these mutants are not deficient in assembly, binding to cells, or intercellular transport. Instead, these pore mutants are unable to expose the VP2 minor capsid protein in the ER during viral entry. This report suggests that contacts between the 5-fold pore and VP2 residues are a critical feature in the JCPyV capsid.

We initially suspected that this 5-fold pore was merely generated by the capsid symmetry and that residues occupying this space would result in significant structural clashes due to the close proximity of symmetry-related monomers. However, our results demonstrate that none of these substitutions results in assembly defects. In the case of N221Q, N221W, and P223M, only subtle changes were seen in the X-ray structure of VP1 pentamers, and similar minor changes were seen in the homology model of P223L. N221Q and N221W alter the pore diameter to only a very minor extent, but (in particular) the bulky N221W substitution mutation and, likely, also N221Y drastically change the molecular surface within the narrow pore and, likely, the interaction between VP1 and VP2. The structural impact of the P223M mutation on the FG1-loop conformation might explain why the prolines at equivalent VP1 positions are highly conserved. Although P223I, P223V, and P223L were not analyzed by X-ray crystallography, one could speculate about similar minor changes within the FG-loop. Leucine, valine, and isoleucine are smaller than methionine, so that the bottleneck radius might be less reduced; however, the nonpolar character at position 223 is preserved in all cases. Substitutions at amino acid position 137 (Q137Y and Q137W) were not reviewed by structural analysis, so the possibility cannot be ruled out that local structural rearrangement may occur. Indeed, differential scanning fluorometry analysis suggested that VP1 pentamers containing the Q137W substitution were altered in their thermal stability. The amino acid at position 137 does not directly line the 5-fold pore.

Our structural studies, demonstrating no significant structural differences between wild-type and mutant pentamers, coupled with the ability of both pentamers and PSVs to incorporate the minor capsid proteins, suggest that the mutant viruses do not display any drastic overall structural defects. This interpretation is additionally supported by the lack of differences in pentamer sta-

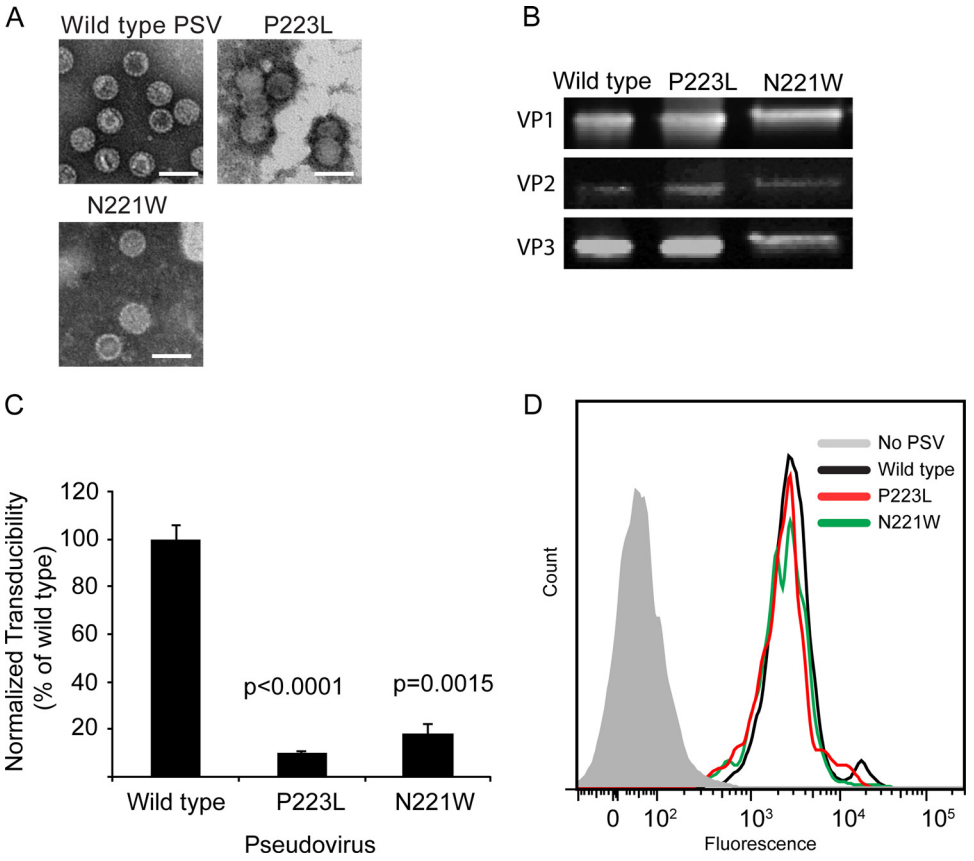


FIG 5 PSVs containing pore mutations are not deficient in assembly or binding to cells. (A) Wild-type or mutant PSVs were produced, and assembly was assessed by TEM. Scale bars denote 50 nm. (B) Western blotting was performed to demonstrate that all of the mutants were able to package VP2 and VP3. (C) Pore mutant PSVs are deficient in viral entry. SVGA cells were inoculated with equivalent numbers of genome copies of the indicated PSV, and luciferase expression was determined 72 hpi. (D) Pore mutant PSVs are not deficient in binding to cells. SVGA cells were inoculated with equivalent numbers of genome copies of the indicated PSV, and binding was determined by flow cytometry.

bility, as measured by differential scanning fluorometry. However, we cannot rule out the possibility that subtle structural changes in the capsid of these mutant viruses are responsible for the decrease seen in our multicycle growth assays.

PSVs containing these pore mutants were severely restricted in their ability to transduce susceptible cells, and the results suggested that these mutants were deficient at some step in viral entry. P223L and N221W were as efficient as wild-type PSVs in binding to cells and in transport to the ER, suggesting that that these pore mutants are still able to productively interact with cellular receptors and host transport factors. Instead, our VP2 exposure experiments demonstrate that these pore mutants do not expose VP2 in the ER. At present, the exact mechanism which causes these mutants to be deficient in exposure of VP2 is unclear.

The viral capsid of polyomaviruses is stabilized by interpentameric disulfide bonds, coordination of calcium ions, and hydrophobic interactions between the C terminus of VP1 and residues on the neighboring VP1 pentamer. It is likely that complete disassembly of the viral capsid requires disruption of the majority of these stabilizing factors. In SV40, virions remain largely intact following disulfide bond isomerization in the ER and retrotranslocation across the ER membrane occurs with virions in the form of large and intact particles (33). *In vitro* studies support this notion and demonstrate that treatment with the disulfide bond-reducing agent DTT is not sufficient for shedding of viral pentamers. Virions are seen to shed VP1 pentamers only after treatment with DTT and the calcium-sequestering molecule EGTA (34). Thin-section transmission electron microscopy studies have demonstrated that SV40 virions in the ER appear smaller in diameter and show visible VP2 in the form of extrusions (16). These studies suggested that VP2 is exposed in the absence of capsid disassembly. One possibility would be that the isomerization of intrapentameric disulfide bonds loosens the contacts between neighboring pentamers enough to allow exposure of the N terminus of VP2 though this opening. Alternatively, the 5-fold pore is situated near VP2, and previous studies have postulated that the 5-fold pore may provide a site of egress for the N terminus of VP2 (7). However, it should be noted that *in vitro* treatment of purified SV40 with ER luminal extracts results in distorted particles from which viral genomic DNA can be extracted, which suggests that capsid disassembly has occurred (16). It is unclear whether an analogous situation occurs during normal infectious entry of polyomaviruses.

At least four distinct possibilities exist for this observed defect in VP2 exposure. First, it is possible that these amino acids alter the interaction of VP1 with VP2, thereby interfering with the ability of the virion to expose VP2 during infectious cellular entry. Second, it is possible that these substitutions alter a recognition

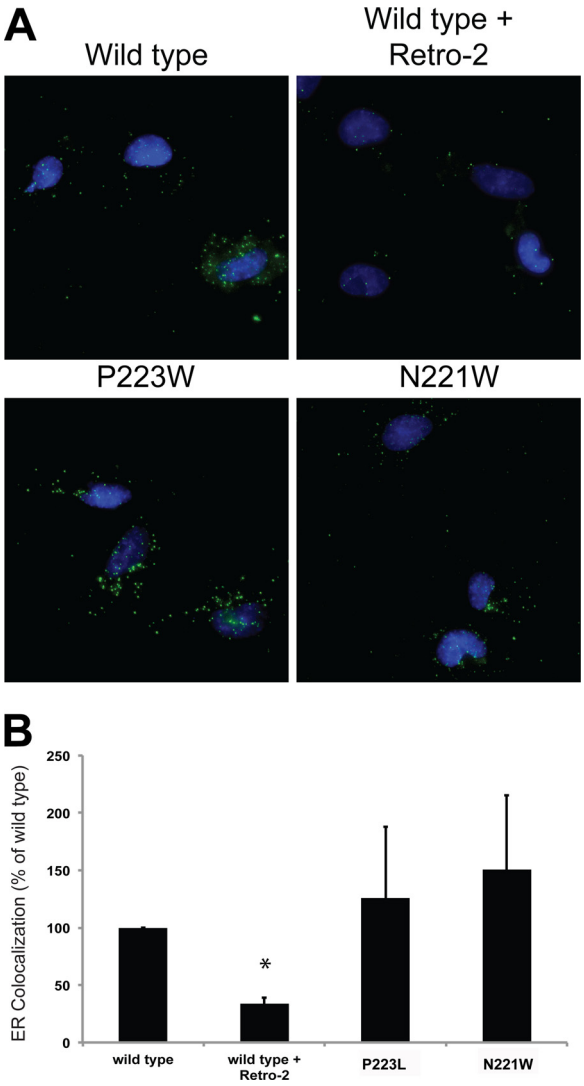


FIG 6 Pore mutant PSVs are transported to the ER as efficiently as wild-type PSVs. (A) SVGA cells were inoculated with equivalent numbers of genome copies of each PSV, and ER colocalization was assessed at 8 hpi by proximity ligation assays. (B) Quantification of the data presented in panel A. Results represent the means of the results of three independent experiments, and error bars denote standard deviations. *, $P < 0.05$.

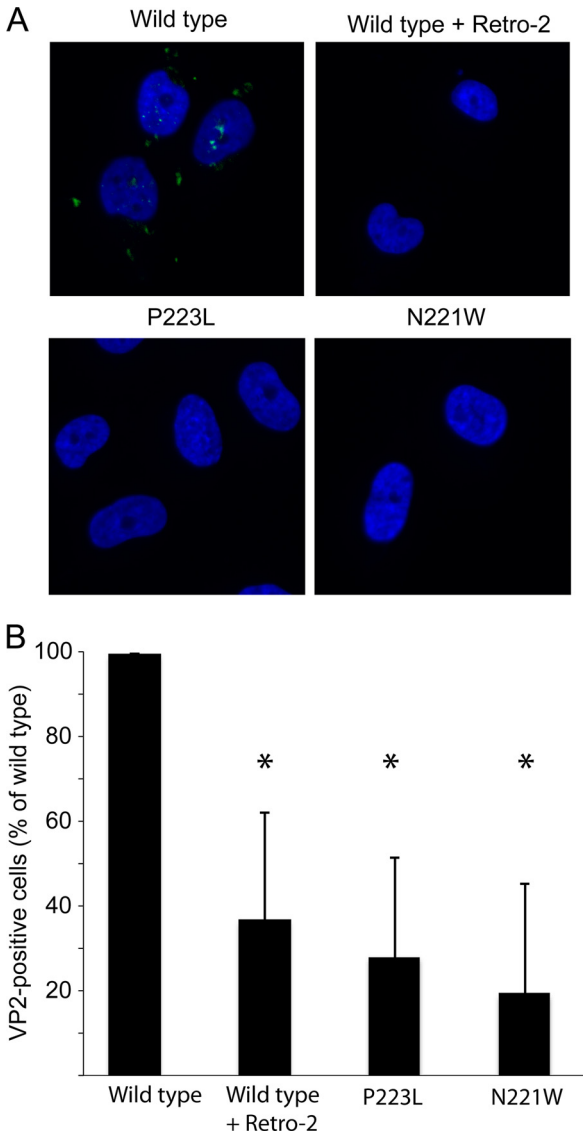


FIG 7 Pore mutants are deficient in exposing VP2 in the ER. (A) SVGA cells were inoculated with equivalent numbers of genome copies of each PSV, and VP2 exposure was assessed at 10 hpi by immunostaining for VP2. (B) Quantification of the data presented in panel A. Results represent the means of the results of three independent experiments, and error bars denote standard deviations. *, $P < 0.05$.

site for an as-yet-unidentified ER-resident chaperone or inhibit productive association with ER-residing chaperones, thereby inhibiting the conformational changes needed to expose the minor capsid proteins. Third, while these mutant viruses do not appear to be defective in ER transport, recent studies have shown that only a small percentage of virions that are transported to the ER actually expose VP2 or VP3. The identity of these parts of the ER is not well understood, but they appear to recruit chaperones such as BIP and DNAJ proteins to the foci where VP2 is exposed (35). It is therefore possible that these pore mutants are successfully transported to the ER but are deficient in transport to critical regions of the ER necessary for VP2 exposure and retrotranslocation. Finally, it is possible that the 5-fold pore of VP1 serves as a conduit for egress of the N terminus of VP2.

Definitively determining the site of VP2 egress from JCPyV is challenging. VP2 must be exposed for productive infection, and yet this exposure does not appear to be accompanied by complete disassembly of the viral capsid (16, 33). If disassembly of the viral capsid prior to VP2 exposure does not occur, then it is possible that VP2 is externalized either through the interpentameric gaps between monomers or through the pore present at the 5-fold axis of symmetry. While our results do not definitively determine the site of VP2 egress, they clearly demonstrate that modulation of this pore hampers exposure of VP2. Increased hydrophobicity within the pore and minor alterations of the pore diameter might alter, and perhaps strengthen, the contacts with VP2. Most of VP2 and VP3 is not resolved in the atomic resolution structures of MPyV, and it is therefore difficult to determine the extent to which the minor capsid proteins interacted with either the wild-type or the pore mutant in these studies. Further studies will be needed to

definitively determine the mechanism by which these minor capsid proteins are exposed.

In conclusion, we demonstrate that the 5-fold pore of JCPyV is an important structural feature and that modification of this pore results in a severe reduction in infectivity. These mutations do not hamper assembly of the virion but instead prevent the exposure of the minor capsid proteins. This work suggests that the 5-fold pore is important for the function of infectious cellular entry and that implementation of small molecules targeting this pore may prove to be an effective antiviral therapy.

ACKNOWLEDGMENTS

We thank members of the Atwood laboratory for critical discussions and reviews of the manuscript.

This work was supported by R01NS043097 (W.J.A.), P01NS065719 (W.J.A. and T.S.), and a Ruth L. Kirschstein National Research Service award (F32NS070687) (C.D.S.N.).

We thank the members of the DiMaio laboratory at Yale University for helpful discussions and for their assistance in performing proximity ligation assays. Quantitative PCR and imaging were performed in the Genomics Core at Brown, which is supported by P30GM103410 (W.J.A.).

REFERENCES

1. **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.
2. **Ferenczy MW, Marshall LJ, Nelson CD, Atwood WJ, Nath A, Khalili K, Major EO.** 2012. Molecular biology, epidemiology, and pathogenesis of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. Clin Microbiol Rev **25:**471–506. http://dx.doi.org/10.1128/CMR.05031-11.
3. **Zurhein G, Chou SM.** 1965. Particles resembling papova viruses in human cerebral demyelinating disease. Science **148:**1477–1479.
4. **Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed).** 2007. Fields virology, 5th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
5. **Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC.** 1991. Structure of simian virus 40 at 3.8-A resolution. Nature **354:**278–284. http://dx.doi.org/10.1038/354278a0.
6. **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.
7. **Chen XS, Stehle T, Harrison SC.** 1998. Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry. EMBO J **17:**3233–3240. http://dx.doi.org/10.1093/emboj/17.12.3233.
8. **Elphick GF, Querbes W, Jordan JA, Gee GV, Eash S, Manley K, Dugan A, Stanifer M, Bhatnagar A, Kroeze WK, Roth BL, Atwood WJ.** 2004. The human polyomavirus, JCV, uses serotonin receptors to infect cells. Science **306:**1380–1383. http://dx.doi.org/10.1126/science.1103492.
9. **Assetta B, Maginnis MS, Gracia Ahufinger I, Haley SA, Gee GV, Nelson CD, O'Hara BA, Allen Ramdial SA, Atwood WJ.** 2013. 5-HT2 receptors facilitate JC polyomavirus entry. J Virol **87:**13490–13498. http://dx.doi.org/10.1128/JVI.02252-13.
10. **Pho MT, Ashok A, Atwood WJ.** 2000. JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. J Virol **74:**2288–2292. http://dx.doi.org/10.1128/JVI.74.5.2288-2292.2000.
11. **Querbes W, Benmerah A, Tosoni D, Di Fiore PP, Atwood WJ.** 2004. A JC virus-induced signal is required for infection of glial cells by a clathrin- and eps15-dependent pathway. J Virol **78:**250–256. http://dx.doi.org/10.1128/JVI.78.1.250-256.2004.
12. **Querbes W, O'Hara BA, Williams G, Atwood WJ.** 2006. Invasion of host cells by JC virus identifies a novel role for caveolae in endosomal sorting of noncaveolar ligands. J Virol **80:**9402–9413. http://dx.doi.org/10.1128/JVI.01086-06.
13. **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.
14. **Nelson C, Carney D, Derdowski A, Lipovsky A, Gee G, O'Hara B, Williard P, DiMaio D, Sello J, Atwood W.** 2013. A retrograde trafficking inhibitor of ricin and Shiga-like toxins inhibits infection of cells by human and monkey polyomaviruses. mBio **4:**e00729–13. http://dx.doi.org/10.1128/mBio.00729-13.
15. **Gasparovic ML, Gee GV, Atwood WJ.** 2006. JC virus minor capsid proteins Vp2 and Vp3 are essential for virus propagation. J Virol **80:**10858–10861. http://dx.doi.org/10.1128/JVI.01298-06.
16. **Geiger R, Andritschke D, Friebe S, Herzog F, Luisoni S, Heger T, Helenius A.** 2011. BAP31 and BiP are essential for dislocation of SV40 from the endoplasmic reticulum to the cytosol. Nat Cell Biol **13:**1305–1314. http://dx.doi.org/10.1038/ncb2339.
17. **Magnuson B, Rainey EK, Benjamin T, Baryshev M, Mkrtchian S, Tsai B.** 2005. ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding. Mol Cell **20:**289–300. http://dx.doi.org/10.1016/j.molcel.2005.08.034.
18. **Rainey-Barger EK, Magnuson B, Tsai B.** 2007. A chaperone-activated nonenveloped virus perforates the physiologically relevant endoplasmic reticulum membrane. J Virol **81:**12996–13004. http://dx.doi.org/10.1128/JVI.01037-07.
19. **Vacante DA, Traub R, Major EO.** 1989. Extension of JC virus host range to monkey cells by insertion of a simian virus 40 enhancer into the JC virus regulatory region. Virology **170:**353–361. http://dx.doi.org/10.1016/0042-6822(89)90425-X.
20. **Aksamit AJ, Mourrain P, Sever JL, Major EO.** 1985. Progressive multifocal leukoencephalopathy: investigation of three cases using in situ hybridization with JC virus biotinylated DNA probe. Ann Neurol **18:**490–496. http://dx.doi.org/10.1002/ana.410180412.
21. **Kabsch W.** 2010. XDS. Acta Crystallogr D Biol Crystallogr **66**(Pt 2)**:**125–132. http://dx.doi.org/10.1107/S0907444909047337.
22. **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.
23. **Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, 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.
24. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr **66:**213–221. http://dx.doi.org/10.1107/S0907444909052925.
25. **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.
26. **Murshudov GN, Vagin AA, Dodson EJ.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr **53:**240–255. http://dx.doi.org/10.1107/S0907444996012255.
27. **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**(Pt 4)**:**439–450. http://dx.doi.org/10.1107/S0907444906005270.
28. **Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L, Sochor J, Damborsky J.** 2012. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput Biol **8:**e1002708. http://dx.doi.org/10.1371/journal.pcbi.1002708.
29. **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.
30. **Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A.** 2006. Curr Protoc Bioinformatics **15:**5.6:5.6.1–5.6.30. http://dx.doi.org/10.1002/0471250953.bi0506s15.
31. **Niesen FH, Berglund H, Vedadi M.** 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc **2:**2212–2221. http://dx.doi.org/10.1038/nprot.2007.321.
32. **Schowalter R, Reinhold W, Buck C.** 2012. Entry tropism of BK and

Merkel cell polyomaviruses in cell culture. PLoS One **7:**e42181. http://dx.doi.org/10.1371/journal.pone.0042181.

33. **Inoue T, Tsai B.** 2011. A large and intact viral particle penetrates the endoplasmic reticulum membrane to reach the cytosol. PLoS Pathog **7:**e1002037. http://dx.doi.org/10.1371/journal.ppat.1002037.
34. **Schelhaas M, Malmstrom J, Pelkmans L, Haugstetter J, Ellgaard L, Grunewald K, Helenius A.** 2007. Simian virus 40 depends on ER protein folding and quality control factors for entry into host cells. Cell **131:**516–529. http://dx.doi.org/10.1016/j.cell.2007.09.038.
35. **Goodwin EC, Lipovsky A, Inoue T, Magaldi TG, Edwards AP, Van Goor KE, Paton AW, Paton JC, Atwood WJ, Tsai B, Dimaio D.** 2011. BiP and multiple DNAJ molecular chaperones in the endoplasmic reticulum are required for efficient simian virus 40 infection. mBio **2:**e00101–11. http://dx.doi.org/10.1128/mBio.00101-11.

JVI Journals.ASM.org

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 α2,3- and α2,6-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.

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 = 7*d* 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 wukipolyomaviruses 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 α2,3- or α2,6-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 μl protein solution and 1 μl 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 CO_2 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 μl 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 μl PBS and pelleted after each wash at 200 × *g* (4 min; 4°C). Cells were incubated in 100 μl of labeled VP1 pentamer solution (50 μg/ml in PBS) on ice for 2 h with agitation every 15 to 20 min. HeLa S3 cells were washed twice in 500 μl PBS and fixed in 500 μl PBS containing 1% formaldehyde for 30 min. 293TT cells were washed twice in PBS and then suspended in 500 μl 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 FACSCanto (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-μl 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 α2,3-sialyllactose and α2,6-sialyllactose (Carbosynth) (each) and a 50 μM 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 D_2O. 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)
R_{meas} (%)	7.0 (56.8)	6.7 (55.9)
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		
R_{work}/R_{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.

$R_{meas} = \frac{\Sigma_{hkl} \sqrt{\frac{n}{n-1}} \Sigma_{j=1}^{n} | I_{hkl,j} - \langle I_{hkl,j} \rangle |}{\Sigma_{hkl} \Sigma_{j}^{n} I_{hkl,j}}$, where n is the number of observations of the reflection and $\langle I_{hkl,j} \rangle$ the intensity of symmetry (or Friedel)-related observations.

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

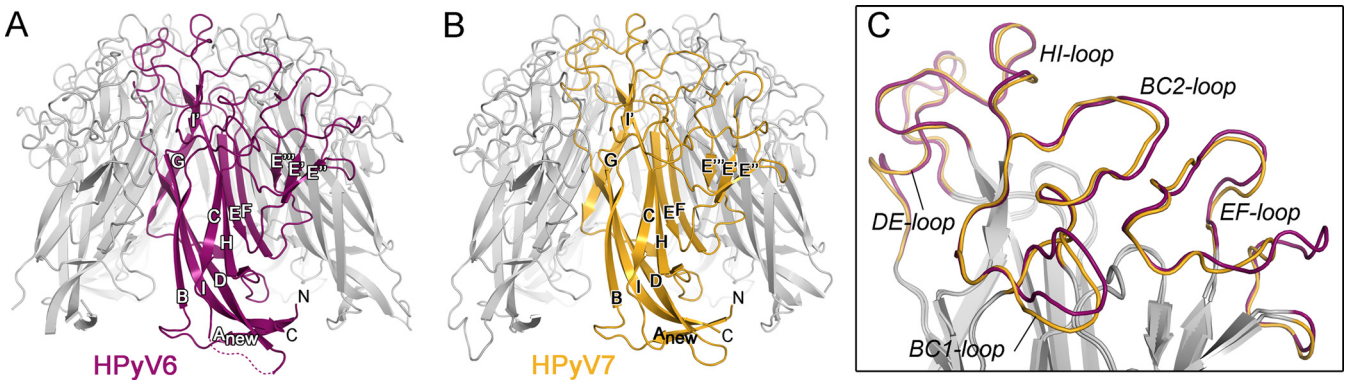


FIG 1 Architecture of HPyV6 and HPyV7 VP1 pentamers. (A and B) Overall folds of HPyV6 (A) and HPyV7 (B) VP1 pentamers shown in a ribbon representation, with VP1 monomers highlighted in magenta and gold, respectively. A dashed line represents the missing HPyV6 CD-loop residues, which are defined only 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

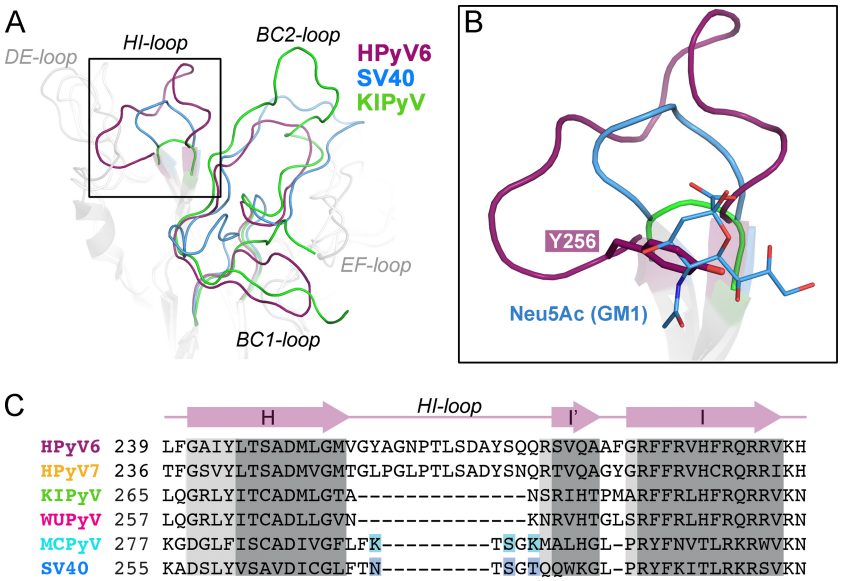


FIG 2 Architecture of VP1 surface loops. Superposition and comparison of VP1 surface loop structures are shown for wukipolyomaviruses HPyV6 and KIPyV (PDB accession no. 3S7V) and orthopolyomavirus SV40 (PDB accession no. 3BWR). VP1 monomers were superposed using the secondary-structure-matching (SSM) 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–α2,3-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 Cα atoms are shaded.

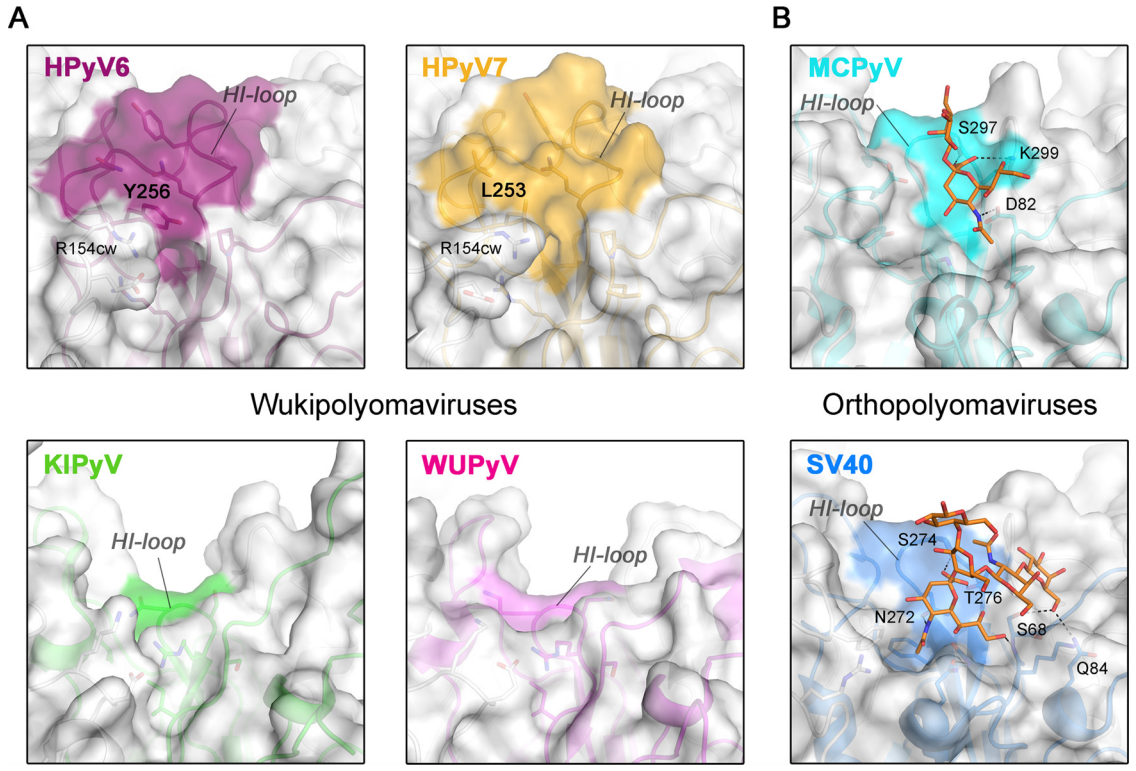


FIG 3 Surface structures of VP1 pentamers. Closeup views of VP1 pentamer top-surface regions that are involved in sialic acid engagement in the case of MCPyV and SV40 are shown. Equivalent surface sections are shown in surface and cartoon representations for VP1 pentamers from HPyV6, HPyV7, and KIPyV (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 α2,3-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. Unassembled 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 α2,3-, α2,6-, and α2,8-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 line with their known use of sialylated receptors for attachment. 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 α2,3- and α2,6-linked sialyllactose by STD NMR spectroscopy (52). STD NMR spectroscopy has been successfully used to define, for

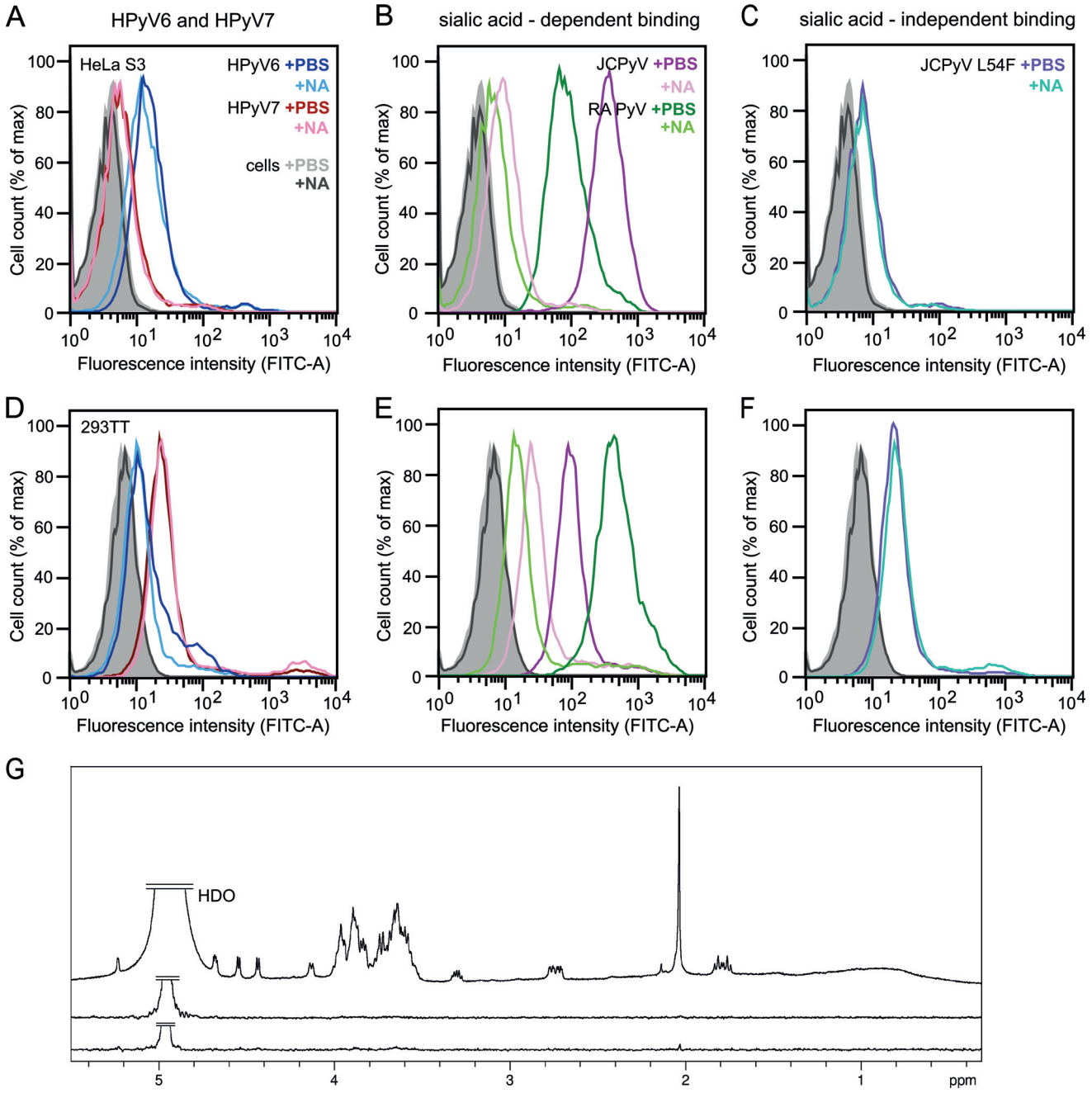


FIG 4 HPyV6 and HPyV7 VP1 do not engage sialic acids. (A to F) Cell binding analysis. (G) Saturation transfer difference (STD) NMR spectroscopy of HPyV6 and HPyV7 VP1 pentamers with α2,3- and α2,8-sialyllactose. (A to F) HeLa S3 (A to C) and 293TT (D to F) cells were subjected to mock treatment (PBS) 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: ^{1}H reference spectrum of 50 μM HPyV7 VP1 with 1 mM α2,3- and α2,6-sialyllactose each; STD NMR difference spectrum recorded with the same sample; STD NMR difference spectrum of 50 μM HPyV6 VP1 with 1 mM (each) α2,3- and α2,8-sialyllactose. No significant saturation transfer to either capsid protein was observed. HDO peaks were truncated for clarity.

example, the interactions of MCPyV and BKPyV 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 simple α2,3- or α2,6-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

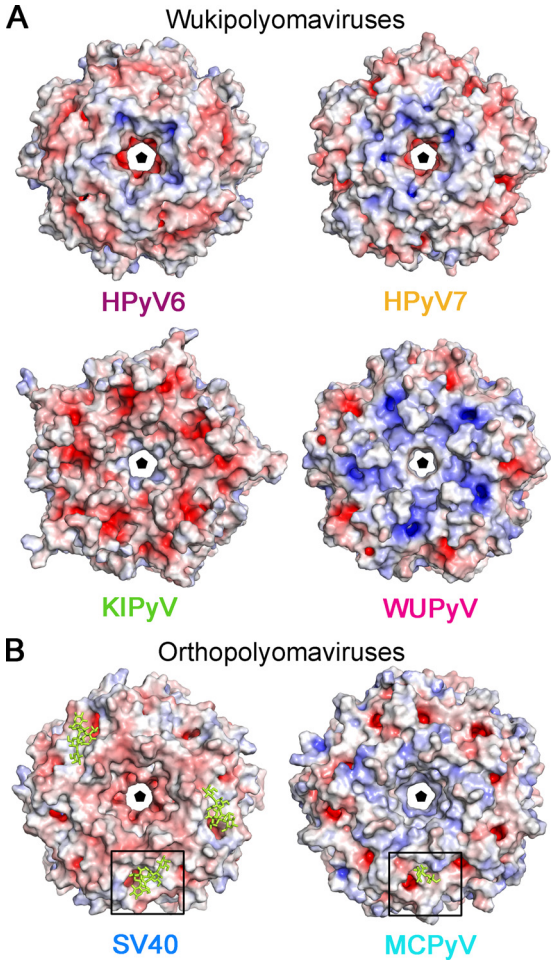


FIG 5 Electrostatic surface potentials of VP1 pentamer from the wuki- and orthopolyomavirus genera. Overall surface representations of HPyV6, HPyV7, and KIPyV (PDB accession no. 3S7V) and WUPyV (PDB accession no. 3S7X) (A) and SV40 VP1 (PDB accession no. 3BWR) and MCPyV (PDB accession no. 4FMI) (B) pentamers 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 α2,3-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 for 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 wukipolyomavirus 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 wukipolyomaviruses 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

1. **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.
2. **DeCaprio JA, Garcea RL.** 2013. A cornucopia of human polyomaviruses. Nat. Rev. Microbiol. **11:**264–276. http://dx.doi.org/10.1038/nrmicro2992.
3. **Feltkamp MC, Kazem S, van der Meijden E, Lauber C, Gorbalenya 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.
4. **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.
5. **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.
6. **Gaynor AM, Nissen MD, Whiley 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.
7. **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.
8. **Schowalter 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.
9. **van der Meijden E, Janssens RW, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, Feltkamp MC.** 2010. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. PLoS Pathog. **6:**e1001024. http://dx.doi.org/10.1371/journal.ppat.1001024.
10. **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.
11. **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.
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.
13. **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.
14. **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.
15. **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.
16. **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.
17. **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.
18. **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.
19. **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.
20. **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.
21. **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.
22. **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.
23. **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.
24. **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.
25. **Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC.** 1991. Structure of simian virus 40 at 3.8-A resolution. Nature **354:**278–284. http://dx.doi.org/10.1038/354278a0.
26. **Stehle T, Gamblin SJ, Yan Y, Harrison SC.** 1996. The structure of simian virus 40 refined at 3.1 A resolution. Structure **4:**165–182. http://dx.doi.org/10.1016/S0969-2126(96)00020-2.
27. **Schowalter 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.
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. **Ravantti 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.

Glycan Engagement by Viruses: Receptor Switches and Specificity

Luisa J. Ströh[1] and Thilo Stehle[1,2]

[1]Interfaculty Institute of Biochemistry, University of Tübingen, D-72076 Tübingen, Germany; email: thilo stehle@uni-tuebingen.de

[2]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Annu. Rev. Virol. 2014. 1:285–306

First published online as a Review in Advance on June 27, 2014

The *Annual Review of Virology* is online at virology.annualreviews.org

This article's doi: 10.1146/annurev-virology-031413-085417

Keywords

sialic acid, sulfated glycosaminoglycan, polyomavirus, coronavirus, reovirus, rotavirus

Abstract

A large number of viruses, including many human pathogens, bind cell-surface glycans during the initial steps of infection. Viral glycan receptors such as glycosaminoglycans and sialic acid–containing carbohydrates are often negatively charged, but neutral glycans such as histo–blood group antigens can also function as receptors. The engagement of glycans facilitates attachment and entry and, consequently, is often a key determinant of the host range, tissue tropism, pathogenicity, and transmissibility of viruses. Here, we review current knowledge about virus-glycan interactions using representative crystal structures of viral attachment proteins in complex with glycans. We illuminate the determinants of specificity utilized by different glycan-binding viruses and explore the potential of these interactions for switching receptor specificities within or even between glycan classes. A detailed understanding of these parameters is important for the prediction of binding sites where structural information is not available, and is invaluable for the development of antiviral therapeutics.

INTRODUCTION

Glycan: generic term for any free or covalently attached mono-, oligo-, or polysaccharide and carbohydrate

Sialic acids: α-acidic keto sugars with a 9-carbon backbone

Glycan microarray: high-throughput screening tool to analyze glycan-binding specificities of viruses and viral proteins

Glycosaminoglycans (GAGs): linear polysaccharides composed of amino sugars linked to uronic acid or galactose

Histo–blood group antigens (HBGAs): noncharged glycans present on red blood cells, epithelial cells, and mucosal secretions

Of the four fundamental building blocks of life—proteins, carbohydrates, lipids, and nucleic acids—carbohydrates (or glycans) arguably are the least well understood and the most underappreciated. Glycans append a wide variety of biological molecules and are found throughout the kingdoms of life. They often contribute to physical and structural integrity, extracellular matrix formation, signal transduction, protein folding, information exchange between cells, and pathogen uptake. However, we know little about many of these functions because glycans are notoriously difficult to work with. Unlike nucleic acids and proteins, they cannot be cloned and sequenced, and they cannot be studied with established molecular biology tools. They are usually difficult to synthesize, they are not always commercially available, and they often have unfamiliar names that are only informative to a small group of scientists. Moreover, information about their three-dimensional structures and conformational dynamics is still limited. All of these factors contribute to the perception that glycans are less exciting than other biologically important molecules such as nucleic acids and proteins.

The paucity of knowledge about glycan structure and function contrasts with the importance of these molecules in many biological processes, and this contrast is particularly obvious when investigating the life cycle of a virus. To gain entry into host cells and initiate an infectious cycle, a virus must first attach to one or several receptors, or attachment factors, that are present on the host cell surface. Engaging these receptors adheres the virus to its target cell and often determines host range and tissue tropism. Many viruses use cell-surface glycans to facilitate attachment and entry. In fact, carbohydrates such as sialic acids were among the first virus receptors known (1–4). However, despite an impressive number of virus-glycan complex structures that have been determined in recent years, we are still largely ignorant of the rules that allow a virus to engage a specific glycan sequence and enable it to switch to a different receptor with a limited change in amino acid sequence. In the majority of cases, little is known about the affinity between a virus and its cognate glycan receptor, as current technologies are often not suitable for determining critical parameters such as dissociation constants. Moreover, the densities and distribution patterns of glycans on cell surfaces are not well understood. However, recent advances in studies of virus-glycan interactions have made it possible to rapidly identify specific receptors by using glycan microarray screening (5–9), to map glycan epitopes that bind to a virus in solution by using saturation transfer difference NMR spectroscopy (10–14), to define the atomic level structure of the virus-glycan interaction by using X-ray crystallography, and to rationalize mutations to determine the precise effect of glycan binding in disease pathogenesis. These advances have contributed to a better understanding of how viruses bind to glycans, and this knowledge will clearly be essential for combating infection and designing improved therapeutic viral vectors.

In this review, we focus on viruses that exemplify different aspects of glycan receptor recognition particularly well and for which structural data are available at near-atomic resolution. Using these cases, we highlight general principles of virus-glycan receptor engagement, including the recognition of sialic acids, glycosaminoglycans (GAGs), and noncharged oligosaccharides such as histo–blood group antigens (HBGAs). Our objectives are to define themes and determinants of glycan recognition and, in particular, to elucidate strategies that viruses use to switch receptor specificities.

SIALIC ACID RECEPTORS

For many enveloped and nonenveloped viruses, cell-surface glycans featuring terminal sialic acids have been identified as functional receptors (**Figure 1**). Sialic acids are ubiquitously expressed in

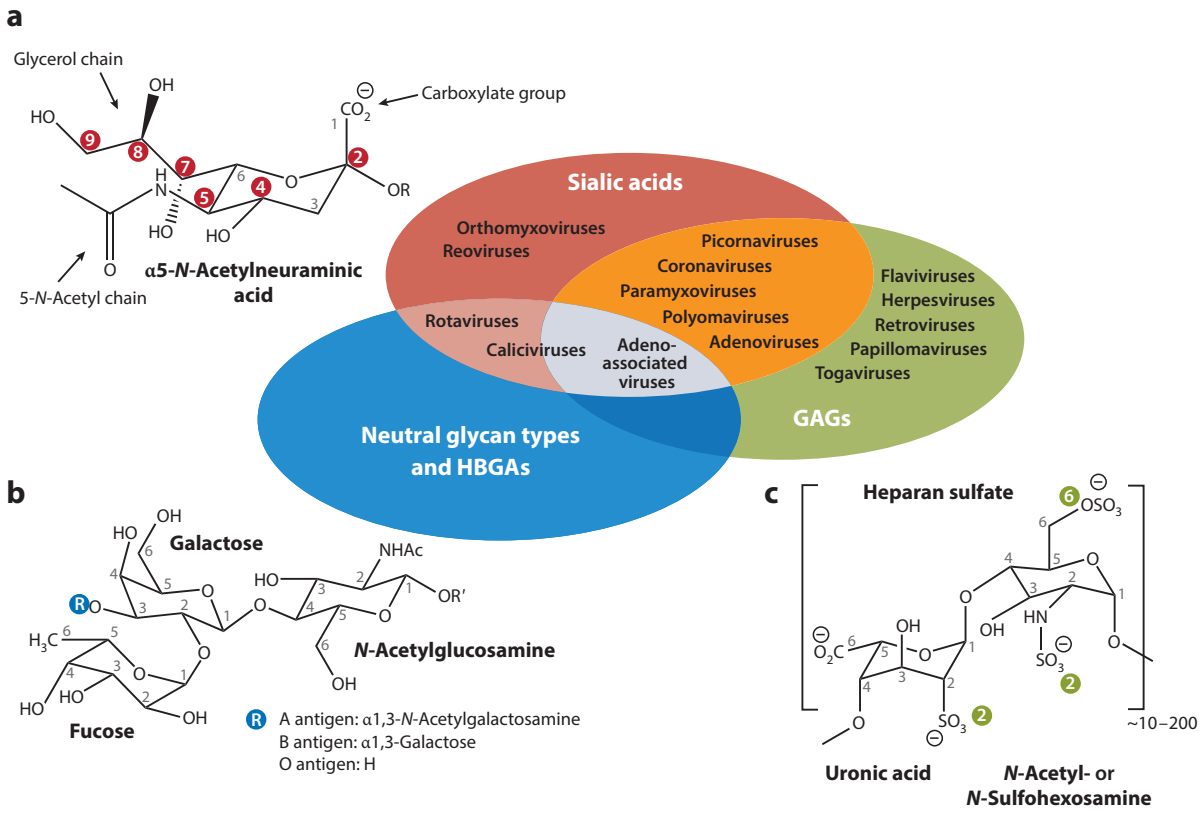


Figure 1

Glycans as viral receptors. Host- and cell-specific variants of (*a*) sialic acids, (*b*) neutral oligosaccharides such as histo–blood group antigens (HBGAs), and (*c*) glycosaminoglycans (GAGs) are used as receptors (or attachment factors) by many viruses. Examples of virus families with members that are known to interact with one or more glycan types are indicated in the respective overlapping ellipses of the Venn diagram. (*a*) The most common form of sialic acid in humans, *N*-acetylneuraminic acid, is shown in the α-conformation. This monosaccharide often caps *N*- and *O*-glycans and glycolipids (R), which are attached via different glycosidic linkages to position C2. Various modifications can occur at carbon positions C4, C5, and C7–9 (reviewed in 16). (*b*) Schematic representation of the type 2 core antigen, which is part of the ABO blood group determinants on *N*- and *O*-glycans. R′ denotes a heterosaccharide core attached to a glycoprotein. (*c*) Disaccharide units consisting of *N*-acetyl- or *N*-sulfohexosamine and either a uronic acid (glucuronic or iduronic acid) or galactose form the long, unbranched GAG chains. As an example, heparan sulfate is shown in the schematic chair representation. Heparan sulfate domains along the oligosaccharide chain usually vary in their modification contents by *N*-sulfation or *N*-acetylation at carbon position C2 of the hexosamine and sulfation at C2 of the uronic acid and at C6 of the hexosamine.

higher vertebrates and are attached at terminal ends of *N*- and *O*-glycans as well as glycolipids (15–17). The most common sialic acid in humans is α5-*N*-acetylneuraminic acid (Neu5Ac), but various modifications (acetylation, methylation, and sulfation) at C4, C5, C7, C8, and C9 give rise to more than 50 different sialic acid variants (**Figure 1*a***) (18), not all of which occur in all species.

Sialic acids are usually connected via α2,3 or α2,6 glycosidic linkages to galactose (Gal) or *N*-acetylgalactosamine (GalNAc) (19) such that protruding functional groups—carboxylate, glycerol, *N*-acetyl, and hydroxyl groups in the case of Neu5Ac (**Figure 1*a***)—are easily accessible for engagement by viral attachment proteins (reviewed in 20). Furthermore, sialic acids can be connected to one another with α2,8 or α2,9 linkages, and thus sialic acids can also occupy internal positions within glycoconjugates (21–23). For example, di- and trisialic acid motifs are common in

***N*-linked glycan:** glycan covalently attached to an asparagine of a polypeptide or protein usually within the consensus sequence -Asn-X-Ser/Thr-

the context of human b- and c-series gangliosides (24). Even longer chains of sialic acids, so-called polysialic acids, can modify proteins such as neural cell adhesion molecules and are aberrantly expressed in tumors (23, 25–28). In addition to various sialic acid modifications, developmentally regulated expression and distribution of different linkages can also strongly influence host range and tissue tropism of sialic acid–engaging viruses. Classic examples are influenza viruses, which infect different species depending on how the terminal Neu5Ac is linked to the underlying glycan structure (reviewed in 29–33).

Structures of viral attachment proteins bound to sialylated glycans reveal that the terminal Neu5Ac often serves as a hook during receptor engagement (reviewed in 20). In most cases, a highly specific network of interactions fastens the Neu5Ac to the protein. The negatively charged carboxylate at one end and the 5-*N*-acetyl group at the opposite end of Neu5Ac create a molecular "fingerprint" that appears to contribute toward selectivity over other glycans (20). However, accumulating structural data on virus-glycan interactions demonstrate that there are many strategies by which a virus can engage Neu5Ac. Here, we review some of these strategies and discuss their potential for specificity switching as well as their relevance for evolution and the emergence of new virus strains.

_O_-linked glycan: glycan linked most frequently to the hydroxyl group of amino acids serine and threonine

Glycolipid: a lipid with an attached glycan

α5-*N*-acetylneuraminic acid (Neu5Ac): one of the most common sialic acid variants (along with Neu5Gc)

Gangliosides: glycosphingolipids usually containing one or more sialic acids

a-, b-, and c-series gangliosides: glycosphingolipids with 1, 2, or 3 sialic acids, respectively, linked to the inner galactose

Glycosphingolipid: glycan linked to a lipid that contains the sphingolipid ceramide

Polysialic acids: linear homopolysaccharides composed of sialic acids; involved in development and plasticity of the brain and cancer metastasis

α5-*N*-glycolylneuraminic acid (Neu5Gc): one of the most common sialic acid variants (along with Neu5Ac)

A Conserved Binding Site for Sialic Acid in Some Polyomaviruses

Polyomaviruses are a rapidly expanding family of small, nonenveloped double-stranded DNA viruses (34, 35). The major capsid protein of these viruses, VP1, is organized into 72 pentamers on the virion surface and often engages sialic acid receptors of different types on host cells during attachment, cell entry, or both. Due to the high similarity among VP1 sequences and the therefore largely conserved VP1 pentamer core structure, polyomaviruses represent an excellent model system for correlating glycan receptor engagement and affinity with virus uptake, tropism, and pathogenesis. Crystal structures of VP1 in complex with sialylated glycans have been determined for the human JC, BK, and Merkel cell polyomaviruses (JCPyV, BKPyV, and MCPyV, respectively) as well as human polyomavirus 9 (HPyV9); the murine polyomavirus (MPyV); and two monkey polyomaviruses, simian virus 40 (SV40) and the B-lymphotropic polyomavirus (LPyV) (**Figure 2**) (14, 36–43).

Phylogenetically, SV40, BKPyV, and JCPyV are especially closely related (44). Their VP1 sequences are >74% identical at the amino acid level, and all three viruses recognize glycans terminating in Neu5Ac (45–48). SV40 engages the branched α2,3-linked GM1 ganglioside, BKPyV binds to α2,8-α2,3-di-Neu5Ac b-series gangliosides, and JCPyV recognizes the linear α2,6-LSTc pentasaccharide. The terminal Neu5Ac is engaged by all three viruses in a highly similar manner within a binding groove formed by surface-exposed loops that emanate from the β-sheet framework of each VP1 monomer (**Figure 2*a–c***). Neighboring VP1 chains in a pentamer close the groove at its ends and contribute contacts. There are conserved hydrogen bonds to the Neu5Ac carboxylate and *N*-acetyl groups. Moreover, the glycerol chain is located in a shallow depression, where it is recognized by van der Waals contacts and a backbone hydrogen bond. The methyl group of the 5-*N*-acetyl chain projects into a hydrophobic cavity. This cavity is somewhat larger and more polar in SV40 VP1, which likely explains why this virus is able to engage sialic acid variants such as α5-*N*-glycolylneuraminic acid (Neu5Gc) (39, 49) that cannot be synthesized by humans (50).

Despite the conserved contacts for terminal Neu5Ac, the SV40, BKPyV, and JCPyV VP1 proteins are highly specific for their respective glycan receptors. This specificity is achieved in each case by an exceedingly small number of additional contacts outside the core Neu5Ac-binding pocket. These limited auxiliary interactions are therefore attractive targets for switching receptor

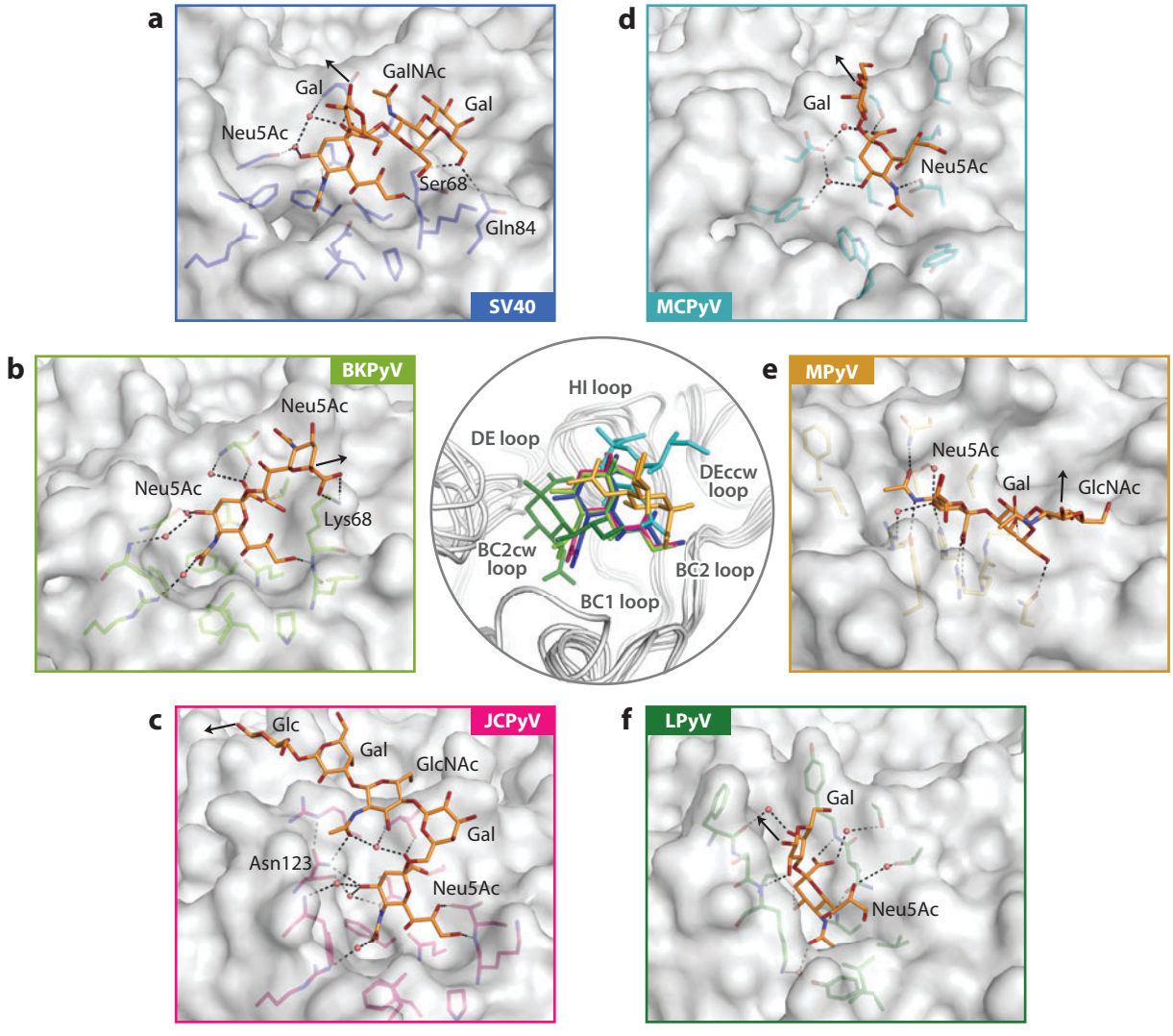


Figure 2

A highly plastic sialic acid–binding site on polyomavirus VP1. Panels *a–f* show views into the ligand-binding sites of six polyomaviruses for which structural data are currently available. The views are similar in each case. The VP1 proteins are shown in surface representation, with glycans and contacting protein represented with sticks. Hydrogen bonds and salt bridges between glycans and proteins are indicated with black dashed lines. (*a*) Simian virus 40 (SV40) VP1–GM1 glycan complex (PDB ID 3BWR; 39), (*b*) BK polyomavirus (BKPyV) VP1–GD3 glycan complex (PDB ID 4MJ0; 14), (*c*) JC polyomavirus (JCPyV) VP1–LSTc pentasaccharide complex (PDB ID 3NXD; 40), (*d*) Merkel cell polyomavirus (MCPyV) VP1–GD1a glycan complex (PDB ID 4FMJ; 41), (*e*) murine polyomavirus (MPyV) VP1 in complex with Neu5Ac–α2,3-Gal–β1,3-[α2,6-Neu5Ac]–GlcNAc–β1,3-Gal–β1,4-Glc (PDB ID 1VPS; 38), (*f*) B-lymphotropic polyomavirus (LPyV) VP1 in complex with α2,3-sialyllactose (PDB ID 4MBY; 42). A superposition of the six terminal Neu5Ac residues from panels *a–f* is shown in the center. The colors of the monosaccharides correspond to the colors assigned to the polyomaviruses in panels *a–f*. The loops surrounding the binding sites are labeled. The superposition highlights a conserved location of the Neu5Ac-binding site, with distinct interaction networks within or adjacent to the core binding sites. Abbreviations: GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; Neu5Ac, 5-*N*-acetylneuraminic acid.

specificities from one sialylated glycan to another. By comparing **Figure 2*a*** with **Figure 2*b***, it is readily apparent that residue 68 is the primary determinant of specificity for the glycans bound by BKPyV and SV40. Lys68 in BKPyV VP1 forms a salt bridge with the carboxylate group of the second Neu5Ac within the α2,8-α2,3-di-Neu5Ac sequence, whereas Ser68 at the equivalent position in SV40 VP1 contacts the Gal at the nonreducing end of the longer branch of GM1. Supporting a critical role of this residue in determining specificity, its mutation in BKPyV VP1 from Lys68 to Ser68 is sufficient to retarget BKPyV from the α2,8-α2,3-di-Neu5Ac sequence to GM1 in vitro and in cell culture (14). Remarkably, whereas SV40 and BKPyV enter cells via caveolae-mediated endocytosis, JCPyV is internalized by clathrin-mediated endocytosis (51, 52). Mutations in the VP1 proteins of SV40, BKPyV, and JCPyV might therefore be useful not only for receptor-switching purposes but also for modulating entry pathways.

A number of studies have investigated similar relationships between VP1 mutations and switches in glycan specificities for BKPyV, SV40, and JCPyV (53–59). It has also become clear that some VP1 mutations in JCPyV are critical for JCPyV pathogenicity, probably by altering the receptor-binding properties of this virus (53–56, 58).

Nonconserved Sialic Acid Binding in Other Polyomaviruses

The three polyomaviruses discussed in the above section all engage Neu5Ac by using highly similar contacts and in a similar orientation. Therefore, it would seem reasonable to assume that other sialic acid–binding polyomaviruses also engage Neu5Ac receptors in this manner. A conserved strategy within a virus family for attachment to sialic acid makes sense from an evolutionary point of view. For example, such a strategy is used by influenza viruses, which are perhaps the best-studied viruses that engage sialic acid. All structures of the influenza virus hemagglutinin determined to date engage the sialic acid portion of the glycan receptor in essentially the same orientation and derive specificity by recognizing different glycosidic linkages.

Strikingly, however, MCPyV, MPyV, and LPyV, which complete the set of polyomavirus VP1-ligand structures currently available, each engages Neu5Ac using a different strategy and in a different orientation (**Figure 2*d–f***), although the three VP1 proteins are >50% identical in their amino acid sequences (44). It is particularly noteworthy that Neu5Ac is bound in the same general area of the protein, involving residues that are in some cases even conserved but nonetheless contact different portions of Neu5Ac in each complex (38, 41, 42). Whereas MPyV and MCPyV have a shallow, surface-exposed binding site that resembles those of SV40, JCPyV, and BKPyV, the Neu5Ac-binding site on LPyV VP1 is much more recessed and essentially swallows the Neu5Ac moiety (**Figure 2*f***). Thus, polyomaviruses have an unusually plastic binding site for Neu5Ac that can engage the very same ligand in at least four different orientations and with different levels of solvent accessibility. This degree of plasticity is to our knowledge unique to polyomaviruses.

Differential Recognition of Sialic Acid Variants by Coronaviruses

The structural diversity of sialic acids far exceeds that of other common monosaccharides due to the many possible chemical modifications. Hence, it is not surprising that some viruses do not bind indiscriminately to all surface-associated sialic acids but instead prefer specific variants (49, 60, 61). In some cases, glycan microarray screening has suggested, and X-ray crystallography studies have confirmed, that distinct Neu5Ac substitutions are essential for the observed interaction. Viral receptor-binding proteins appear to favor particular sialic aid subtypes that carry additional acetyl or hydroxyl groups (6, 49, 62, 63). For example, the Neu5Gc variant of Neu5Ac carries an extra

hydroxyl group at the 5-*N*-acetyl moiety. Neu5Gc is not synthesized in humans (50) but can be absorbed from dietary sources and is metabolically incorporated into certain cell types such as intestinal endothelium and epithelium (64). Interestingly, Neu5Gc also accumulates in cancer cells (65). An important role of Neu5Gc in viral attachment and in defining species tropism has been firmly established for several viruses (49, 62, 66), and these are not reviewed here. Instead, we examine the implications of the engagement of *O*-acetylated sialic acids, which have recently been investigated by structural analysis of coronaviruses (63, 67).

O-acetylation of sialic acid appears in host-, organ-, and cell-specific patterns (68). Some members of the Coronaviridae family, including toroviruses and betacoronaviruses, possess hemagglutinin esterases (HEs) that allow reversible attachment to 5-*N*,9-*O*-diacetylneuraminic acid (Neu5,9Ac_2) or 4-*O*,5-*N*,9-*O*-triacetylneuraminic acid (Neu4,5,9Ac_3) through the combined action of two locally separated protein domains with either receptor-binding or receptor-destroying function (sialate-*O*-acetylesterase) (69, 70). Whereas most HEs bind Neu5,9Ac_2, some murine coronaviruses can use 4-*O*-acetylated Neu5Ac (Neu4,5Ac_2) as a component of their receptors (69).

The crystal structure of bovine coronavirus (BCoV) HE was determined in complex with Neu4,5,9Ac_3 (67), and the murine coronavirus mouse hepatitis virus strain S (MHV-S) HE protein was crystallized bound to Neu4,5Ac_2 (63). The two proteins share an amino acid sequence identity of approximately 60% and largely conserved overall structures. Comparison of the two structures nicely illustrates the surprisingly modest structural changes underlying the difference in specificity and demonstrates a strategy of subtly adjusting existing binding pockets to engage different sialic acid functional groups (**Figure 3**). The two binding sites possess largely conserved amino acids and reveal only minor changes in architecture in the receptor-binding region. Nevertheless, comparatively minor differences in MHV-S HE result in a shift in receptor-binding specificity toward 4-*O*- instead of 9-*O*-acetylated sialic acid (63). Interestingly, the hydrophobic pocket that accommodates the 9-*O*-acetyl moiety in the BCoV complex (Leu161, Tyr184,

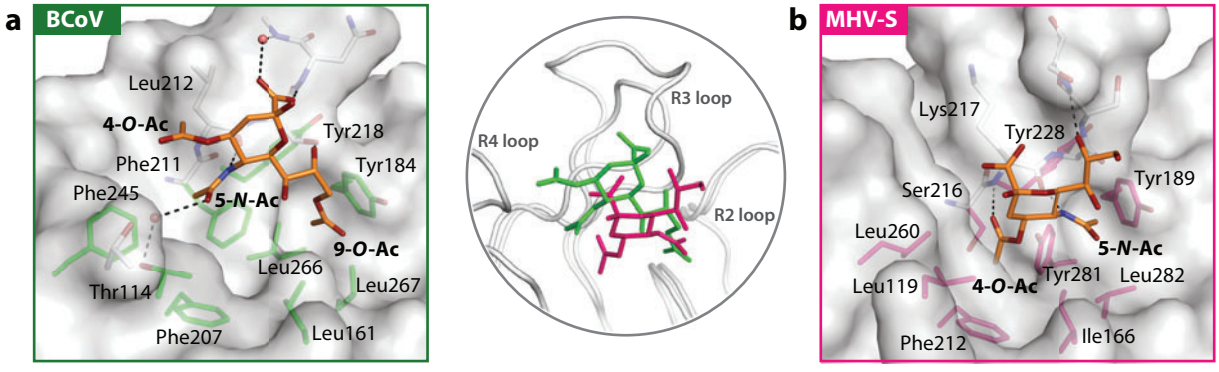


Figure 3

Engagement of differently *O*-acetylated 5-*N*-acetylneuraminic acid (Neu5Ac) by the hemagglutinin esterase (HE) proteins of bovine coronavirus (BCoV) and mouse hepatitis virus strain S (MHV-S). (*a*) A view into the ligand-binding site of BCoV HE bound to 4-*O*,5-*N*,9-*O*-triacetylneuraminic acid (Neu4,5,9Ac_3) (PDB ID 3CL5; 67). (*b*) MHV-S HE bound to 4-*O*,5-*N*-diacetylneuraminic acid (Neu4,5Ac_2) (PDB ID 4C7W; 63). The views are similar in each case. The HE proteins are shown in surface representation, with glycans and contacting protein represented with sticks. Hydrogen bonds and salt bridges between glycans and proteins are indicated with black dashed lines. A superposition of the two monosaccharides, shown in the middle, highlights minor structural changes within a topologically conserved receptor-binding site on HE. These changes allow for an alternative engagement of methyl groups from Neu4,5,9Ac_3 and Neu4,5Ac_2 in two conserved hydrophobic pockets. The colors of the monosaccharides correspond to the colors assigned to the viruses in panels *a* and *b*.

Leu266, and Leu267) (**Figure 3*a***) is conserved in MHV-S HE (Ile166, Tyr189, Tyr281, and Leu282) but in this case accommodates the 5-*N*-acetyl group (**Figure 3*b***). Consequently, to engage the 4-*O*-acetyl instead of 5-*N*-acetyl moiety in BCoV, a second hydrophobic pocket has adapted in MHV-S. The R3 loop is extended to allow the backbone carbonyl group (Lys217) to recognize the 5-*N*-acetyl group via hydrogen bonding, whereas the equivalent BCoV backbone carbonyl group (Leu212) interacts with the 5-*N*-acetyl group in an altered loop conformation. Hence, existing pockets engage topologically different acetyl groups in the two viruses, leading to a rotation of the sialic acid in the binding site (**Figure 3*b***). Due to this rotation, the glycerol chain is solvent exposed in the MHV-S complex but recognized at the C8 hydroxyl group by a backbone amide from the extended R3 loop. Thus, minor structural changes allow an alternative docking of methyl groups from Neu5,9Ac_2 and Neu4,5Ac_2 into two conserved hydrophobic pockets, with additional polar interactions assisting in specific recognition.

Proteoglycan: core protein in the mammalian extracellular matrix with one or more covalently attached glycosaminoglycan chains

Not all coronaviruses have HEs, and in fact coronaviruses such as BCoV and human coronavirus OC43 (HCo-OC43) also rely mostly on the N-terminal domain (NTD) of their spike (S) proteins for binding to Neu5,9Ac_2 (71, 72). Despite significant sequence similarity, the MHV-S NTD binds a protein receptor, and the virus exclusively uses HE for attachment to Neu4,5Ac_2 (73). However, structural details of the Neu5,9Ac_2 engagement mode by S NTDs in comparison with that of HEs remain to be elucidated (72, 74). It is not known whether severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), two emerging pathogens, engage glycan-based receptors, although HE proteoglycans appear to be involved in the attachment of SARS-CoV (75).

Physically Distinct Sialic Acid–Binding Sites in Reoviruses

The viruses discussed thus far use subtle modifications to modulate sialic acid binding specificities in a conserved binding site. However, some viruses use entirely different locations and binding modes on the same attachment proteins for glycan engagement. This strategy is exemplified by studies that provided a structural and functional basis for the interactions of two reovirus serotypes with sialylated glycan receptors (76, 77).

The three major serotypes of mammalian orthoreoviruses (reoviruses) are represented by prototype strains type 1 Lang (T1L), type 2 Jones (T2J), and type 3 Dearing (T3D). The well-studied T1L and T3D reoviruses differ markedly in viral spread and cell tropism, and these differences have been linked to sequence variations in the outer-capsid protein σ1 (78, 79). The fiber-like homotrimeric σ1 protein projects from the 12 vertices of the icosahedral reovirus virion and mediates the attachment of the virus to target cells (78, 80). The σ1 protein can be partitioned into three functionally and structurally distinct domains: the tail, body, and head. The N-terminal tail is predicted to form an α-helical coiled coil (81, 82), the body primarily consists of β-spiral repeats (76, 83), and the C-terminal head folds into a compact domain composed of eight antiparallel β-strands (83). The head binds with high affinity to the immunoglobulin superfamily member JAM-A (junctional adhesion molecule A) (84). As all known reovirus serotypes bind JAM-A (83), interactions with JAM-A are unlikely to dictate the serotype-specific differences in cell tropism. Instead, these differences in tropism are likely a consequence of virus binding to serotype-specific, sialylated glycan coreceptors.

Structure-function studies show that T3D σ1 engages terminal α2,3-, α2,6-, and α2,8-linked sialic acids in a binding site in the β-spiral section, near the midpoint of the molecule (76). Contacts are limited to the Neu5Ac moiety itself (**Figure 4*a***). The coreceptor for T1L reovirus was identified by glycan microarray screening as the GM2 glycan, a branched, sialylated tetrasaccharide (77). Structural analyses of T1L σ1 bound to GM2 demonstrated that the protein, much like T3D

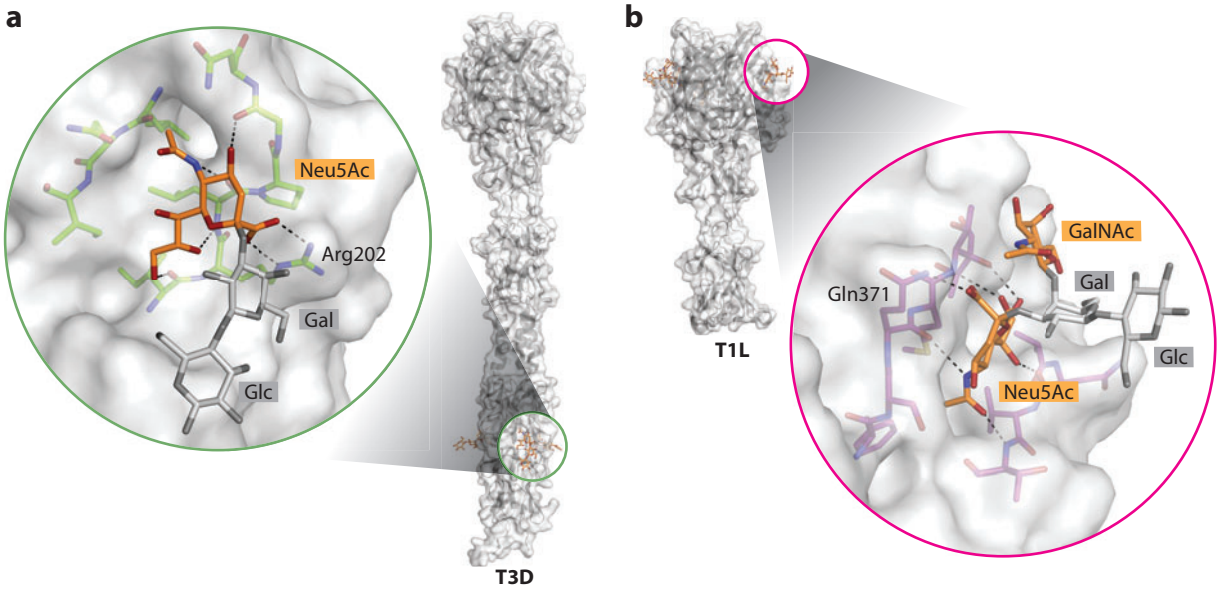


Figure 4

Adaptation of a new glycan-binding site on reovirus σ1. Crystal structures of reovirus (*a*) type 3 Dearing (T3D) σ1 in complex with the GM3 glycan (PDB ID 3S6X; 76) and (*b*) type 1 Lang (T1L) σ1 in complex with the GM2 glycan (PDB ID 4GU3; 77). In both panels, σ1 is shown as a semitransparent surface. In the close-up views of the two binding sites, key contact residues and bound glycans are represented with sticks. Black dashed lines display hydrogen bonds and salt bridges. Glycan moieties that are not involved in interactions with the protein are shown in gray. Abbreviations: GalNAc, *N*-acetylgalactosamine; Neu5Ac, 5-*N*-acetylneuraminic acid.

σ1, also primarily engages Neu5Ac. Surprisingly, however, the Neu5Ac-binding site is located in the head domain of σ1 (**Figure 4*b***), and is thus approximately 150 Å distant from the Neu5Ac-binding site of the homologous T3D σ1 protein. Both glycan-binding sites are distinct from the respective JAM-A-binding sites and would allow sequential receptor engagement in a multistep process. In addition to the overall location of the binding site, the strategy to engage Neu5Ac also differs drastically between the two serotypes. In the case of T3D σ1, a bidentate salt bridge with Arg202 anchors the Neu5Ac carboxylate, whereas this functional group is hydrogen bonded to a noncharged residue (Gln371) in T1L σ1. The other functional groups of Neu5Ac are also contacted in a different manner, although both proteins rely heavily on backbone interactions to ligate the glycan.

Reovirus σ1 thus represents a rare example of a viral attachment protein for which different strains have evolved physically distinct binding sites for essentially the same ligand, the Neu5Ac compound. We do not know to which structures, such as a carrier protein, the glycans are attached in each case, and the different locations of the two binding sites on σ1 might thus be dictated by the size and shape of that carrier. It is also possible that mechanisms of attachment are not conserved between the reovirus serotypes, and this could contribute to the observed differences in viral tropism and spread.

HISTO–BLOOD GROUP ANTIGEN RECEPTORS

In addition to sialic acids, some viruses also bind to neutral glycans such as HBGAs (**Figure 1**), which are present on red blood cells as well as epithelial cells and mucosal secretions (85, 86).

Well-studied examples for HBGA recognition are human noroviruses, which are members of the Caliciviridae family and cause outbreaks of acute gastroenteritis worldwide (reviewed in 87, 88). A human rotavirus also binds HBGAs, and this finding is of particular interest as most known rotaviruses engage sialylated glycan receptors. Therefore, it appears that rotaviruses can switch their specificity not only among sequences within a glycan family but also from one class of glycan to another.

Switching Glycan Specificity from Sialic Acids to Histo–Blood Group Antigens in a Rotavirus

Binding of rotaviruses to cells is mediated by outer-capsid protein VP4, a spike-like trimer that projects from the virion surface and serves as a major target for neutralizing antibodies (89). VP4 dictates the rotavirus classification system into P[*x*] genotypes (90). Rotavirus infectivity is enhanced by proteolytic cleavage of the VP4 trimer into N-terminal (VP8*) and C-terminal (VP5*) subunits. The VP8* subunit serves as the viral hemagglutinin and mediates interactions with glycan receptors (91), whereas the VP5* subunit facilitates membrane penetration (92). Animal rotaviruses engage terminal sialic acid–containing receptors, and crystal structures in complex with sialic acid show VP8* assuming a galectin fold, with an interesting twist (62, 93–98). Galectins are carbohydrate-binding proteins that typically engage galactose-terminating ligands via a conserved binding site at the top of the molecule. However, this site is blocked in VP8*, and the virus instead engages its sialic acid receptor via a different interface that lies on the side of the spike-shaped VP8* protein (99).

Rotaviruses are grouped historically into strains that are sensitive to treatment of cells with neuraminidase (sialidase-sensitive strains) and those that are not (sialidase-insensitive strains). Animal rotaviruses are usually sialidase sensitive, as they engage terminal sialic acid–containing receptors (62, 93–98). The receptor specificities of human rotaviruses have been the subject of considerable debate. Sialidase-insensitive strains can nevertheless engage and productively use sialic acid–based receptors in which the sialic acid is attached to one branch of biantennary glycans such as the GM1 and GD1a gangliosides (100). Because commonly used neuraminidases cannot remove the sialic acid in these branched glycans, treatment with these enzymes has no effect on infectivity.

A study combining glycan microarray screening and crystallographic analysis demonstrated that VP8* from the human P[14] rotavirus strain HAL1166 specifically binds HBGAs at the same site that binds sialic acid in other rotavirus strains (**Figure 5**) (101). Interestingly, the VP8* protein of HAL1166 rotavirus exhibits subtle modifications in its binding site that lead to its inability to bind sialic acid and allow it to instead engage HBGAs (101). The change in specificity is attributable to the insertion of a single amino acid in the HAL1166 VP8* binding pocket (Ser187) that reorients the neighboring residue Tyr188 such that its side chain clashes with sialic acid. At the same time, the reoriented Tyr188 side chain can now participate in hydrophobic contacts with the alternate receptor HBGA. As the remaining residues in the binding site are largely conserved among rotaviruses that do and do not bind sialic acid, this analysis demonstrates the profound effect on glycan specificity of an exceedingly subtle change in the receptor-binding pocket. It is remarkable that rotaviruses can also switch between entirely different classes of glycans through such modest changes in VP8*.

Interestingly, the switch between glycan classes observed for rotaviruses resembles a characteristic feature of caliciviruses. The protruding (P) domain of the major dimeric capsid protein VP1 constitutes the bulk of the outer surface of the calicivirus capsid. Sequence alignments in

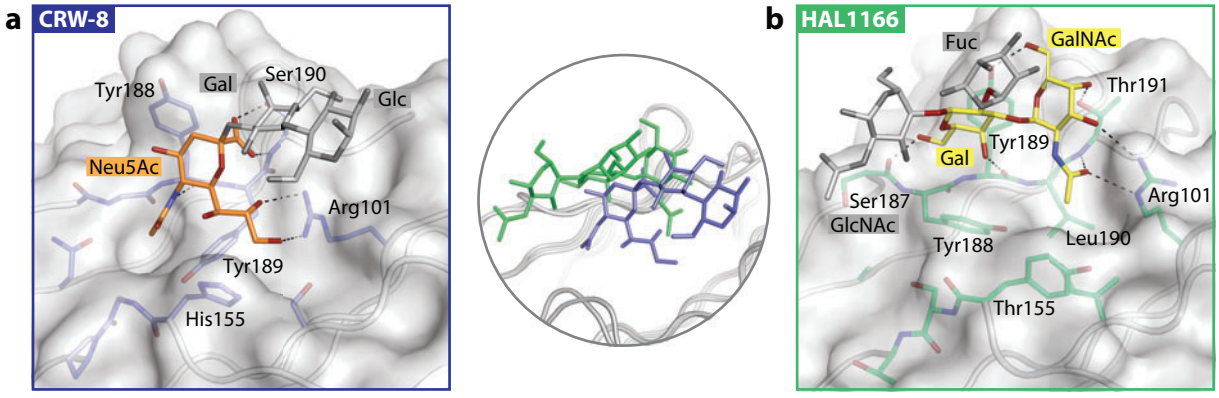


Figure 5

Engagement of different classes of glycans in a conserved location on the rotavirus spike protein VP8*. (*a*) A view into the ligand-binding site of porcine rotavirus CRW-8 VP8* bound to the GM3 glycan (PDB ID 3SIT; 66). (*b*) The same view for human rotavirus strain HAL1166 VP8* bound to a tetrasaccharide corresponding to the terminal structure of A-type histo–blood group antigen (HBGA) (PDB ID 4DS0; 101). The VP8* structures are shown in semitransparent surface representation, and key contact residues and bound glycans are depicted as sticks. Direct hydrogen bonds between VP8* and glycans are indicated by black dashed lines. Carbon atoms of A-type HBGA residues are shown in yellow and carbons of GM3 residues in orange to indicate the different classes of the two glycans. Glycan moieties that are not engaged by the protein are colored gray. A superposition of the two structures, performed by using secondary structure matching, is shown in the middle. The close-up view shows that both glycans bind in the same region of VP8*. Subtle changes in architecture and sequence result in a switch between different classes of glycan receptors. Abbreviations: GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; Neu5Ac, 5-*N*-acetylneuraminic acid.

combination with structural analysis revealed highly strain-specific HBGA-binding interfaces for engagement of different ABH and Lewis epitopes by human noroviruses (87, 88, 102–104). In contrast, the murine norovirus and the feline calicivirus engage sialic acids, possibly as parts of gangliosides (105, 106). A region topologically similar to the HBGA-binding site of the human norovirus strain VA387 has been identified on the murine norovirus P domain as a potential sialic acid–binding site, but atomic-level structural data are not yet available (107). Hence, functional and high-resolution structural studies are required to better understand glycan engagement by caliciviruses, and in particular to examine the molecular aspects of the specificity switch from HBGAs to Neu5Ac and how this switch is linked to possible adaptation to different hosts.

GLYCOSAMINOGLYCAN RECEPTORS

The third class of glycan receptors or attachment factors used by viruses is formed by the polyanionic GAGs. These linear glycans consist of β1,4-linked disaccharide blocks, which contain an *N*-acetylated or *N*-sulfated amino sugar (GlcNAc or GalNAc) linked to either a uronic acid (glucuronic or iduronic acid) or galactose (**Figure 1*c***) (108). Major GAG types are the nonsulfated hyaluronan (HA) and the sulfated compounds heparan sulfate (HS), chondroitin sulfate (CS), dermatan sulfate (DS), and keratan sulfate (KS). GAGs are components of proteoglycans and thus are present in the extracellular matrix as well as on the cell surface of most eukaryotic tissues and organs.

Heparan sulfate (HS): complex glycosaminoglycan that contains regions with modification by sulfation and epimerization to various extents

Given their prominent location at the surface of cells, it is not surprising that a large number of viruses use GAGs as attachment or entry factors (see also 109–111). The list of

viruses currently includes adenoviruses species B and C (112, 113), adeno-associated viruses (AAVs) (114), herpes simplex virus types 1 and 2 (115–117), human cytomegalovirus (118, 119), hepatitis viruses (120–123), dengue virus (124–126), papillomaviruses (127–131), MCPyV (132), human immunodeficiency virus type 1 (133), foamy virus (134), norovirus genotype II (135), respiratory syncytial virus (136), astroviruses (137), and members of the Togaviridae (138–140) and Picornaviridae (141–147) families.

Heparin: highly sulfated glycosaminoglycan that is produced and stored in mast cells and is used as anticoagulant

Despite this impressive list, our understanding of the rules underlying virus-GAG interactions at the molecular and atomic levels is limited. Few reliable high-resolution structures of viral receptor-binding proteins in complex with GAGs or GAG analogs exist, in contrast to the much larger databases of viral proteins bound to HBGAs or sialic acid. In addition, chemical heterogeneity and modifications of GAGs have challenged the development of analytical techniques to accurately define the GAG structures on target cells. Furthermore, observations that several viruses, such as foot-and-mouth disease virus (FMDV), undergo cell culture adaptation resulting in increased binding to HS without any selective pressure questioned the utilization of HS by clinical, non-culture-adapted strains (143).

The picornavirus FMDV was one of the first examples for which high-resolution studies of GAG-virus interactions became available (141, 142). The FMDV particle engages heparin in a shallow, surface-exposed groove, with contacts involving several heparin sulfate groups. Heparin, a highly sulfated form of HS, is widely used for in vitro binding experiments. Structures of human papillomavirus (HPV) capsid protein L1 from strains HPV-16 and HPV-18 in complex with oligomeric heparin have more recently been reported (129). However, our inspection of the available coordinates and electron densities (PDB IDs 3OAE and 3OFL) shows that the ligands are not well defined in the complexes and likely have low occupancies. The reported contacts between the HPV L1 proteins and heparin must therefore be treated with caution. A number of lower-resolution structures of complexes with GAGs or analogs have also been determined for different AAVs (148–150). These viruses are promising vectors for gene therapy, and current challenges lie in targeting AAVs to specific cells and enhancing cargo delivery while avoiding neutralization by the immune system (151). GAGs, sialic acids, and Gal can serve as attachment receptors for different AAV serotypes (114, 152–156). Crystal structures of human and simian AAVs in combination with mutagenesis have identified GAG-binding sites with clusters of positively charged residues on the particle surfaces (154, 157–159). However, cocrystallization and soaking experiments have not yielded high-resolution data that would allow for a detailed understanding of the interactions with GAGs. Cryo–electron microscopy (cryo-EM) has also shed some light on AAV-GAG interactions, but the exact binding sites and conformational changes triggered by GAG engagement remain controversial (148, 149), perhaps owing to the modest resolution of the structures and the heterogeneous nature of the ligands used for complex formation. The more homogeneous compound sucrose octasulfate (SOS) was used as a GAG analog in a cryo-EM reconstruction of AAV-DJ particles to 4.8 Å (150). The structural analysis revealed only modest local changes upon SOS complex formation relative to the crystal structure of the unliganded virus (157). It furthermore showed that the ligand is primarily contacted by salt bridges with arginines that were predicted earlier to be involved in GAG binding (157, 160). However, the short SOS disaccharide has low occupancy and thus probably has lower affinity for the virus capsid compared with longer GAG chains. It is also conceivable that longer GAG chains achieve higher avidity by bridging GAG-binding sites on the virion (161). Interestingly, a terminal Gal monosaccharide is a cell-surface receptor for AAV serotype 9 (156, 162), providing support for a switch between glycan receptor classes. AAV-9 is of particular relevance for gene delivery applications due to its efficient targeting and transduction of heart, liver, skeletal muscle, and alveolar epithelium. The Gal-binding site has been mapped by modeling and mutagenesis onto the native AAV structure

(162, 163), and moreover, initial studies report improved transduction efficiency by exploiting both glycan receptors for cell entry after engraftment of the Gal receptor footprint onto the GAG-binding serotype 2 virion (164). Interestingly, AAVs interact differentially with purified, homogeneous heparin oligosaccharides (165). However, high-resolution structural information is required to achieve a comprehensive understanding of diverse receptor recognition of AAVs and facilitate strategies for retargeting this virus.

CONCLUSIONS

For many viruses, interactions with cell-surface glycans are the first step in what can be a complex attachment and entry process. This initial interaction is often of lower affinity and, in some cases, is followed by second binding event to a coreceptor that tethers the virus more firmly to the cell surface. However, our understanding of how viruses engage glycans remains incomplete. The vast majority of available atomic-resolution structures are of viruses (or viral proteins) bound to either sialylated glycans or HBGAs, and these structures most often highlight a striking diversity in receptor engagement rather than conserved strategies. Viruses have, for example, evolved many different ways to engage sialic acid. Sometimes several binding modes can even be seen in closely related viruses within the same family, as demonstrated by the polyomaviruses. Although fascinating, this diversity in binding can clearly complicate efforts to predict interactions in even closely related viruses.

Although the individual interactions differ from virus to virus, most viruses do seem to bind glycan receptors in shallow, surface-exposed grooves that feature a limited number of contacts and tend to have low affinity for their ligand. A virus capsid protein that fully buries a portion of its glycan receptor upon binding, as seen in complexes of LPyV with sialylated glycans (**Figure 2*f***), is clearly the exception. Burying polar or charged compounds within a protein is obviously costly in terms of energy. However, low affinity and exposed binding surfaces may also offer important advantages to viruses, as these features can modulate binding more easily and effect changes in receptor binding with substitutions of few residues. Such changes could be required to allow a virus to target new cell types, quickly adapt to a new host, or modulate its antigenicity in response to immune surveillance. As demonstrated by the rotaviruses (**Figure 5**), switches between sialic acid and HBGAs, which have quite different chemical structures, can also be achieved with a surprisingly small set of substitutions in glycan-binding surfaces. We still lack detailed views of GAG binding by viruses, which is particularly frustrating given the involvement of GAGs in attachment by many viruses. For some viruses, GAGs are also predicted to induce conformational changes in viral capsids during attachment and entry, but the nature and function of such changes are unclear (166).

Coupled with tractable reverse genetics systems and high-resolution structural analyses, advances in glycan microarray screening now provide a powerful platform for the identification and characterization of glycan receptors and enable a thorough investigation of their interactions with viruses. We expect that such studies will lead to new approaches for antiviral therapy and help to establish a comprehensive toolbox for retargeting viruses from one type of glycan to another, which is of particular relevance for therapeutic viral vector development. An important remaining challenge lies in analyzing and quantifying the densities and developmentally regulated distribution patterns of glycans on host- and tissue-specific cell surfaces. Moreover, an improved correlation of affinity measurements with structural studies is required for a more rigorous evaluation of the determinants of specificity and for designing receptor switches. It is our hope that this review will contribute to meeting these challenges by helping to stimulate interdisciplinary research in the virus–glycan receptor field.

SUMMARY POINTS

1. Many enveloped and nonenveloped viruses, including many human pathogens, engage glycans such as sialic acids, sulfated glycosaminoglycans, and histo–blood group antigens as receptors (or attachment factors).
2. The engagement of specific glycan receptors often determines host range and tissue tropism and, ultimately, the outcome of the infection.
3. Glycan microarray screening has made it possible to rapidly identify specific virus glycan receptors.
4. X-ray structures of viral attachment protein–glycan complexes define virus-glycan interactions at the atomic level.
5. Conservations in sequence and three-dimensional structure as well as critical changes between more homologous viral attachment proteins from different serotypes or closely related viruses are often essential for glycan-binding specificities.
6. Switches of glycan receptor engagement can be due to limited changes within or approximate to a conserved core binding site or may result in the adaptation of a completely new binding site on the viral attachment protein.
7. Common interaction patterns and glycan binding modes on the molecular level have been described in more detail for sialic acids and histo–blood group antigens, but high-resolution complex structures of virus attachment proteins with glycosaminoglycans are rare.
8. A detailed understanding of the glycan-binding specificity, interaction affinity, and receptor-switching potential of viruses provides a new way forward in the development of antiviral strategies.

FUTURE ISSUES

1. Specific glycan receptors remain to be identified for many viruses.
2. Parameters that allow viruses to switch glycan receptors are still incompletely understood.
3. Modeling of glycan-virus interactions remains challenging due to the complexity and often rather shallow nature of the glycan-binding sites.
4. A more rigorous correlation of structural data with affinity measurements is required to fully evaluate glycan-binding properties and receptor switches.
5. Densities and developmentally regulated distributions of glycans on host- and tissue-specific cell surfaces need to be further analyzed and quantified.
6. Knowledge about virus-glycan interactions can provide a platform for the design of antiviral compounds.
7. Knowledge about virus-glycan interactions facilitates the retargeting of viruses to alternate receptors for gene delivery applications.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (P01 NS065719 and R01 AI076983) and the German Research Foundation (SFB-685) to T.S. We apologize to our numerous colleagues who have contributed to this area of research but whose work could not be cited or discussed here due to space limitations.

LITERATURE CITED

1. Hirst GK. 1941. The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. *Science* 94:22–23
2. Fazekas de St Groth S. 1948. Viropexis, the mechanism of influenza virus infection. *Nature* 162:294–95
3. Hirst GK. 1948. The nature of the virus receptors of red cells. I. Evidence on the chemical nature of the virus receptors of red cells and of the existence of a closely analogous substance in normal serum. *J. Exp. Med.* 87:301–14
4. Weis W, Brown JH, Cusack S, Paulson JC, Skehel JJ, Wiley DC. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* 333:426–31
5. Nilsson EC, Storm RJ, Bauer J, Johansson SMC, Lookene A, et al. 2011. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* 17:105–9
6. Song X, Yu H, Chen X, Lasanajak Y, Tappert MM, et al. 2011. A sialylated glycan microarray reveals novel interactions of modified sialic acids with proteins and viruses. *J. Biol. Chem.* 286:31610–22
7. Heimburg-Molinaro J, Tappert M, Song X, Lasanajak Y, Air G, et al. 2012. Probing virus-glycan interactions using glycan microarrays. *Methods Mol. Biol.* 808:251–67
8. Padler-Karavani V, Song X, Yu H, Hurtado-Ziola N, Huang S, et al. 2012. Cross-comparison of protein recognition of sialic acid diversity on two novel sialoglycan microarrays. *J. Biol. Chem.* 287:22593–608
9. Palma AS, Feizi T, Childs RA, Chai W, Liu L. 2014. The neoglycolipid (NGL)-based oligosaccharide microarray system poised to decipher the *meta*-glycome. *Curr. Opin. Chem. Biol.* 18:87–94
10. Mayer M, Meyer B. 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* 38:1784–88
11. Haselhorst T, Blanchard H, Frank M, Kraschnefski MJ, Kiefel MJ, et al. 2007. STD NMR spectroscopy and molecular modeling investigation of the binding of *N*-acetylneuraminic acid derivatives to rhesus rotavirus VP8*core. *Glycobiology* 17:68–81
12. Rademacher C, Krishna NR, Palcic M, Parra F, Peters T. 2008. NMR experiments reveal the molecular basis of receptor recognition by a calicivirus. *J. Am. Chem. Soc.* 130:3669–75
13. Fiege B, Rademacher C, Cartmell J, Kitov PI, Parra F, Peters T. 2012. Molecular details of the recognition of blood group antigens by a human norovirus as determined by STD NMR spectroscopy. *Angew. Chem. Int. Ed.* 51:928–32
14. **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 Pathog.* 9:e1003688**
15. Varki NM, Varki A. 2007. Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab. Investig.* 87:851–57
16. Varki A, Schauer R. 2009. Sialic acids. In *Essentials of Glycobiology*, ed. A Varki, RD Cummings, JD Esko, HH Freeze, P Stanley, et al., pp. 199–217. Cold Spring Harbor, NY: Cold Spring Harb. Lab. Press
17. Varki A, Gagneux P. 2012. Multifarious roles of sialic acids in immunity. *Ann. N. Y. Acad. Sci.* 1253:16–36
18. Angata T, Varki A. 2002. Chemical diversity in the sialic acids and related α-keto acids: an evolutionary perspective. *Chem. Rev.* 102:439–69
19. Nie H, Li Y, Sun XL. 2012. Recent advances in sialic acid–focused glycomics. *J. Proteomics* 75:3098–112

14. Identifies a single-amino-acid switch between different receptors.

20. Neu U, Bauer J, Stehle T. 2011. Viruses and sialic acids: rules of engagement. *Curr. Opin. Struct. Biol.* 21:610–18
21. Brisson JR, Baumann H, Imberty A, Perez S, Jennings HJ. 1992. Helical epitope of the group B meningococcal α(2–8)-linked sialic acid polysaccharide. *Biochemistry* 31:4996–5004
22. Nakata D, Troy FA II. 2005. Degree of polymerization (DP) of polysialic acid (polySia) on neural cell adhesion molecules (N-CAMS): development and application of a new strategy to accurately determine the DP of polySia chains on N-CAMS. *J. Biol. Chem.* 280:38305–16
23. Sato C, Kitajima K. 2013. Disialic, oligosialic and polysialic acids: distribution, functions and related disease. *J. Biochem.* 154:115–36
24. Yu RK, Tsai YT, Ariga T, Yanagisawa M. 2011. Structures, biosynthesis, and functions of gangliosides—an overview. *J. Oleo Sci.* 60:537–44
25. Finne J. 1982. Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. *J. Biol. Chem.* 257:11966–70
26. Roth J, Zuber C, Wagner P, Taatjes DJ, Weisgerber C, et al. 1988. Reexpression of poly(sialic acid) units of the neural cell adhesion molecule in Wilms tumor. *Proc. Natl. Acad. Sci. USA* 85:2999–3003
27. Hildebrandt H, Becker C, Gluer S, Rosner H, Gerardy-Schahn R, Rahmann H. 1998. Polysialic acid on the neural cell adhesion molecule correlates with expression of polysialyltransferases and promotes neuroblastoma cell growth. *Cancer Res.* 58:779–84
28. Janas T, Janas T. 2011. Membrane oligo- and polysialic acids. *Biochim. Biophys. Acta* 1808:2923–32
29. Viswanathan K, Chandrasekaran A, Srinivasan A, Raman R, Sasisekharan V, Sasisekharan R. 2010. Glycans as receptors for influenza pathogenesis. *Glycoconj. J.* 27:561–70
30. Ge SQ, Wang ZL. 2011. An overview of influenza A virus receptors. *Crit. Rev. Microbiol.* 37:157–65
31. Imai M, Kawaoka Y. 2012. The role of receptor binding specificity in interspecies transmission of influenza viruses. *Curr. Opin. Virol.* 2:160–67
32. Wilks S, de Graaf M, Smith DJ, Burke DF. 2012. A review of influenza haemagglutinin receptor binding as it relates to pandemic properties. *Vaccine* 30:4369–76
33. Matrosovich M, Herrler G, Klenk HD. 2013. Sialic acid receptors of viruses. *Top. Curr. Chem.* doi: 10.1007/128_2013_466
34. Dalianis T, Hirsch HH. 2013. Human polyomaviruses in disease and cancer. *Virology* 437:63–72
35. DeCaprio JA, Garcea RL. 2013. A cornucopia of human polyomaviruses. *Nat. Rev. Microbiol.* 11:264–76
36. Stehle T, Yan Y, Benjamin TL, Harrison SC. 1994. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* 369:160–63
37. Stehle T, Harrison SC. 1996. Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. *Structure* 4:183–94
38. Stehle T, Harrison SC. 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J.* 16:5139–48
39. Neu U, Woellner K, Gauglitz G, Stehle T. 2008. Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc. Natl. Acad. Sci. USA* 105:5219–24
40. Neu U, Maginnis MS, Palma AS, Ströh LJ, Nelson CD, et al. 2010. Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif. *Cell Host Microbe* 8:309–19
41. 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 Pathog.* 8:e1002738

42. Identifies an unusual sialic acid–binding site that essentially buries the glycan.

42. 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 Pathog.* 9:e1003714

43. Khan ZM, Liu Y, Neu U, Gilbert M, Ehlers B, et al. 2014. Crystallographic and glycan microarray analysis of human polyomavirus 9 VP1 identifies *N*-glycolyl neuraminic acid as a receptor candidate. *J. Virol.* 88:6100–11
44. van der Meijden E, Janssens RWA, Lauber C, Bouwes Bavinck JN, Gorbalenya AE, Feltkamp MC. 2010. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathog.* 6:e1001024

45. Liu CK, Wei G, Atwood WJ. 1998. Infection of glial cells by the human polyomavirus JC is mediated by an *N*-linked glycoprotein containing terminal α(2–6)-linked sialic acids. *J. Virol.* 72:4643–49
46. Komagome R, Sawa H, Suzuki T, Suzuki Y, Tanaka S, et al. 2002. Oligosaccharides as receptors for JC virus. *J. Virol.* 76:12992–3000
47. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, Rapoport TA. 2003. Gangliosides are receptors for murine polyoma virus and SV40. *EMBO J.* 22:4346–55
48. Low JA, Magnuson B, Tsai B, Imperiale MJ. 2006. Identification of gangliosides GD1b and GT1b as receptors for BK virus. *J. Virol.* 80:1361–66
49. Campanero-Rhodes MA, Smith A, Chai WG, Sonnino S, Mauri L, et al. 2007. *N*-Glycolyl GM1 ganglioside as a receptor for simian virus 40. *J. Virol.* 81:12846–58
50. Varki A. 2009. Multiple changes in sialic acid biology during human evolution. *Glycoconj. J.* 26:231–45
51. Pho MT, Ashok A, Atwood WJ. 2000. JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. *J. Virol.* 74:2288–92
52. Eash S, Querbes W, Atwood WJ. 2004. Infection of Vero cells by BK virus is dependent on caveolae. *J. Virol.* 78:11583–90
53. Zheng HY, Takasaka T, Noda K, Kanazawa A, Mori H, et al. 2005. New sequence polymorphisms in the outer loops of the JC polyomavirus major capsid protein (VP1) possibly associated with progressive multifocal leukoencephalopathy. *J. Gen. Virol.* 86:2035–45
54. 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
55. Delbue S, Branchetti E, Bertolacci S, Tavazzi E, Marchioni E, et al. 2009. JC virus VP1 loop-specific polymorphisms are associated with favorable prognosis for progressive multifocal leukoencephalopathy. *J. Neurovirol.* 15:51–56
56. Gorelik L, Reid C, Testa M, Brickelmaier M, Bossolasco S, 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.* 204:103–14
57. Magaldi TG, Buch MH, Murata H, Erickson KD, Neu U, et al. 2012. Mutations in the GM1 binding site of simian virus 40 VP1 alter receptor usage and cell tropism. *J. Virol.* 86:7028–42
58. Maginnis MS, Ströh LJ, Gee GV, O'Hara BA, Derdowski A, et al. 2013. Progressive multifocal leukoencephalopathy–associated mutations in the JC polyomavirus capsid disrupt lactoseries tetrasaccharide c binding. *mBio* 4:e00247-13
59. Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, Buck CB. 2013. BK polyomavirus genotypes represent distinct serotypes with distinct entry tropism. *J. Virol.* 87:10105–13
60. Rolsma MD, Kuhlenschmidt TB, Gelberg HB, Kuhlenschmidt MS. 1998. Structure and function of a ganglioside receptor for porcine rotavirus. *J. Virol.* 72:9079–91
61. Lofling J, Lyi SM, Parrish CR, Varki A. 2013. Canine and feline parvoviruses preferentially recognize the non-human cell surface sialic acid *N*-glycolylneuraminic acid. *Virology* 440:89–96
62. Yu X, Dang VT, Fleming FE, von Itzstein M, Coulson BS, Blanchard H. 2012. Structural basis of rotavirus strain preference toward *N*-acetyl- or *N*-glycolylneuraminic acid–containing receptors. *J. Virol.* 86:13456–66
63. Langereis MA, Zeng QH, Heesters B, Huizinga EG, de Groot RJ. 2012. The murine coronavirus hemagglutinin-esterase receptor-binding site: a major shift in ligand specificity through modest changes in architecture. *PLoS Pathogens* 8:e1002492
64. Bardor M, Nguyen DH, Diaz S, Varki A. 2005. Mechanism of uptake and incorporation of the non-human sialic acid *N*-glycolylneuraminic acid into human cells. *J. Biol. Chem.* 280:4228–37
65. Inoue S, Sato C, Kitajima K. 2010. Extensive enrichment of *N*-glycolylneuraminic acid in extracellular sialoglycoproteins abundantly synthesized and secreted by human cancer cells. *Glycobiology* 20:752–62
66. Yu X, Coulson BS, Fleming FE, Dyason JC, von Itzstein M, Blanchard H. 2011. Novel structural insights into rotavirus recognition of ganglioside glycan receptors. *J. Mol. Biol.* 413:929–39
67. Zeng QH, Langereis MA, van Vliet ALW, Huizinga EG, de Groot RJ. 2008. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc. Natl. Acad. Sci. USA* 105:9065–69

63. Highlights a glycan specificity switch from 9-*O*- to 4-*O*-acetylated sialic acid.

67. BCoV HE bound to the acetylated receptor reveals extended remodeling of the binding site.

68. Mandal C, Schwartz-Albiez R, Vlasak R. 2012. Functions and biosynthesis of *O*-acetylated sialic acids. *Top. Curr. Chem.* doi: 10.1007/128_2011_310
69. de Groot RJ. 2006. Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj. J.* 23:59–72
70. Langereis MA, Zeng Q, Gerwig GJ, Frey B, von Itzstein M, et al. 2009. Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases. *Proc. Natl. Acad. Sci. USA* 106:15897–902
71. Kunkel F, Herrler G. 1993. Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* 195:195–202
72. Peng G, Xu L, Lin YL, Chen L, Pasquarella JR, et al. 2012. Crystal structure of bovine coronavirus spike protein lectin domain. *J. Biol. Chem.* 287:41931–38
73. Langereis MA, van Vliet ALW, Boot W, de Groot RJ. 2010. Attachment of mouse hepatitis virus to *O*-acetylated sialic acid is mediated by hemagglutinin-esterase and not by the spike protein. *J. Virol.* 84:8970–74
74. Peng G, Sun D, Rajashankar KR, Qian Z, Holmes KV, Li F. 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proc. Natl. Acad. Sci. USA* 108:10696–701
75. Lang J, Yang N, Deng J, Liu K, Yang P, et al. 2011. Inhibition of SARS pseudovirus cell entry by lactoferrin binding to heparan sulfate proteoglycans. *PLoS ONE* 6:e23710
76. Reiter DM, Frierson JM, Halvorson EE, Kobayashi T, Dermody TS, Stehle T. 2011. Crystal structure of reovirus attachment protein σ1 in complex with sialylated oligosaccharides. *PLoS Pathog.* 7:e1002166
77. Reiss K, Stencel JE, Liu Y, Blaum BS, Reiter DM, et al. 2012. The GM2 glycan serves as a functional coreceptor for serotype 1 reovirus. *PLoS Pathog.* 8:e1003078
78. Weiner HL, Powers ML, Fields BN. 1980. Absolute linkage of virulence and central nervous system cell tropism of reoviruses to viral hemagglutinin. *J. Infect. Dis.* 141:609–16
79. Tyler KL, McPhee DA, Fields BN. 1986. Distinct pathways of viral spread in the host determined by reovirus S1 gene segment. *Science* 233:770–74
80. Lee PW, Hayes EC, Joklik WK. 1981. Protein σ1 is the reovirus cell attachment protein. *Virology* 108:156–63
81. Nibert ML, Dermody TS, Fields BN. 1990. Structure of the reovirus cell-attachment protein: a model for the domain organization of σ1. *J. Virol.* 64:2976–89
82. Duncan R, Horne D, Cashdollar LW, Joklik WK, Lee PW. 1990. Identification of conserved domains in the cell attachment proteins of the three serotypes of reovirus. *Virology* 174:399–409
83. Chappell JD, Prota AE, Dermody TS, Stehle T. 2002. Crystal structure of reovirus attachment protein σ1 reveals evolutionary relationship to adenovirus fiber. *EMBO J.* 21:1–11
84. Barton ES, Forrest JC, Connolly JL, Chappell JD, Liu Y, et al. 2001. Junction adhesion molecule is a receptor for reovirus. *Cell* 104:441–51
85. Ravn V, Dabelsteen E. 2000. Tissue distribution of histo–blood group antigens. *APMIS* 108:1–28
86. Marionneau S, Cailleau-Thomas A, Rocher J, Le Moullac-Vaidye B, Ruvoen N, et al. 2001. ABH and Lewis histo–blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie* 83:565–73
87. Tan M, Jiang X. 2010. Norovirus gastroenteritis, carbohydrate receptors, and animal models. *PLoS Pathog.* 6:e1000983
88. Ruvoën-Clouet N, Belliot G, Le Pendu J. 2013. Noroviruses and histo–blood groups: the impact of common host genetic polymorphisms on virus transmission and evolution. *Rev. Med. Virol.* 23:355–66
89. Yeager M, Dryden KA, Olson NH, Greenberg HB, Baker TS. 1990. Three-dimensional structure of rhesus rotavirus by cryoelectron microscopy and image reconstruction. *J. Cell Biol.* 110:2133–44
90. Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Banyai K, et al. 2011. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Arch. Virol.* 156:1397–413
91. Fiore L, Greenberg HB, Mackow ER. 1991. The VP8 fragment of VP4 is the rhesus rotavirus hemagglutinin. *Virology* 181:553–63
92. Dowling W, Denisova E, LaMonica R, Mackow ER. 2000. Selective membrane permeabilization by the rotavirus VP5* protein is abrogated by mutations in an internal hydrophobic domain. *J. Virol.* 74:6368–76

76. First example of a fiber-like structure engaging a sialic acid receptor.

77. Glycan microarray screening and crystallography provide an explanation for known serotype-dependent differences of reoviruses.

93. Dormitzer PR, Greenberg HB, Harrison SC. 2001. Proteolysis of monomeric recombinant rotavirus VP4 yields an oligomeric VP5* core. *J. Virol.* 75:7339–50
94. Monnier N, Higo-Moriguchi K, Sun ZY, Prasad BV, Taniguchi K, Dormitzer PR. 2006. High-resolution molecular and antigen structure of the VP8* core of a sialic acid–independent human rotavirus strain. *J. Virol.* 80:1513–23
95. Blanchard H, Yu X, Coulson BS, von Itzstein M. 2007. Insight into host cell carbohydrate recognition by human and porcine rotavirus from crystal structures of the virion spike associated carbohydrate-binding domain (VP8*). *J. Mol. Biol.* 367:1215–26
96. Kraschnefski MJ, Bugarcic A, Fleming FE, Yu X, von Itzstein M, et al. 2009. Effects on sialic acid recognition of amino acid mutations in the carbohydrate-binding cleft of the rotavirus spike protein. *Glycobiology* 19:194–200
97. Haselhorst T, Fiebig T, Dyason JC, Fleming FE, Blanchard H, et al. 2011. Recognition of the GM3 ganglioside glycan by rhesus rotavirus particles. *Angew. Chem. Int. Ed.* 50:1055–58
98. Martinez MA, Lopez S, Arias CF, Isa P. 2013. Gangliosides have a functional role during rotavirus cell entry. *J. Virol.* 87:1115–22
99. Dormitzer PR, Sun ZYJ, 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.* 21:885–97
100. Haselhorst T, Fleming FE, Dyason JC, Hartnell RD, Yu X, et al. 2009. Sialic acid dependence in rotavirus host cell invasion. *Nat. Chem. Biol.* 5:91–93
101. Hu L, Crawford SE, Czako R, Cortes-Penfield NW, Smith DF, et al. 2012. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo–blood group antigen. *Nature* 485:256–59
102. Huang PW, Farkas T, Zhong WM, Thornton S, Morrow AL, Xi J. 2005. Norovirus and histo–blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *J. Virol.* 79:6714–22
103. Tan M, Xia M, Chen Y, Bu W, Hegde RS, et al. 2009. Conservation of carbohydrate binding interfaces: evidence of human HBGA selection in norovirus evolution. *PLoS ONE* 4:e5058
104. 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 Pathog.* 7:e1002152
105. Stuart AD, Brown TDK. 2007. α2,6-Linked sialic acid acts as a receptor for feline calicivirus. *J. Gen. Virol.* 88:177–86
106. Taube S, Perry JW, Yetming K, Patel SP, Auble H, et al. 2009. Ganglioside-linked terminal sialic acid moieties on murine macrophages function as attachment receptors for murine noroviruses. *J. Virol.* 83:4092–101
107. Taube S, Perry JW, McGreevy E, Yetming K, Perkins C, et al. 2012. Murine noroviruses bind glycolipid and glycoprotein attachment receptors in a strain-dependent manner. *J. Virol.* 86:5584–93
108. Esko JD, Kimata K, Lindahl U. 2009. Proteoglycans and sulfated glycosaminoglycans. In *Essentials of Glycobiology*, ed. A Varki, RD Cummings, JD Esko, HH Freeze, P Stanley, et al., pp. 229–48. Cold Spring Harbor, NY: Cold Spring Harb. Lab. Press
109. Liu J, Thorp SC. 2002. Cell surface heparan sulfate and its roles in assisting viral infections. *Med. Res. Rev.* 22:1–25
110. Olofsson S, Bergstrom T. 2005. Glycoconjugate glycans as viral receptors. *Ann. Med.* 37:154–72
111. Zhu WY, Li JJ, Liang GD. 2011. How does cellular heparan sulfate function in viral pathogenicity? *Biomed. Environ. Sci.* 24:81–87
112. Dechecchi MC, Tamanini A, Bonizzato A, Cabrini G. 2000. Heparan sulfate glycosaminoglycans are involved in adenovirus type 5 and 2–host cell interactions. *Virology* 268:382–90
113. Tuve S, Wang H, Jacobs JD, Yumul RC, Smith DF, Lieber A. 2008. Role of cellular heparan sulfate proteoglycans in infection of human adenovirus serotype 3 and 35. *PLoS Pathog.* 4:e1000189
114. Summerford C, Samulski RJ. 1998. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J. Virol.* 72:1438–45
115. Nahmias AJ, Kibrick S. 1964. Inhibitory effect of heparin on herpes simplex virus. *J. Bacteriol.* 87:1060–66

101. Conserved location of HBGA- and sialic acid–binding sites on human and animal rotaviruses.

116. WuDunn D, Spear PG. 1989. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J. Virol.* 63:52–58
117. Shieh MT, WuDunn D, Montgomery RI, Esko JD, Spear PG. 1992. Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J. Cell Biol.* 116:1273–81
118. Neyts J, Snoeck R, Schols D, Balzarini J, Esko JD, et al. 1992. Sulfated polymers inhibit the interaction of human cytomegalovirus with cell surface heparan sulfate. *Virology* 189:48–58
119. Compton T, Nowlin DM, Cooper NR. 1993. Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology* 193:834–41
120. Barth H, Schafer C, Adah MI, Zhang F, Linhardt RJ, et al. 2003. Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J. Biol. Chem.* 278:41003–12
121. Schulze A, Gripon P, Urban S. 2007. Hepatitis B virus infection initiates with a large surface protein–dependent binding to heparan sulfate proteoglycans. *Hepatology* 46:1759–68
122. Kalia M, Chandra V, Rahman SA, Sehgal D, Jameel S. 2009. Heparan sulfate proteoglycans are required for cellular binding of the hepatitis E virus ORF2 capsid protein and for viral infection. *J. Virol.* 83:12714–24
123. Longarela OL, Schmidt TT, Schöneweis K, Romeo R, Wedemeyer H, et al. 2013. Proteoglycans act as cellular hepatitis delta virus attachment receptors. *PLoS ONE* 8:e58340
124. Chen Y, Maguire T, Hileman RE, Fromm JR, Esko JD, et al. 1997. Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nat. Med.* 3:866–71
125. Dalrymple N, Mackow ER. 2011. Productive dengue virus infection of human endothelial cells is directed by heparan sulfate–containing proteoglycan receptors. *J. Virol.* 85:9478–85
126. Roehrig JT, Butrapet S, Liss NM, Bennett SL, Luy BE, et al. 2013. Mutation of the dengue virus type 2 envelope protein heparan sulfate binding sites or the domain III lateral ridge blocks replication in Vero cells prior to membrane fusion. *Virology* 441:114–25
127. Joyce JG, Tung JS, Przysiecki CT, Cook JC, Lehman ED, et al. 1999. The L1 major capsid protein of human papillomavirus type 11 recombinant virus–like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. *J. Biol. Chem.* 274:5810–22
128. Giroglou T, Florin L, Schafer F, Streeck RE, Sapp M. 2001. Human papillomavirus infection requires cell surface heparan sulfate. *J. Virol.* 75:1565–70
129. Dasgupta J, Bienkowska-Haba M, Ortega ME, Patel HD, Bodevin S, et al. 2011. Structural basis of oligosaccharide receptor recognition by human papillomavirus. *J. Biol. Chem.* 286:2617–24
130. Cruz L, Meyers C. 2013. Differential dependence on host cell glycosaminoglycans for infection of epithelial cells by high-risk HPV types. *PLoS ONE* 8:e68379
131. Cerqueira C, Liu Y, Kuhling L, Chai W, Hafezi W, et al. 2013. Heparin increases the infectivity of human papillomavirus type 16 independent of cell surface proteoglycans and induces L1 epitope exposure. *Cell Microbiol.* 15:1818–36

132. A novel two-step glycan-binding entry mechanism within the polyomavirus family.

132. Schowalter RM, Pastrana DV, Buck CB. 2011. Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry. *PLoS Pathog.* 7:e1002161

133. Saphire ACS, Bobardt MD, Zhang Z, David G, Gallay PA. 2001. Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages. *J. Virol.* 75:9187–200
134. Nasimuzzaman M, Persons DA. 2012. Cell membrane–associated heparan sulfate is a receptor for prototype foamy virus in human, monkey, and rodent cells. *Mol. Ther.* 20:1158–66
135. Tamura M, Natori K, Kobayashi M, Miyamura T, Takeda N. 2004. Genogroup II noroviruses efficiently bind to heparan sulfate proteoglycan associated with the cellular membrane. *J. Virol.* 78:3817–26
136. Hallak LK, Spillmann D, Collins PL, Peeples ME. 2000. Glycosaminoglycan sulfation requirements for respiratory syncytial virus infection. *J. Virol.* 74:10508–13
137. Dong J, Dong L, Mendez E, Tao Y. 2011. Crystal structure of the human astrovirus capsid spike. *Proc. Natl. Acad. Sci. USA* 108:12681–86
138. Byrnes AP, Griffin DE. 1998. Binding of Sindbis virus to cell surface heparan sulfate. *J. Virol.* 72:7349–56
139. Gardner CL, Ebel GD, Ryman KD, Klimstra WB. 2011. Heparan sulfate binding by natural eastern equine encephalitis viruses promotes neurovirulence. *Proc. Natl. Acad. Sci. USA* 108:16026–31

140. Silva LA, Khomandiak S, Ashbrook AW, Weller R, Heise MT, et al. 2014. A single-amino-acid polymorphism in chikungunya virus E2 glycoprotein influences glycosaminoglycan utilization. *J. Virol.* 88:2385–97

141. Fry EE, Lea SM, Jackson T, Newman JW, Ellard FM, et al. 1999. The structure and function of a foot-and-mouth disease virus–oligosaccharide receptor complex. *EMBO J.* 18:543–54

142. Fry EE, Newman JW, Curry S, Najjam S, Jackson T, et al. 2005. Structure of foot-and-mouth disease virus serotype A10 61 alone and complexed with oligosaccharide receptor: receptor conservation in the face of antigenic variation. *J. Gen. Virol.* 86:1909–20

143. Jackson T, Ellard FM, Ghazaleh RA, Brookes SM, Blakemore WE, et al. 1996. Efficient infection of cells in culture by type O foot-and-mouth disease virus requires binding to cell surface heparan sulfate. *J. Virol.* 70:5282–87

144. Khan AG, Pickl-Herk A, Gajdzik L, Marlovits TC, Fuchs R, Blaas D. 2011. Entry of a heparan sulphate–binding HRV8 variant strictly depends on dynamin but not on clathrin, caveolin, and flotillin. *Virology* 412:55–67

145. Vlasak M, Goester I, Blaas D. 2005. Human rhinovirus type 89 variants use heparan sulfate proteoglycan for cell attachment. *J. Virol.* 79:5963–70

146. McLeish NJ, Williams CH, Kaloudas D, Roivainen MM, Stanway G. 2012. Symmetry-related clustering of positive charges is a common mechanism for heparan sulfate binding in enteroviruses. *J. Virol.* 86:11163–70

147. Tan CW, Poh CL, Sam IC, Chan YF. 2013. Enterovirus 71 uses cell surface heparan sulfate glycosaminoglycan as an attachment receptor. *J. Virol.* 87:611–20

148. O'Donnell J, Taylor KA, Chapman MS. 2009. Adeno-associated virus-2 and its primary cellular receptor—cryo-EM structure of a heparin complex. *Virology* 385:434–43

149. Levy HC, Bowman VD, Govindasamy L, McKenna R, Nash K, et al. 2009. Heparin binding induces conformational changes in adeno-associated virus serotype 2. *J. Struct. Biol.* 165:146–56

150. Xie Q, Spilman M, Meyer NL, Lerch TF, Stagg SM, Chapman MS. 2013. Electron microscopy analysis of a disaccharide analog complex reveals receptor interactions of adeno-associated virus. *J. Struct. Biol.* 184:129–35

151. Asokan A, Schaffer DV, Samulski RJ. 2012. The AAV vector toolkit: poised at the clinical crossroads. *Mol. Ther.* 20:699–708

152. Kaludov N, Brown KE, Walters RW, Zabner J, Chiorini JA. 2001. Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J. Virol.* 75:6884–93

153. Walters RW, Yi SMP, Keshavjee S, Brown KE, Welsh MJ, et al. 2001. Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J. Biol. Chem.* 276:20610–16

154. Wu Z, Miller E, Agbandje-McKenna M, Samulski RJ. 2006. α2,3 and α2,6 *N*-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J. Virol.* 80:9093–103

155. Agbandje-McKenna M, Kleinschmidt J. 2011. AAV capsid structure and cell interactions. *Methods Mol. Biol.* 807:47–92

156. Shen S, Bryant KD, Brown SM, Randell SH, Asokan A. 2011. Terminal *N*-linked galactose is the primary receptor for adeno-associated virus 9. *J. Biol. Chem.* 286:13532–40

157. Xie Q, Bu W, Bhatia S, Hare J, Somasundaram T, et al. 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. USA* 99:10405–10

158. Lerch TF, Xie Q, Chapman MS. 2010. The structure of adeno-associated virus serotype 3B (AAV-3B): insights into receptor binding and immune evasion. *Virology* 403:26–36

159. 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

160. Kern A, Schmidt K, Leder C, Muller OJ, Wobus CE, et al. 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J. Virol.* 77:11072–81

161. Zhang F, Aguilera J, Beaudet JM, Xie Q, Lerch TF, et al. 2013. Characterization of interactions between heparin/glycosaminoglycan and adeno-associated virus. *Biochemistry* 52:6275–85

162. Bell CL, Gurda BL, van Vliet K, Agbandje-McKenna M, Wilson JM. 2012. Identification of the galactose binding domain of the adeno-associated virus serotype 9 capsid. *J. Virol.* 86:7326–33

141. First structure of a virus in complex with a heparin trisaccharide at higher resolution.

163. DiMattia MA, Nam HJ, van Vliet K, Mitchell M, Bennett A, et al. 2012. Structural insight into the unique properties of adeno-associated virus serotype 9. *J. Virol.* 86:6947–58
164. Shen S, Horowitz ED, Troupes AN, Brown SM, Pulicherla N, et al. 2013. Engraftment of a galactose receptor footprint onto adeno-associated viral capsids improves transduction efficiency. *J. Biol. Chem.* 288:28814–23
165. Mietzsch M, Broecker F, Reinhardt A, Seeberger PH, Heilbronn R. 2014. Differential AAV serotype-specific interaction patterns with synthetic heparins and other glycans. *J. Virol.* 88:2991–3003
166. Raff AB, Woodham AW, Raff LM, Skeate JG, Yan L, et al. 2013. The evolving field of human papillomavirus receptor research: a review of binding and entry. *J. Virol.* 87:6062–72

165. Differential glycan-binding patterns of AAV serotypes.

OPEN ACCESS Freely available online

PLOS | PATHOGENS

A Structure-Guided Mutation in the Major Capsid Protein Retargets BK Polyomavirus

Ursula Neu[1☯¤a], Stacy-ann A. Allen[2☯], Bärbel S. Blaum[1,3], Yan Liu[4], Martin Frank[5], Angelina S. Palma[4¤b], Luisa J. Ströh[1], Ten Feizi[4], Thomas Peters[3], Walter J. Atwood[2]*, Thilo Stehle[1,6]*

1 Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany, **2** Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, United States of America, **3** Department of Chemistry, University of Luebeck, Luebeck, Germany, **4** Glycosciences Laboratory, Faculty of Medicine, Imperial College London, London, United Kingdom, **5** Biognos AB, Gothenburg, Sweden, **6** Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America

Abstract

Viruses within a family often vary in their cellular tropism and pathogenicity. In many cases, these variations are due to viruses switching their specificity from one cell surface receptor to another. The structural requirements that underlie such receptor switching are not well understood especially for carbohydrate-binding viruses, as methods capable of structure-specificity studies are only relatively recently being developed for carbohydrates. We have characterized the receptor specificity, structure and infectivity of the human polyomavirus BKPyV, the causative agent of polyomavirus-associated nephropathy, and uncover a molecular switch for binding different carbohydrate receptors. We show that the b-series gangliosides GD3, GD2, GD1b and GT1b all can serve as receptors for BKPyV. The crystal structure of the BKPyV capsid protein VP1 in complex with GD3 reveals contacts with two sialic acid moieties in the receptor, providing a basis for the observed specificity. Comparison with the structure of simian virus 40 (SV40) VP1 bound to ganglioside GM1 identifies the amino acid at position 68 as a determinant of specificity. Mutation of this residue from lysine in BKPyV to serine in SV40 switches the receptor specificity of BKPyV from GD3 to GM1 both *in vitro* and in cell culture. Our findings highlight the plasticity of viral receptor binding sites and form a template to retarget viruses to different receptors and cell types.

Citation: Neu U, Allen S-aA, Blaum BS, Liu Y, Frank M, et al. (2013) A Structure-Guided Mutation in the Major Capsid Protein Retargets BK Polyomavirus. PLoS Pathog 9(10): e1003688. doi:10.1371/journal.ppat.1003688

Editor: Denise A. Galloway, Fred Hutchinson Cancer Research Center, United States of America

Received July 1, 2013; **Accepted** August 21, 2013; **Published** October 10, 2013

Funding: This project was supported by SFB-685 to TS, 5R01CA71878-13 to WJA, P01-NS065719 to WJA and TS, as well as Wellcome Trust grants WT093378MA and WT099197MA, the UK Research Councils' Basic Technology Initiative 'Glycoarrays' (GRS/79268) and EPSRC Translational Grant (EP/G037604/1) to TF. ASP was supported by Fundação para a Ciência e Tecnologia (BPD SFRH/26515/2006, Portugal). 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. Dr. Martin Frank is employed by a company, Biognos AB in Gothenburg, Sweden. He declares that he does not have any financial, non-financial, professional or personal competing interests. His employment does not alter his and the other authors' adherence to all the PLOS PLoS Pathogens policies on sharing data and materials (as outlined in the guide for authors).

* E-mail: walter_atwood@brown.edu (WJA); thilo.stehle@uni-tuebingen.de (TS)

☯ These authors contributed equally to this work.

¤a Current address: National Institute of Medical Research, The Ridgeway, Mill Hill, London, United Kingdom.
¤b Current address: REQUIMTE, CQFB, Faculty of Science and Technology /UNL, Caparica, Portugal.

Author Summary

Viruses need to bind to receptors on host cells for viral entry and infection, and the type of receptor bound determines the range of hosts and tissues the virus can infect. Viruses within a family often vary in their tissue distribution and pathogenicity because changes in receptor specificity can produce a virus with different spread and infectivity. In fact, many transmissions between species are based on a virus acquiring binding capability for a new receptor. The structural changes that underlie such receptor switching are not well understood. We have analyzed the structural requirements for receptor binding and switching of the human BK polyomavirus (BKPyV), the causative agent of polyomavirus-associated nephropathy. We show that BKPyV uses specific gangliosides that all contain a common α2,8-disialic acid motif to infect cells, and have characterized the interaction in atomic detail. Our data explains the requirement for this disialic acid motif and in particular highlights a single amino acid that is central to determining specificity. Mutation of this residue switches the receptor specificity, enabling BKPyV to infect cells bearing a different class of gangliosides. Our findings highlight the plasticity of viral receptor binding sites and form a template to retarget viruses to different receptors and cell types.

Introduction

Interactions of a virus with receptors on host cells are crucial for viral entry and infection, and determine host range and tissue tropism of the virus. As a result, receptor specificity and affinity are tightly regulated, with changes in either producing a virus with different spread and infectivity. Many zoonotic transmissions are based on a virus acquiring binding capability for a new receptor. Despite their importance, most viral specificity switches are poorly characterized, especially for carbohydrate-binding viruses, as methods capable of structure-specificity studies are only being developed for carbohydrates. Here we characterize the receptor specificity of the human BK Polyomavirus in depth, and retarget it to use the Simian Virus 40 (SV40) receptor GM1.

The human polyomavirus BK Virus (BKPyV) is a non-enveloped, double-stranded DNA (dsDNA) virus that belongs to the family *Polyomaviridae*. Other members of the family include Simian Virus 40 (SV40), the human JC Virus (JCPyV), Merkel Cell Polyomavirus (MCPyV) and at least eight other recently discovered human polyomaviruses [1]. BKPyV was first isolated from a kidney transplant recipient in 1971 [2]. It establishes a persistent asymptomatic infection in the genitourinary tract of approximately 70% of the adult population [3,4,5]. A key modulator of BKPyV reactivation is immunosuppression of the host that leads to an increase in viral replication [3]. Complications of BKPyV reactivation include the development of polyomavirus-induced nephropathy (PVN) in kidney transplant

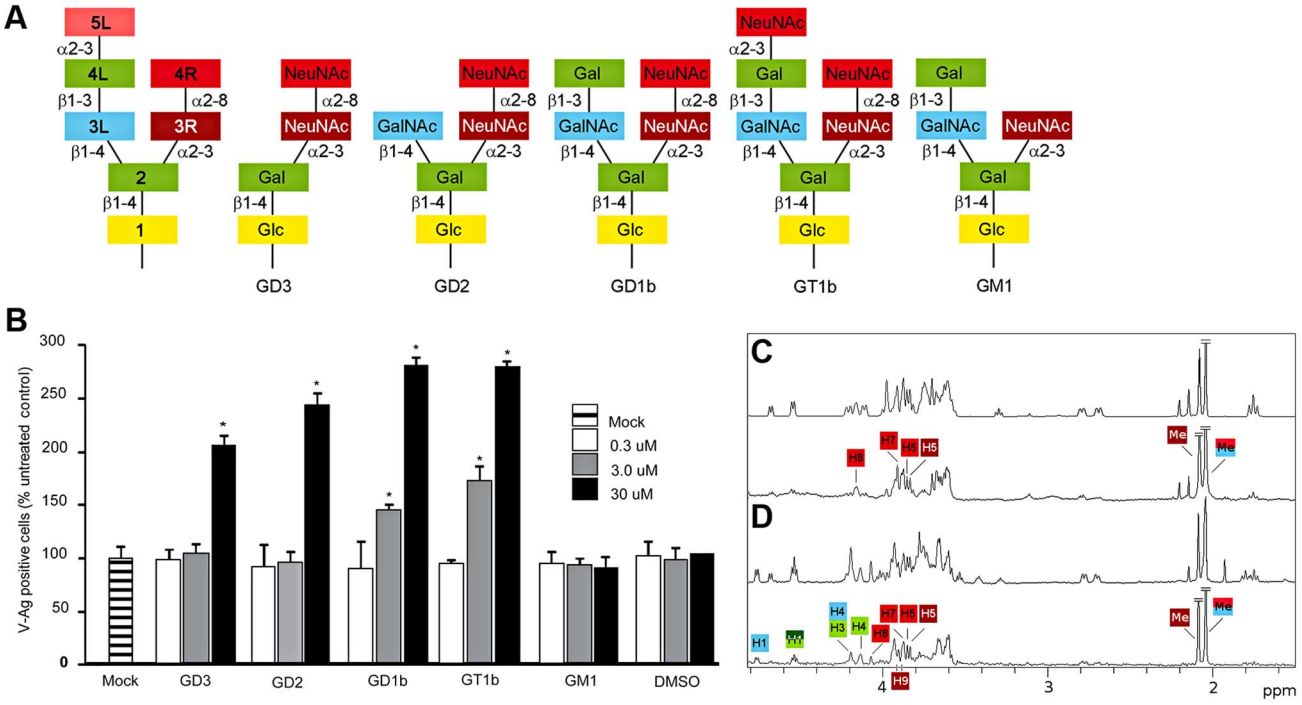


Figure 1. B-series gangliosides are receptors for BKPyV. (A) Schematic representation of structures of disialic acid-containing b-series gangliosides GD3, GD2, GD1b, GT1b and of the monosialylated a-series ganglioside GM1 indicating the abbreviated designations used here for individual residues in the gangliosides. (B) Ganglioside supplementation assays. Vero cells were incubated with gangliosides, challenged with BKPyV and scored for infection. The average number of VP1 positive cells is plotted compared to controls. Error bars represent the standard deviation for 3 independent experiments. Asterisks indicate p-value (*$p<0.05$). (C) STD off-resonance (top) and difference (bottom) spectra of WT BKPyV VP1 in the presence of 50-fold excess GD3 oligosaccharide. The off-resonance spectrum was scaled to 3%. GD3 resonances labeled in the difference spectrum receive considerable saturation transfer from the protein. Regions with strong signal overlap are not labeled because saturation effects in this region cannot be unambiguously assigned. Signals that were truncated are denoted by diagonal bars. (D) STD off-resonance (top) and difference spectrum (bottom) of wild type BKPyV in the presence of 50-fold excess GD1b oligosaccharide. The spectrum is labeled as in C.
doi:10.1371/journal.ppat.1003688.g001

recipients, and hemorrhagic cystitis in bone marrow transplant recipients [3,6,7].

BKPyV attachment is mediated by cell-surface sialic acid [8]. The most common sialic acid type in humans is 5-*N*-acetyl neuraminic acid (NeuNAc) [9]. Simians and most other mammals, however, possess an enzyme that can attach an additional hydroxyl group to NeuNAc, yielding 5-*N*-glycolyl neuraminic acid (NeuNGc). In contrast to humans, these animals therefore carry both NeuNAc and NeuNGc. Gangliosides were found to mediate cell attachment of BKPyV [10], and gangliosides GD1b and GT1b were later shown to function as specific receptors for BKPyV [11] (Fig. 1). Gangliosides are ceramide-based glycolipids, which are used as receptors for most of the well-characterized polyomaviruses, for example GD1a and GT1b for Polyoma, or GM1 for SV40 [12].

A polyomavirus capsid consists of 72 pentamers of the major capsid protein VP1 [13,14]. Crystal structures of Polyoma, SV40, MCPyV and JCPyV VP1 show that receptors are bound in shallow grooves formed by VP1 loop structures on the outer surface of the virus. These loops contribute to different receptor specificities and are the only parts of VP1 that are not well conserved [14,15,16,17,18].

In this study, we use viral infection assays to demonstrate that the b-series gangliosides GD3, GD2, GD1b and GT1b enhance BKPyV infection. We then define the common α2,8-disialic acid motif on these gangliosides as the primary binding epitope for BKPyV by NMR spectroscopy. In order to understand how the disialic acid motif is recognized by the virus, we solved the crystal structure of a BKPyV VP1 pentamer in complex with GD3 and generated a model of the pentamer with the larger GD1b oligosaccharide. Analysis of these complexes reveals extensive interaction with the terminal sialic acid, specificity-defining contacts with the internal sialic acid, thus explaining the requirement for a disialic acid motif, and additional contacts with the branched GD1b. Mutagenesis of residues contacting the disialic acid motif abolishes infectivity and a comparison with the SV40 VP1-GM1 complex attributes the different viral receptor specificities to one point mutation. Introduction of this mutation into BKPyV switches specificity, enabling BKPyV to bind GM1 and abolishes binding to GD3, as shown by saturation transfer difference (STD) NMR spectroscopy and carbohydrate microarray analyses. The microarray analyses moreover reveal that the mutant is specific for the 'human' sialic acid NeuNAc. This contrasts with SV40 VP1, which has a preference for the more prevalent NeuNGc found in simian species and many other nonhuman mammals. The specificity of the mutant thus sheds light on the influence of sialic acid on species tropism.

Results

All b-series gangliosides support infection of BKPyV

To date, only four gangliosides had been tested to support BKPyV infection, two of which, the b-series gangliosides GD1b and GT1b, were confirmed as receptors [11]. The carbohydrate moieties of gangliosides typically consist of two "arms" (Fig. 1A). In this manuscript, we number the carbohydrates sequentially (starting from the lipid anchor) and use "L" and "R" designations to indicate whether a carbohydrate is part of the left or the right arm, respectively. For example, NeuNAc 3R is the third carbohydrate and located in the right arm (Fig. 1A, schematic on the left hand side). The right arm consists of only sialic acids and is used to classify gangliosides. The b-series gangliosides, e.g. GD1b and GT1b, carry both NeuNAc 3R and NeuNAc 4R, while a-series gangliosides such as GM1 carry only NeuNAc 3R (Fig. 1A). We tested the effects of all common b-series gangliosides on BKPyV infection by supplementing permissive Vero cells with GD3, GD2, GD1b or GT1b (Fig. 1B). Consistent with previous reports [11], gangliosides GD1b and GT1b enhanced infectivity of Vero cells. However, gangliosides GD2 and GD3, which had not been tested previously, also enhanced infection of the cells (Fig. 1B). Incorporation of all b-series gangliosides into the plasma membrane of Vero cells increased the binding of labeled VP1 pentamers to cells (data not shown). In control experiments, supplementing cells with the a-series ganglioside GM1 had no effect on infection or binding (Fig. 1B, and data not shown). The ability of b-series gangliosides to enhance BKPyV infection was greater in the presence of the left arm, with GD1b and GT1b supporting infection best (Fig. 1B). Taken together, these data show that the α2,8-disialic acid motif in b-series gangliosides is the minimal requirement for binding, with the left arm of GD1b and GT1b contributing some additional interactions.

BKPyV carbohydrate epitope mapping on b-series gangliosides

We analyzed pentamer binding to GD3 and GD1b oligosaccharides by STD NMR spectroscopy [19], which identifies ligand atoms that contact a protein in solution. The strongest saturation transfer from BKPyV VP1 to GD3 was observed for the methyl group of the terminal NeuNAc 4R, followed by the methyl group of the internal NeuNAc 3R (Fig. 1C). No significant transfer was observed to any of the anomeric protons, or to the NeuNAc H3 protons. Interestingly, no transfer was observed for the Glc and Gal residues of GD3 (Fig. 1C), suggesting that they may not participate in binding.

We repeated the same experiment for GD1b oligosaccharide, observing transfer to essentially the same set of protons from the disialyl moiety plus additional transfer to Gal 4L and GalNAc 3L in the GD1b left arm as well as the branching Gal 2 residue (Fig. 1D). Resonances H4 and H1 from Gal 2 and H1 from GalNAc 3L can be unambiguously assigned, but H3 from Gal 2 and H4 from GalNAc 3L overlap and cannot be distinguished. From the Gal 4L ring, only the anomeric proton can be assigned in the STD difference spectrum. The STD spectra thus show that while the right arms of both GD3 and GD1b interact with BKPyV VP1 in a similar way, additional contacts are provided by the left arm of GD1b.

Structure of a BKPyV VP1-GD3 complex

To define the structural features underlying the receptor-binding specificity of BKPyV, we solved the structure of the BKPyV VP1 pentamer at 2.0 Å resolution (Table 1). The VP1 pentamer is a doughnut-shaped ring, with the five monomers arranged around a central pore that aligns with the five-fold symmetry axis (Fig. 2A). The monomers adopt a β-sandwich fold with jelly-roll topology that is present in many viral capsid proteins. The β-strands B, I, D, G and C, H, E, F (designated alphabetically from the N-terminus of the full-length protein) are linked by extensive loops that decorate the surface of the protein. For clarity, the long BC-loop is subdivided into loops BC1 and BC2, which face in different directions.

We incubated BKPyV VP1 crystals in 20 mM GD3 oligosaccharide solution and solved the structure of the resulting complex at 1.7 Å resolution (Table 1). Attempts to form a complex with GD1b oligosaccharide by soaking or co-crystallization failed in several crystal forms, likely because of the size of the GD1b hexasaccharide. However, the GD3 structure encompasses the

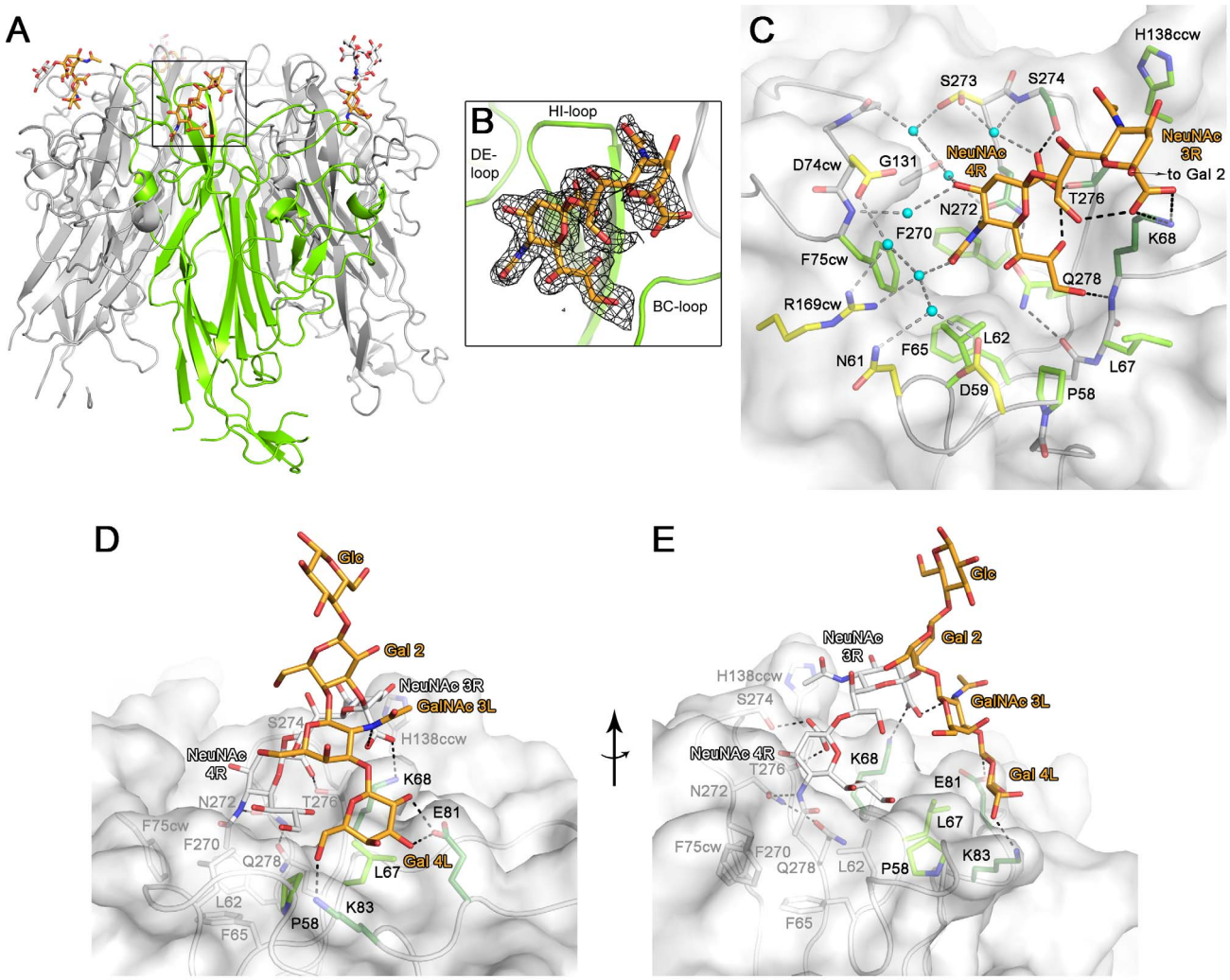


Figure 2. Structure of a BKPyV VP1-GD3 oligosaccharide complex. (A) Structure of BKPyV VP1 pentamer in complex with GD3 oligosaccharide. One VP1 monomer is highlighted in green. Monosaccharides unbiased by crystal contacts are colored orange, while the ones binding to crystal contacts are colored white. (B) Composite annealed difference electron density for the terminal disialic acid motif of GD3 oligosaccharide bound to BKPyV VP1 at a σ level of 2.5. cw/ccw = belonging to the clockwise/counterclockwise neighboring VP1 monomer within the pentamer. (C) Interactions of GD3 oligosaccharide with BKPyV VP1. The oligosaccharide is shown in orange. Side chains contacting the sugar are colored as follows: those making hydrogen bonds are colored dark green, those making van der Waals interactions are light green, and those making water-mediated hydrogen bonds are colored yellow. Atoms of the protein backbone are shown in gray, and water molecules are in cyan. Direct hydrogen bonds are indicated as black dashed lines, water-mediated ones are grey. (D+E) Model of the interaction of BKPyV VP1 with GD1b oligosaccharide. The left arm and the stem of GD1b are colored orange. Residues interacting with the left arm of GD1b are colored as in (C). The disialic acid motif of GD1b and the protein residues contacting it are colored white.
doi:10.1371/journal.ppat.1003688.g002

minimal motif required for binding and provides a template for understanding interactions with the larger GD1b oligosaccharide. The structures of unliganded and liganded BKPyV VP1 are virtually identical (r.m.s.d. of 0.4 Å for all atoms of one monomer), indicating that GD3 binding does not induce a conformational change in the protein. GD3 engages BKPyV VP1 at the top of the pentamer, which corresponds to the outer surface of the virion (Fig. 2A, B). Contacts involve residues in the BC1-, HI- and DE-loops of one monomer, as well as the BC2-loop of the clockwise neighboring VP1 monomer (BC2cw) and the DE-loop of the counterclockwise neighbor (DEccw). Four of the five binding sites within one pentamer are occupied by ligand, while one is inaccessible due to crystal packing. Two of the four occupied binding sites, however, participate in crystal contacts. As their conformations are influenced by these non-physiologic interactions, they will not be considered further. The two remaining GD3 oligosaccharides do not participate in crystal contacts and assume an identical conformation, which therefore should represent a physiologically relevant complex. In both cases, only the terminal NeuNAc-α2,8-NeuNAc motif interacts with BKPyV VP1, while the Gal-Glc moiety projects into solution. This is in agreement with the STD NMR data that showed little saturation transfer between BKPyV VP1 and the Gal-Glc portion of GD3.

Contacts between BKPyV VP1 and GD3

The terminal NeuNAc 4R is the main contact of GD3 with BKPyV VP1. In all four occupied binding sites on the BKPyV VP1 pentamer, NeuNAc 4R adopts the same conformation and makes identical interactions with the protein. The sialic acid

Table 1. Crystallographic data collection and refinement statistics.

	BKPyV VP1		BKPyV VP1 - GD3 complex	
Data collection				
Space group	$P2_12_12$		$P2_12_12$	
Unit cell [Å]	144.9, 152.3, 62.7		144.7, 152.6, 63.2	
Resolution [Å]	50 - 2.0	2.05 - 2.00	50 - 1.7	1.74 - 1.70
Total reflections	388,470	27,741	1,103,641	64,273
Unique reflections	94,342	6,826	154,169	11,266
R_{meas} [%]	12.7	67.0	7.1	69.4
Completeness [%]	99.8	98.4	100.0	99.8
I/σI	11.01	2.40	19.44	2.66
Refinement				
R_{work} [%]	16.1	22.0	15.0	23.1
R_{free} [%]*	19.0	25.1	17.7	24.9
Protein atoms	9,957		10,181	
Average B-factor [Å^2]	23.1		21.5	
Carbohydrate atoms	-		197	
Average B-factor [Å^2]	-		33.1	
Solvent atoms	899		1,327	
Average B-factor [Å^2]	30.0		32.6	
r.m.s.d.				
Bond lengths [Å]	0.010		0.010	
Bond angles [°]	1.246		1.311	
B-factor Wilson [Å^2]	27.1		26.4	

r.m.s.d. = root-mean-square deviation.
*R_{free} was calculated with 3.5% of the data.
doi:10.1371/journal.ppat.1003688.t001

carboxylate group is recognized by two hydrogen bonds to the side chains of S274 and T276 (Fig. 2C). Additional, water-mediated hydrogen bonds are formed to the side chain of S273 and the backbone of S274. The O4 hydroxyl group of NeuNAc 4R interacts via water-mediated hydrogen bonds with the side chain of N272, the backbone of G131, and the backbone nitrogen of F75cw. The *N*-acetyl group makes a hydrogen bond to N272 and a water-mediated hydrogen bond to R169cw. Its methyl group inserts into a tight-fitting, hydrophobic pocket on BKPyV VP1 that is formed by four non-polar residues (L62, F65, F270 and F75cw). Finally, the glycerol chain of NeuNAc lies in a shallow groove and makes van der Waals contacts to the side chains of P58, L62, L67, K68 and Q278. The glycerol chain also forms a single hydrogen bond to the K68 backbone nitrogen. The conformation of the NeuNAc 4R ring is stabilized by an intramolecular hydrogen bond from the carboxylate group to the O8 hydroxyl group.

The second sialic acid of GD3, NeuNAc 3R, has weaker electron density and makes fewer contacts with the protein (Fig. 2C). Its carboxylate group forms a salt bridge with the K68 side chain. The methyl group of its *N*-acetyl chain stacks against a hydrophobic surface created by parts of the side chains of H138ccw, S274 and T276. The prominent role of van der Waals interactions with the methyl group mirrors the NMR results, which feature prominent saturation transfer to the methyl groups of both NeuNAcs (Fig. 1C).

Modelling of the BKPyV-GD1b interaction reveals additional contacts

The BKV-GD3 complex structure enabled us to model the interaction of BKV VP1 with the longer GD1b oligosaccharide, using the tightly bound terminal NeuNAc 4R as an anchor. A large number of possible GD1b oligosaccharide conformations was calculated and superposed on the terminal NeuNAc 4R in the BKV-GD3 complex structure. The oligosaccharide conformations were filtered for presence of the specificity-defining contacts between the protein and the internal NeuNAc 3R. The remaining conformations could be classified into two groups: one in which the left arm of GD1b pointed away from the protein, not engaging in interactions, and one in which this arm made additional contacts with BKV VP1. Some of the latter conformations were then subjected to a final round of molecular dynamics (MD) simulations in explicit water. We found that the 'left' arm is involved in several weaker interactions with amino acids P58, D59, L67, K68, E81, and K83. A snapshot of the complex is shown in (Fig. 2D,E). According to this model, the Gal 4L residue of GD1b can adopt a position enabling hydrogen bonds between its 2- and 3-hydroxyl groups and the side chain of E81, as well as hydrogen bonds between its 6-hydroxyl group and the side chain of K83 and the backbone carbonyl of P58. Moreover, the left arm of GD1b is supported by van der Waals interactions with the side chain of L67, and there is an intramolecular hydrogen bond between the carboxylate group of the internal NeuNAc 3R and the *N*-acetyl group of GalNAc 3L. The model is in accord with the observed increase in BKV infection with increasing length of the left arm of b-series gangliosides, and also with the STD NMR data that show signals for some protons in the left arm.

Carbohydrate binding is crucial for BKPyV spread and infectivity

To test the biological relevance of these interactions mutations were introduced into an infectious clone of BKPyV. We first probed the interaction with the tightly bound terminal NeuNAc 4R (Fig. 3A) with mutations designed to abolish carbohydrate binding either by reducing the number of hydrogen bonds (S274A, T276A, and S274A/T276A), eliminating van der Waals contacts (F75V), or by introducing steric hindrance (L62W, F75W). Vero cells were transfected with mutant or wild-type (WT) BKPyV plasmid DNA. Viral gene expression was scored every 3 days over a 22 day growth period. The first data point after transfection indicated no difference between wild-type and the mutants in terms of protein expression and localization (data not shown). While WT BKPyV resulted in viral production that continued to spread with time, all mutants that targeted the binding site for terminal sialic acid did not propagate, highlighting the importance of these interactions (Fig. 3B). We then targeted the binding site of the internal NeuNAc 3R. Mutant H138A, in which a van der Waals contact is removed from the second sialic acid, propagated at a significantly reduced level compared to WT (Fig. 3B). We also introduced the mutations in our recombinant pentamer construct. Flow cytometry binding assays using WT and mutant pentamers show that the mutants have greatly reduced cell binding (Fig. 3C), suggesting that the loss of viral propagation is due to an attachment defect. The structural integrity of mutant pentamers was verified with circular dichroism spectroscopy, and their ability to assemble into pentamers was confirmed with gel filtration (data not shown). All the mutations were in the receptor binding site, which is distant from the sites important for capsid assembly. Thus, the mutations are very unlikely to cause defects in capsid assembly.

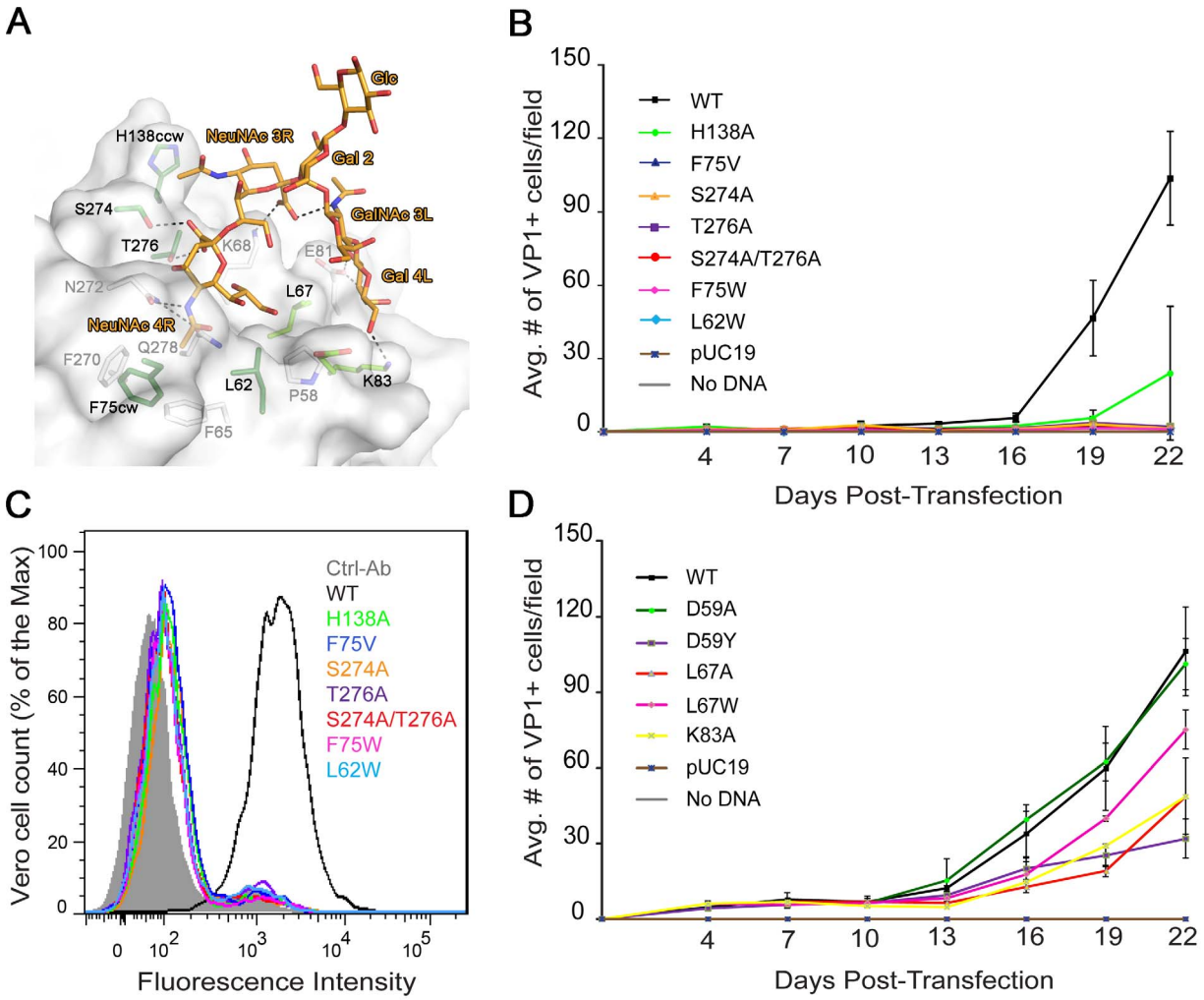


Figure 3. Mutation of residues that interact with the left and right branches of GD1b abolish or reduce growth. (A) Structural model of the BKPyV-GD1b complex. Residues targeted for mutation are colored dark green for the disialic acid motif and light green for the left arm binding sites, respectively. (B) Growth assay for the sialic acid binding site mutants. Vero cells were transfected with linearized WT or mutant BKPyV DNA. Cells were fixed permeabilized and stained for VP1 at 3 day intervals for 22 days and analyzed by indirect immunofluorescence. Viral spread was quantified by scoring for cells expressing VP1. The average number of VP1 positive cells is plotted from 3 independent experiments with each time point representing the average number of infected cells per visual field for 8 fields. (C) Binding of the sialic acid binding site mutants to Vero cells. Cells were incubated with purified His-tagged WT or mutant BKPyV pentamers and an Alexa Fluor 488 conjugated penta-His secondary antibody. Cells were fixed and pentamer binding to cells assessed by flow cytometry. Histograms show the fluorescence intensity of the Alexa 488 antibody alone (gray-filled), WT pentamer (black) and mutants (color) for 1×10^4 events. (D) Growth assay for the left arm binding site mutants assessed as in B.
doi:10.1371/journal.ppat.1003688.g003

Finally, we mutated residues in the putative binding site for the left arm of b-series gangliosides (Fig. 3D). Again, we designed mutations to either introduce steric hindrance (D59Y and L67W) or to remove contacts (D59A, L67A and K83A). The mutations that created steric hindrance significantly reduced BKPyV spread in culture. The D59A mutation had no significant effect, but L67A and K83A both reduced BKPyV growth. In addition, the E81A mutant, which also removes a contact from the second arm, was described in an earlier paper to have slightly reduced growth [20]. The mutations likely did not abolish growth altogether because they still permit interactions with the primary contact, terminal sialic acid. While the D59Y and L67 mutations might in theory also interfere with primary sialic acid binding due to the branched nature of GD1b, the K83A and E81A mutations certainly only target the second arm. Thus, mutagenesis confirms our structural model and highlights the importance of specific contacts with the second arm of GD1b. Taken together, our biological data confirm that the binding site for terminal sialic acid is indispensable for viral infection, while peripheral interactions further enhance binding and infection.

Structural basis of specificity

BKPyV is most closely related to SV40 and JCPyV, with amino acid identity among their VP1 proteins as high as 74%. Nevertheless, the three viruses recognize different sialic acid containing receptors. BKPyV interacts with α2,8-linked b-series gangliosides, while SV40 binds the branched α2,3-linked GM1 ganglioside [12] and JCPyV attaches to the linear α2,6-linked sequence in LSTc [17].

In receptor complexes of all three viruses, the terminal sialic acid engages in critical and highly conserved interactions that anchor the ligand to VP1 (Fig. 4) [16,17]. Specificity for the three

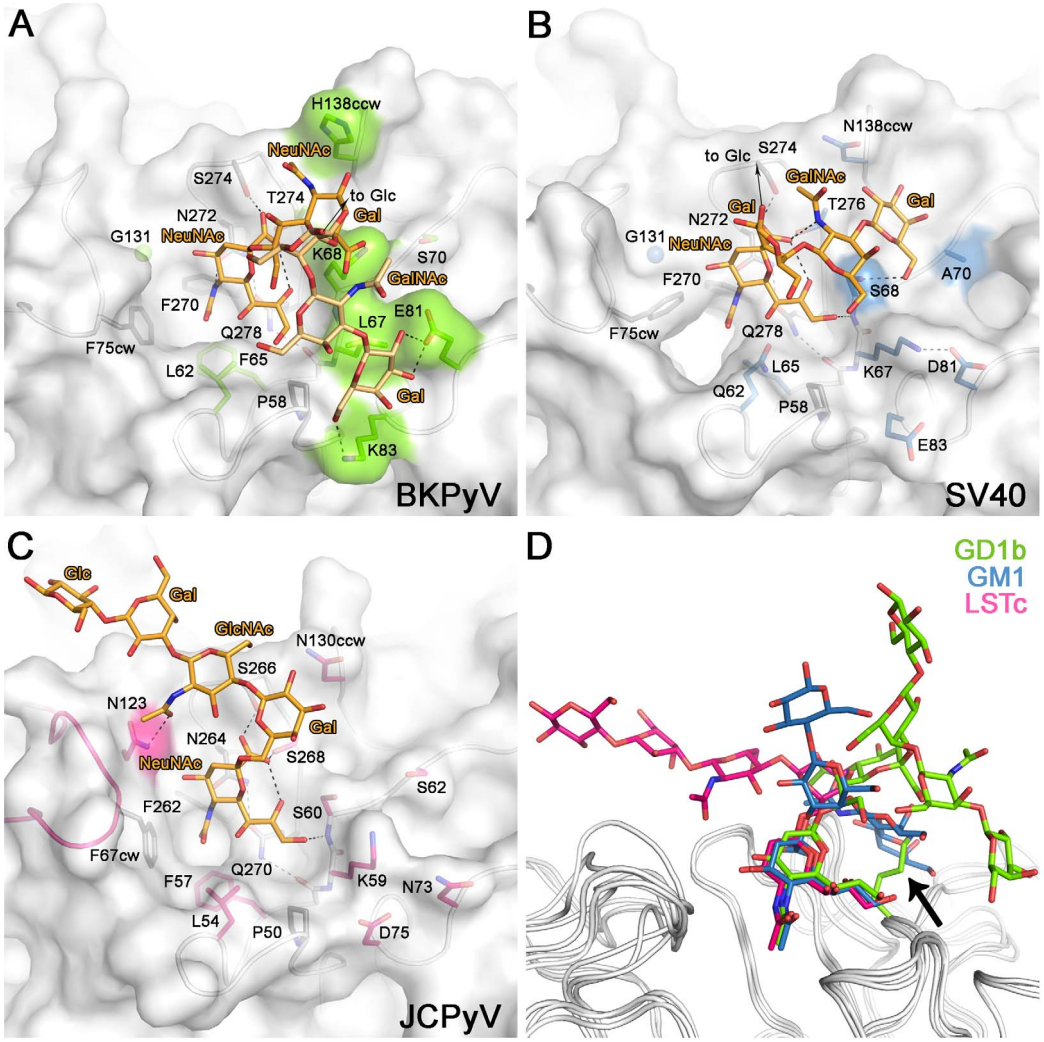


Figure 4. Carbohydrate binding sites of BKPyV, JCPyV and SV40. Recognition of carbohydrate receptors by BKPyV (A), SV40 (B), and JCPyV (C). Only residues making direct hydrogen bonds or van der Waals interactions are shown. Residues that make conserved interactions are colored in yellow, side chains whose positions are not conserved are shown in green (BKPyV), blue (SV40) and pink (JCPyV). (D) Comparison of carbohydrate ligands of BKPyV, SV40 and JCPyV. The oligosaccharides are colored green, blue and pink for the BKPyV, SV40 and JCPyV ligands, respectively. The structures were aligned using the protein main chain.
doi:10.1371/journal.ppat.1003688.g004

different oligosaccharide receptor sequences arises in each case from a small number of unique contacts outside the sialic acid binding site. JCPyV recognizes an L-shaped conformation of the LSTc oligosaccharide. The key residue that makes contacts to both legs of this L-shaped glycan is N123 [17] (Fig.4C). BKPyV (Fig. 4A) and SV40 (Fig. 4B) both have a glycine at the equivalent position and thus cannot form similar contacts. BKPyV VP1 specificity for the α2,8-disialyl motif can be attributed to residues K68 and H138, which form contacts with the internal NeuNAc 3R. These two residues are not conserved in SV40 or JCPyV, and thus neither virus is able to specifically interact with α2,8-disialic acid carrying glycans in the same manner [16,17]. Moreover, none of the BKPyV residues contacting the left arm of gangliosides are conserved in either SV40 or JCPyV. Although the SV40 receptor, GM1, resembles GD1b with an identical left arm, BKPyV and SV40 VP1 bind the left arm at different sites on the proteins (Fig. 4A,B). The two pockets recognize the left arm in different ways. The GalNAc 3L methyl group of GM1 in the SV40 complex is bound at a similar position on VP1 as the NeuNAc 3R methyl group in the BKPyV complex. BKPyV cannot bind the left arm of GM1 in the orientation seen in the SV40 complex because the binding site is blocked by the large side chain of K68, which would lead to clashes with GalNAc (Fig. 4A,B). Apart from this difference, BKPyV and SV40 VP1 share similar surface features at the SV40 left arm binding site and display the same main chain conformation in their surface loops. Thus, the inability of BKPyV to bind to GM1 appears to be determined by the amino acid at position 68.

A single point mutation enables BKPyV to recognize GM1

To validate the conclusions derived from the structural comparisons, we introduced a K68S mutation into the BKPyV VP1 pentamer expression construct. Purified K68S pentamers were analyzed by STD NMR for binding to GD3 and GM1. Unlike the WT BKPyV-GD3 pair, almost no saturation transfer was observed for BKPyV K68S and GD3, indicating that the mutation virtually abolished binding to the disialic acid motif of GD3 (Fig. 5A). However, saturation transfer from BKPyV K68S

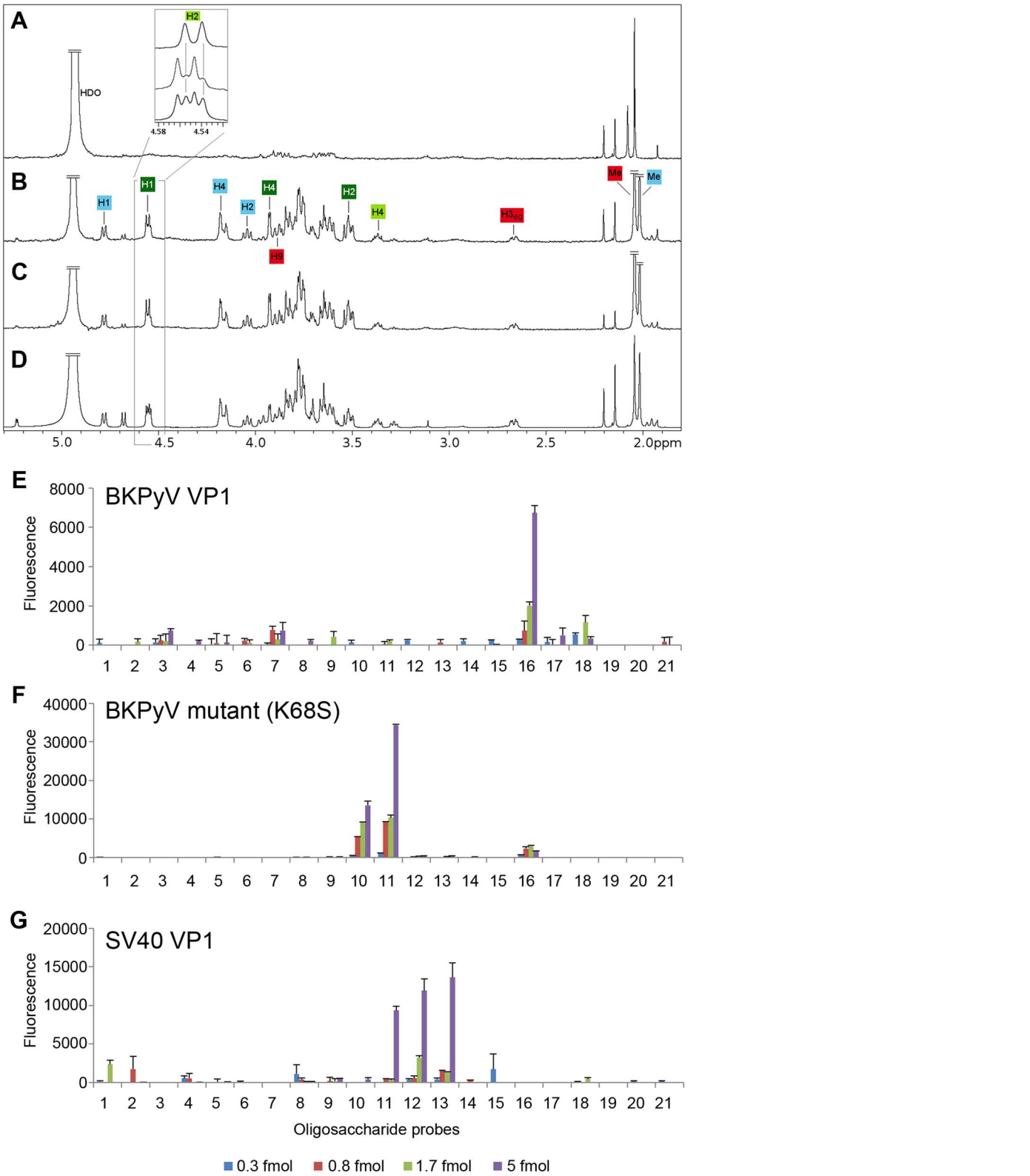


Figure 5. The K68S mutation targets BKPyV to the SV40 receptor GM1. STD difference spectra of (A) BKPyV K68S with GD3, (B and C, respectively) BKPyV K68S and SV40 with GM1. (D) SV40-GM1 off-resonance spectrum. A 50-fold excess of oligosaccharide was used for each spectrum. The off-resonance spectrum was scaled to 3%. Resonances labeled in the difference spectra with GM1 (B and C) receive considerable saturation transfer from BKPyV K68S and SV40. Regions with strong signal overlap are not labeled because they cannot be unambiguously assigned. Binding of BKPyV K68S to GD3, previously seen for WT BKPyV (Fig. 1C), is abolished by the mutation (A). Carbohydrate microarray analyses of recombinant VP1 of (E) BKPyV, (F) BKPyV K68S and (G) SV40 using 21 ganglioside-related saccharide probes, which included the b-series gangliosides as well as GM1 variants NeuNAc-GM1 and NeuNGc-GM1. The doses of probes arrayed per spot are indicated. Numerical scores of the binding signals are means of duplicate spots (with error bars). The complete list of probes and their sequences are in Supplemental Table S1.
doi:10.1371/journal.ppat.1003688.g005

VP1 to GM1 was as efficient as for the SV40 VP1-GM1 pair, which was included for comparison (Fig. 5B–D). This indicates that the K68S mutation switches the binding preference of BKPyV VP1 from GD3 to GM1. The STD NMR spectra of SV40 VP1 and BKPyV K68S VP1 with GM1 are almost indistinguishable, suggesting that GM1 engages in the same contacts with both proteins. Saturation transfer is primarily observed to protons of the NeuNAc 3R and Gal 4L rings. In addition, both the GalNAc and the NeuNAc methyl groups in GM1 received considerable saturation in the complexes, with the NeuNAc methyl group being more affected. Our observations are in good agreement with the crystal structure of the SV40 VP1-GM1 complex [16] and demonstrate that a single amino acid mutation suffices for BKPyV to adapt to the SV40 receptor.

BKPyV K68S mutant is specific for the human sialic acid NeuNAc

BKPyV, JCPyV and SV40 differ in one aspect of their sialic acid binding site. The cavity engaging the methyl group is tight-fitting and lined with hydrophobic residues in BKPyV and JCPyV, but significantly enlarged and partially hydrophilic in SV40 (Fig. 4A–C). This difference may reflect the different hosts of these viruses, humans and simians, and the different types of sialic acids characteristic for each host. While the most prominent sialic acid in humans is NeuNAc, simians carry in addition to NeuNAc larger amounts of *N*-glycolyl neuraminic acid (NeuNGc), in which the methyl group is replaced by the bigger and more hydrophilic glycolyl (CH_2-OH) group [9,21]. SV40 preferentially binds to NeuNGc-GM1, and the glycolyl group likely engages polar residues in the cavity [16,22]. By contrast, the smaller and more hydrophobic cavity of BKPyV and JCPyV cannot accommodate the glycolyl group in a similar manner, thus making BKPyV and JCPyV specific for NeuNAc.

To assess their specificity for human-type and simian-type sialic acids, the VP1 proteins of WT BKPyV, mutant BKPyV K68S, and SV40 were analyzed using a focused ganglioside microarray comprised of 21 ganglioside-related saccharide probes, which included the b-series gangliosides as well as GM1 variants NeuNAc-GM1 and NeuNGc-GM1 (Supplemental Table S1). Microarray analyses revealed differing binding specificities of the three VP1 proteins (Fig. 5E–G). With the WT BKPyV, the only detectable binding was to the b-series ganglioside GD1b (position 16) and the signal intensity was relatively low. No binding to any other b-series gangliosides GD3, GD2, GT1b or GQ1b was detected, possibly due the lower binding avidity to these probes compared to GD1b in the array (Fig. 5E). The BKPyV K68S mutant showed barely detectable binding to GD1b, but highly specific and strong binding to the two NeuNAc-GM1 probes (positions 10 and 11), which differed only in the composition of their lipid moieties (Fig. 5F). Interestingly, there was no binding to the simian-type NeuNGc-GM1 probes, in contrast to SV40 VP1, which showed preferential binding to the two NeuNGc-GM1 probes (positions 12 and 13) (Fig. 5G). This finding is in accord with earlier observations and consistent with our structural analysis (Fig. 4).

BKPyV K68S uses NeuNAc-GM1 as a receptor on human cells

We next tested whether BKPyV K68S was able to use human-type NeuNAc-GM1 to attach to cells. Purified K68S or WT BKPyV VP1 pentamers were incubated with simian (Vero) and human (HEK) cells, and binding was detected by flow cytometry. In both Vero and HEK cells, K68S mutant pentamers had reduced binding compared to WT pentamers, reflecting either lower affinity or a lower number of receptor molecules (Fig. 6A,C). There was no significant change in the binding of K68S or WT VP1 pentamers to Vero cells that were supplemented with 3 μM NeuNAc-GM1 prior to incubation with the pentamers (Fig. 6A). However, binding of K68S VP1 to HEK cells was increased upon supplementation with GM1, whereas WT binding levels were unchanged (Fig. 6C). This finding might be linked to the enzyme CMP-sialic acid hydroxylase, which converts NeuNAc to NeuNGc and is present on simian but not human cells [9]. We performed a competitive binding assay with GM1 treated HEK cells in the presence and absence of Cholera toxin subunit B (CTX), which uses GM1 as a receptor [23]. CTX abolished binding of K68S VP1, confirming that the K68S mutant is in fact retargeted to GM1 (Fig. 6C). CTX had no effect on the binding of WT BKPyV VP1 (Fig. 6C).

The K68S mutation was also assayed in long-term viral growth assays. We found that while transfection of WT BKPyV plasmid into Vero cells resulted in viral propagation and spread, transfection with the K68S plasmid failed to propagate (Fig. 6B). In human cells however the K68S mutant spread as efficiently as WT regardless of supplementation with GM1 (Fig. 6D).

Discussion

In this structure-function study, we investigated the interaction of BKPyV with its glycan receptors and identified key determinants of specificity. We show that the conserved α2,8-disialic acid motif on the right arm of b-series gangliosides is the minimal binding epitope for BKPyV, with the variable left arm contributing some additional contacts. Point mutations in the receptor binding site abolish viral spread and infectivity, demonstrating the physiological relevance of the observed interactions.

Our data demonstrate that all of the b-series gangliosides tested support BKPyV infection. As attachment likely requires multiple interactions, the virus is predicted to engage a mixture of gangliosides on the cell surface, depending on lipid composition. While gangliosides are likely entry receptors for BKPyV, the main binding epitope of BKPyV, α2,8-disialic acid, is not only present on gangliosides, but also on glycoproteins. It has been shown for another ganglioside-binding polyomavirus that such sialylated glycoproteins act as decoy receptors [24]. The additional contacts with the left arm of b-series gangliosides therefore may increase BKPyV binding affinity for ganglioside ligands and distinguish those from glycoproteins, which likely would lead the virus along non-infectious entry pathways.

The importance of b-series gangliosides for BKPyV infection may have implications for BKPyV tropism and pathogenesis. Biochemical analyses indicate that the kidney, where BKPyV persists, is rich in diverse sphingolipids and particularly gangliosides. The most abundant gangliosides in adult human kidney are GM3 and GD3, but small amounts of more complex gangliosides were also detected [25,26]. The relative abundance of simple gangliosides differs between the kidney and the brain, where complex gangliosides are most abundant in adults [27]. Therefore, the differences in affinity toward b-series gangliosides are only one determinant of their usage as receptors *in vivo*, as a lower affinity can be balanced by a greater abundance in the host tissue. Moreover, gangliosides are differentially expressed in cortical tubular, medullary and glomerular tissues of adult human kidney and developmental changes in ganglioside expression have been observed in bovine kidney [26]. Thus, our observation of differing affinities toward b-series gangliosides raises the question whether developmental or drug-induced changes in

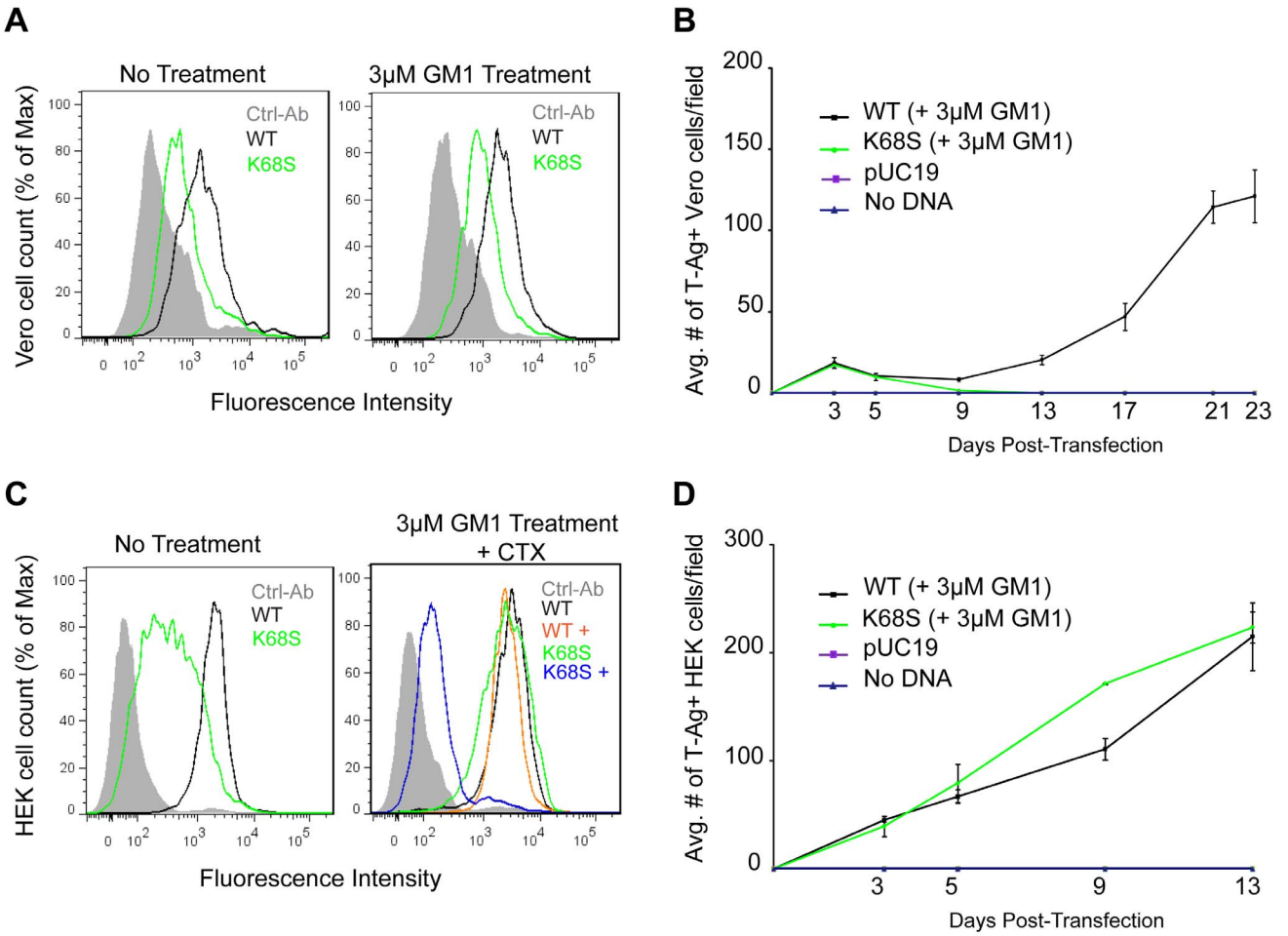


Figure 6. K68S BKPyV uses NeuAc-GM1 as a receptor for attachment and infection. (A) Binding of BKPyV K68S to Vero cells. Cells were treated as in 3C fixed and pentamer binding to cells untreated (left) or treated with NeuNAc-GM1 (right) assessed by flow cytometry. Histograms show the fluorescence intensity of the Alexa 488 antibody alone (gray-filled), WT pentamer (black) and K68S pentamer (green) for 1×10^4 events. (B) Growth assay for BKPyV K68S in Vero cells. Cells were transfected as previously described, treated with NeuNAc-GM1, fixed and stained over 23 days. Viral spread was quantified by scoring for cells expressing T-Ag. The anti-V antigen monoclonal antibody 597 used in Figures 1 and 3 recognizes an epitope that is disrupted by the K68S mutation, requiring the use of an mAb against T-Ag. The average number of T-Ag positive cells is plotted from 3 independent experiments. (C) Binding of BKPyV K68S to HEK cells. Cells were treated as previously described, fixed and pentamer binding to cells untreated (left), or treated with NeuNAc-GM1 and CTX (right) assessed by flow cytometry. Histograms show the fluorescence intensity of the Alexa 488 antibody alone (gray-filled), K68S pentamer (green) WT pentamer (black), WT pentamer with CTX (orange) and K68S pentamer with CTX (blue) for 1×10^4 events. (D) Growth assay for BKPyV K68S in HEK cells. Cells were transfected as previously described, treated with NeuNAc-GM1, fixed and stained over 13 days. Viral spread was quantified as above. Gangliosides were added every 3 days.
doi:10.1371/journal.ppat.1003688.g006

ganglioside distribution may play a role in BKPyV latency and reactivation.

Structural comparison of BKPyV-GD3 with the closely related SV40-GM1 complex suggested that a protruding lysine at position 68 may prevent BKPyV from binding GM1. To test this hypothesis, we introduced the smaller SV40 residue S68 into BKPyV VP1. This single mutation switched the oligosaccharide specificity of BKPyV from b-series gangliosides to GM1, altering a key attachment property of BKPyV. Known BKPyV VP1 sequences contain lysine, arginine or histidine at position 68. All of these amino acids would block engagement of GM1 but promote or at least tolerate binding of b-series gangliosides. BKPyV strains do not carry a serine at position 68, and this may indicate that recognition of GM1 instead of b-series gangliosides may not be advantageous to the virus in the context of the host organism. Possible explanations could be that GM1 is not very abundant or is differentially localized in human kidneys [25,26]. There likely exists an evolutionary constraint on BKPyV to bind b-series gangliosides, not GM1, especially as the remainder of the SV40 GM1 binding site is mostly conserved in BKPyV.

We have shown that unlike SV40, the BKPyV K68S mutant is specific for GM1 containing the human sialic acid NeuNAc, and cannot engage its simian counterpart NeuNGc due to steric hindrance. As the WT BKPyV and K68S binding sites for terminal sialic acid are identical, WT BKPyV shares the same sialic acid specificity. The inability of K68S to propagate in simian Vero cells and its ability to attach to and propagate in human HEK cells highlights the importance of sialic acid specificity for viral species tropism. Collectively, our data on BKPyV, JCPyV and SV40 suggest that each virus has adapted to the most prominent sialic acid in its host.

SV40, JCPyV and BKPyV all feature a conserved platform of core residues that allows them to efficiently engage terminal sialic acid in a similar manner. These core residues mediate the vast

majority of interactions. However, each virus achieves its distinct receptor specificity with a small number of strategically positioned satellite residues, such as K68 in the case of BKPyV, that form distinct contacts with additional carbohydrate moieties. Thus, these satellite residues define the context in which a terminal sialic acid can be bound, and as demonstrated here they present attractive opportunities for switching receptor specificities.

It is tempting to speculate that at least some members of the polyomavirus family have evolved from an initial sialic-acid binding template through subtle modification of their satellite residues, thereby expanding their host range and tropism. The switching of specificity can occur naturally in viruses, and often triggers altered pathogenicity and species tropism. In many cases, switching is due to exceedingly small changes in the virus capsid structure. Prominent examples include different serotypes of adenoviruses, the canine and feline parvoviruses, as well as avian, swine and human influenza viruses [28,29,30]. In many of these cases, however, the molecular functions of these switches, such as how specific mutations alter the interaction with receptors, are not well understood at the atomic level. Switching polyomavirus receptor specificities, as demonstrated here in a first example, may therefore be a useful tool to study parameters that define host receptor recognition, viral uptake, and entry pathways.

Materials and Methods

Virus infection

Cells (ATCC, Manassas, VA) were maintained at 37°C in Cellgro Minimum Essential Medium Eagle (MEM) supplemented with 5% heat inactivated fetal bovine serum (Atlanta Biologicals) and penicillin (10,000 U/ml) and streptomycin (10,000 µg/ml) (Gibco). Cells seeded in 24-well dishes were pre-incubated with media, dimethyl sulfoxide (DMSO) or gangliosides GM1, GD2, GD3, GD1b, and GT1b (Matreya) at 0.3–30 µM for 17 h at 37°C. Prior to infection cells were chilled for 20 min at 4°C and washed with 2% MEM. Cells were infected with 8×10^5 Fluorescent Forming Units (FFU) per ml of BKPyV for 1 h at 37°C. The infectious media was then removed and replaced with fresh growth media. Infection was scored 72 h post infection by staining for VP1 and analyzed by indirect immunofluorescence. Construction of the BKPyV pUC-19 expression plasmid was previously described [20]. BKPyV VP1 mutants were generated by site directed mutagenesis using QuickChange XL (Stratagene, La Jolla, CA). Mutant and WT plasmids were digested with BamHI (Promega). Vero or HEK cells were transfected with (0.5 µg) of linearized mutant or WT BKPyV plasmid DNA using Fugene 6 (Roche).

Indirect immunofluorescence

To detect expression of viral antigens cells were fixed with 2% paraformaldehyde in phosphate buffered saline (PBS) for 20 min at 25°C and permeabilized with 1% Triton-X 100 in PBS for 15 min at 37°C. Cells were incubated with the primary mouse monoclonal antibody PAb597 (1:10 [20], or PAb416 (Ab-2) (0.2 mg/ml) (Calbiochem), [31], used at 8 ng/µl to stain for BKPyV T-Ag. After incubation cells were washed with PBS and incubated with Alexa Fluor 488-labeled goat anti-mouse antibody in PBS (Invitrogen).

Flow cytometry

Vero or HEK cells were incubated with media or 3 µM GM1 for 17–18 h. Cells were washed and suspended in 100 µl (10 µg/mL) (Sigma) of CTX or PBS for 30 min on ice with 10 min agitation. Cells were washed and then incubated with 100 µl of purified wild type or mutant BKPyV VP1 pentamers (100 µg/mL) in PBS on ice for 2 h with 30 min agitations or with PBS alone. Cells were washed and suspended in 100 µl of Penta-His-AlexaFlour 488 conjugated antibody (10 µg/mL) (Qiagen) in PBS on ice for 1 h with 15 min agitations. Cells were washed and fixed in 1% paraformaldehyde and binding analyzed using a BD FACSCanto II flow cytometer (Benton, Dickinson, and Company). Data were analyzed using Flow Jo (Tree Star Inc.) software.

Recombinant protein expression and purification

We expressed and purified a truncated form of BKPyV VP1 that assembles into pentamers but does not form capsids. DNA coding for amino acids 30–300 of BKPyV VP1 was amplified by PCR and cloned into the pET15b expression vector (Novagen) in frame with an N-terminal hexahistidine tag (His-tag) and a thrombin cleavage site. The protein was overexpressed in *E. coli* BL21(DE3) and purified by nickel affinity chromatography and gel filtration on Superdex-200. For crystallization, the tag was cleaved with thrombin before gel filtration, leaving non-native amino acids GSHM at the N-terminus.

STD NMR measurements

All NMR spectra were recorded using 3 mm tubes on a Bruker DRX 500 MHz spectrometer fitted with a 5 mm cryogenic probe at 283 K and processed with TOPSPIN 2.0 (Bruker). For all proteins used for STD NMR, two NMR samples were prepared, containing either 1 mM GM1 oligosaccharide (Alexis) or 1 mM GD3 oligosaccharide (Sigma). Protein concentrations were between 19 µM and 22 µM. An additional sample contained 20 µM WT BKPyV VP1 and 1 mM GD1b oligosaccharide (Elicityl, F). Additional protein-free samples were prepared that only contained 1 mM GM1, GD3 or GD1b oligosaccharide. These samples were used to verify that no direct excitation of ligand resonances occurred during STD NMR measurements, and they served as samples for the spectral assignment. 0.1 mM trimethylsilyl propionate was then added to the GD3 sample to allow ^{1}H referencing. The buffer used for all NMR measurements contained 20 mM deutero-Tris pH 7.5, 150 mM NaCl, and 20 mM deutero-DTT. Samples were prepared in D_2O and no additional water suppression was used to avoid affecting the anomeric proton signals. The off- and on-resonance frequencies were set to 80 ppm and 7 ppm, respectively. The total relaxation delay was 4 s. A cascade of 40 Gaussian-shaped pulses with 50 ms duration each, corresponding to a strength of 65 Hz, and a saturation time of 2 s was used for selective excitation. A 10 ms continuous-wave spin lock filter with a strength of 3.7 kHz was employed in order to suppress residual protein signals. 32 k points were collected and zero filling to 64 k data points was employed. Spectra were multiplied with an exponential line broadening factor of 1 Hz prior to Fourier transformation.

To assign oligosaccharide proton resonances, series of 1D ^{1}H-TOCSY and COSY spectra as well as ^{1}H, ^{13}C-HSQC spectra were acquired. Literature values for related oligosaccharides served as additional assignment controls [32,33,34,35]. Assignment of the acetate methyl groups was taken from [34] for GD3 and [32] for GM1.

Crystallization and structure determination

For crystallization, BKPyV VP1 was supplemented with 20 mM DTT and concentrated to 6.6–7.0 mg/ml. The protein was crystallized at 20°C by sitting drop vapor diffusion against a reservoir of 16–18% PEG 3,350, 0.1 M HEPES pH 7.5 and 0.25 M LiCl (drop size 300 nl protein+300 nl reservoir). Crystals were

harvested into reservoir solution containing 14–16% PEG 3,350 and cryoprotected by soaking in harvesting solution supplemented with 30% (v/v) glycerol for 10 s before flash-freezing them in liquid nitrogen. For oligosaccharide complex formation, crystals were soaked in harvesting solution containing 20 mM GD3 oligosaccharide for 15 min before cryoprotection.

Diffraction data were collected at the SLS (Villigen, CH) and processed with xds [36], and the structure was solved by molecular replacement with Phaser [37] using the core of the SV40 VP1 pentamer (3BWQ) as the search model. After rigid body and simulated annealing refinement in Phenix [38], missing parts of the model were built in Coot [39]. Refinement proceeded by alternating rounds of refinement in Refmac5 [40] and model building in Coot. Fivefold non-crystallographic symmetry restraints were used throughout refinement. Oligosaccharide residues were located in weighted 2mFo-DFc and mFo-DFc electron density maps and refined with restraints from the CCP4 monomers library; only the α2,3- and α2,8- glycosidic linkages had to be user-defined. Data collection and refinement statistics are given in Table 1. Coordinates and structure factor amplitudes were deposited with the RCSB data bank (www.rcsb.org) with entry codes 4MJ0 (BKPyV VP1 bound to GD3) and 4MJ1 (unliganded BKPyV VP1). Structure figures were prepared with PyMol (Schrödinger Inc.).

Molecular modeling

To generate a model for the complex, we first explored the conformational space of GD1b alone using high-temperature molecular dynamics (MD) [41] and subsequently positioned the individual sampled snapshots into the binding site using the α2,8-disialic acid motif of crystal structure as an anchor point. Molecular dynamics simulation of GD1b was performed at 700 K for 100 ns using the MM3 force field as implemented in the TINKER software (http://dasher.wustl.edu/tinker/). Torsion restraints were applied on the ring torsions to avoid inversion of the carbohydrate rings during MD. Snapshots were recorded every 0.5 ps resulting in a conformational ensemble consisting of 200000 frames. Further processing of the data was performed using Conformational Analysis Tools (CAT) (http://www.md-simulations.de/CAT/). Conformational maps were calculated as described [41] in order to check that the accessible conformational space of the glycosidic linkages was sufficiently explored (Supplemental Fig. S1). All snapshots were positioned into the crystal structure using three atoms of residue NeuNAc 4R as an anchor and conformations that result in atom-overlaps with the protein were removed. Additionally two filters were applied on the remaining snapshots that control the position and orientation of residue NeuNAc 3R: Only snapshots were allowed to pass that have the center of the carboxylate and the *N*-acetyl group within 3 Å of the corresponding residue of the crystal structure. Several conformations were manually selected and refined based on 5 ns MD simulations at 300K in explicit water using YASARA [42]. Model coordinates are available from the authors upon request.

Carbohydrate microarray analyses

Microarrays comprised lipid-linked oligosaccharide probes, neoglycolipids (NGLs) and glycolipids, robotically printed on nitrocellulose-coated glass slides using a non-contact instrument [43,44]. For the analyses, an array set of 21 ganglioside-related probes (18 sialylated and 3 non-sialylated, in house designation Ganglioside Dose Response Array set 1) was used, in which each probe was arrayed at four levels: 0.3, 0.8, 1.7 and 5.0 fmol/spot. The microarray analyses were performed essentially as described [17]. In brief, microarrays were blocked in 5 mM HEPES (pH 7.4), 150 mM NaCl, 0.3% (v/v) Blocker Casein (Pierce), 0.3% (w/v) bovine serum albumin (Sigma) and 5 mM $CaCl_2$ (referred to as HBS-Casein/BSA). WT and mutant BKPyV VP1 were diluted in HBS-Casein/BSA and overlaid at 300 μg/ml and 150 μg/ml, respectively, followed by incubation with mouse monoclonal anti-poly-histidine and biotinylated antimouse IgG antibodies (both from Sigma). SV40 VP1 was tested as a protein-antibody complex that was prepared by preincubating with mouse monoclonal anti-poly-histidine and biotinylated anti-mouse IgG antibodies at a ratio of 4:2:1 (by weight) and diluted in HBS-Casein/BSA to provide a final SV40 VP1 concentration of 150 μg/ml. The SV40 and K68S protein samples had been supplemented with DTT to prevent dimerization of pentamers, while WT BKPyV VP1 was analysed without DTT. Binding was detected with Alexa Fluor-647-labelled streptavidin (Molecular Probes). Microarray data analyses were as described [45].

Supporting Information

Figure S1 Conformational energy maps for the GD1b oligosaccharide. The maps show the accessible conformational space of the glycosidic linkages and are calculated based on 200000 snapshots sampled from a 100 ns MD simulation at 700 K using TINKER/MM3 as described in (Frank et al., 2007). Three local energy minima are predicted for the internal NeuNAc-α2,3-Gal linkage.
(PDF)

Figure S2 Growth of BKPyV K68S in the absence of exogenous GM1. (A) Vero cells were transfected as previously described, treated with media lacking GM1 ganglioside addition, fixed and stained over 23 days. Viral spread was quantified by scoring for cells expressing T-Ag. (B) HEK cells were transfected as previously described, treated with media lacking GM1 ganglioside addition, fixed and stained over 13 days. Viral spread was quantified as above.
(TIF)

Table S1 Oligosaccharide probes. The table lists probes with their sequences included in the ganglioside 'dose- response' set 1.
(DOC)

Acknowledgments

We are grateful to members of our laboratories, especially Holger Hengel and Tina Wagner for help with protein expression and purification, and members of the Glycosciences Laboratory, especially Wengang Chai, Robert Childs and Mark Stoll, for their essential role in the neoglycolipid-based microarray system. We thank the SLS staff for assistance with data collection.

Author Contributions

Conceived and designed the experiments: UN SAA BSB YL MF ASP TF WJA TS. Performed the experiments: UN SAA BSB YL MF ASP LJS. Analyzed the data: UN SAA BSB YL MF ASP LJS TF TP WJA TS. Wrote the paper: UN SAA BSB YL MF ASP TF WJA TS.

References

1. Decaprio JA, Garcea RL (2013) A cornucopia of human polyomaviruses. Nat Rev Microbiol 11: 264–276.
2. Gardner SD, Field AM, Coleman DV, Hulme B (1971) New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet 1: 1253–1257.
3. Hirsch HH, Steiger J (2003) Polyomavirus BK. Lancet Infect Dis 3: 611–623.
4. Shinohara T, Matsuda M, Cheng SH, Marshall J, Fujita M, et al. (1993) BK virus infection of the human urinary tract. J Med Virol 41: 301–305.
5. Nickeleit V, Hirsch HH, Binet IF, Gudat F, Prince O, et al. (1999) Polyomavirus infection of renal allograft recipients: from latent infection to manifest disease. J Am Soc Nephrol 10: 1080–1089.
6. Bedi A, Miller CB, Hanson JL, Goodman S, Ambinder RF, et al. (1995) Association of BK virus with failure of prophylaxis against hemorrhagic cystitis following bone marrow transplantation. J Clin Oncol 13: 1103–1109.
7. Hirsch HH (2005) BK virus: opportunity makes a pathogen. Clin Infect Dis 41: 354–360.
8. Seganti L, Mastromarino P, Superti F, Sinibaldi L, Orsi N (1981) Receptors for BK virus on human erythrocytes. Acta Virol 25: 177–181.
9. 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.
10. Sinibaldi L, Goldoni P, Pietropaolo V, Longhi C, Orsi N (1990) Involvement of gangliosides in the interaction between BK virus and Vero cells. Arch Virol 113: 291–296.
11. 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.
12. 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.
13. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, et al. (1991) Structure of simian virus 40 at 3.8-A resolution. Nature 354: 278–284.
14. Stehle T, Yan Y, Benjamin TL, Harrison SC (1994) Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. Nature 369: 160–163.
15. 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.
16. 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.
17. Neu U, Maginnis MS, Palma AS, Stroh LJ, Nelson CD, 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.
18. 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 Pathog 8: e1002738.
19. Mayer M, Bernd M (1999) Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. Angew Chem Int Ed 38: 1784–1788.
20. Dugan AS, Gasparovic ML, Tsomaia N, Mierke DF, O'Hara BA, et al. (2007) Identification of amino acid residues in BK virus VP1 that are critical for viability and growth. J Virol 81: 11798–11808.
21. Muchmore EA, Diaz S, Varki A (1998) A structural difference between the cell surfaces of humans and the great apes. Am J Phys Anthropol 107: 187–198.
22. Campanero-Rhodes MA, Smith A, Chai W, Sonnino S, Mauri L, et al. (2007) N-glycolyl GM1 ganglioside as a receptor for simian virus 40. J Virol 81: 12846–12858.
23. Merritt EA, Sarfaty S, van den Akker F, L'Hoir C, Martial JA, et al. (1994) Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. Protein Sci 3: 166–175.
24. Qian M, Tsai B (2010) Lipids and proteins act in opposing manners to regulate polyomavirus infection. J Virol 84: 9840–9852.
25. Shayman JA, Radin NS (1991) Structure and function of renal glycosphingolipids. Am J Physiol 260: F291–302.
26. Holthofer H, Reivinen J, Miettinen A (1994) Nephron segment and cell-type specific expression of gangliosides in the developing and adult kidney. Kidney Int 45: 123–130.
27. Yu RK, Nakatani Y, Yanagisawa M (2009) The role of glycosphingolipid metabolism in the developing brain. J Lipid Res 50 Suppl: S440–445.
28. Persson BD, Muller S, Reiter DM, Schmitt BB, Marttila M, et al. (2009) An arginine switch in the species B adenovirus knob determines high-affinity engagement of cellular receptor CD46. J Virol 83: 673–686.
29. Hueffer K, Parrish CR (2003) Parvovirus host range, cell tropism and evolution. Curr Opin Microbiol 6: 392–398.
30. Gamblin SJ, Skehel JJ (2010) Influenza hemagglutinin and neuraminidase membrane glycoproteins. J Biol Chem 285: 28403–28409.
31. Eash S, Atwood WJ (2005) Involvement of cytoskeletal components in BK virus infectious entry. J Virol 79: 11734–11741.
32. Houliston RS, Jacobs BC, Tio-Gillen AP, Verschuuren JJ, Khieu NH, et al. (2009) STD-NMR used to elucidate the fine binding specificity of pathogenic anti-ganglioside antibodies directly in patient serum. Biochemistry 48: 220–222.
33. Brisson JR, Baumann H, Imberty A, Perez S, Jennings HJ (1992) Helical epitope of the group B meningococcal alpha(2–8)-linked sialic acid polysaccharide. Biochemistry 31: 4996–5004.
34. Michon F, Brisson JR, Jennings HJ (1987) Conformational differences between linear alpha (2–8)-linked homosialooligosaccharides and the epitope of the group B meningococcal polysaccharide. Biochemistry 26: 8399–8405.
35. Haselhorst T, Blanchard H, Frank M, Kraschnefski MJ, Kiefel MJ, et al. (2007) STD NMR spectroscopy and molecular modeling investigation of the binding of N-acetylneuraminic acid derivatives to rhesus rotavirus VP8* core. Glycobiology 17: 68–81.
36. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr 66: 133–144.
37. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760–763.
38. 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.
39. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132.
40. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240–255.
41. Frank M, Lutteke T, von der Lieth CW (2007) GlycoMapsDB: a database of the accessible conformational space of glycosidic linkages. Nucl Acid Res 35: 287–290.
42. Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. Proteins 57: 678–683.
43. Palma AS, Feizi T, Zhang Y, Stoll MS, Lawson AM, et al. (2006) Ligands for the beta-glucan receptor, Dectin-1, assigned using "designer" microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. J Biol Chem 281: 5771–5779.
44. Steinberg SF (2008) Structural basis of protein kinase C isoform function. Physiolog Rev 88: 1341–1378.
45. Stoll MS, Feizi T. Software tools for storing, processing and displaying carbohydrate microarray data. In: Kettner C, editor; 2009; Potsdam, Germany. Beilstein Institute for the Advancement of Chemical Sciences.

Progressive Multifocal Leukoencephalopathy-Associated Mutations in the JC Polyomavirus Capsid Disrupt Lactoseries Tetrasaccharide c Binding

Melissa S. Maginnis,[a] Luisa J. Ströh,[b] Gretchen V. Gee,[a] Bethany A. O'Hara,[a] Aaron Derdowski,[a] Thilo Stehle,[b,c] Walter J. Atwood[a]

Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, USA[a]; Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany[b]; Department of Pediatrics, Vanderbilt University, School of Medicine, Nashville, Tennessee, USA[c]

M.S.M. and L.J.S. contributed equally to this work.

ABSTRACT **The human JC polyomavirus (JCPyV) is the causative agent of the fatal, demyelinating disease progressive multifocal leukoencephalopathy (PML). The Mad-1 prototype strain of JCPyV uses the glycan lactoseries tetrasaccharide c (LSTc) and serotonin receptor 5-HT$_{2A}$ to attach to and enter into host cells, respectively. Specific residues in the viral capsid protein VP1 are responsible for direct interactions with the α2,6-linked sialic acid of LSTc. Viral isolates from individuals with PML often contain mutations in the sialic acid-binding pocket of VP1 that are hypothesized to arise from positive selection. We reconstituted these mutations in the Mad-1 strain of JCPyV and found that they were not capable of growth. The mutations were then introduced into recombinant VP1 and reconstituted as pentamers in order to conduct binding studies and structural analyses. VP1 pentamers carrying PML-associated mutations were not capable of binding to permissive cells. High-resolution structure determination revealed that these pentamers are well folded but no longer bind to LSTc due to steric clashes in the sialic acid-binding site. Reconstitution of the mutations into JCPyV pseudoviruses allowed us to directly quantify the infectivity of the mutants in several cell lines. The JCPyV pseudoviruses with PML-associated mutations were not infectious, nor were they able to engage sialic acid as measured by hemagglutination of human red blood cells. These results demonstrate that viruses from PML patients with single point mutations in VP1 disrupt binding to sialic acid motifs and render these viruses noninfectious.**

IMPORTANCE **Infection with human JC polyomavirus (JCPyV) is common and asymptomatic in healthy individuals, but during immunosuppression, JCPyV can spread from the kidney to the central nervous system (CNS) and cause a fatal, demyelinating disease, progressive multifocal leukoencephalopathy (PML). Individuals infected with HIV, those who have AIDS, or those receiving immunomodulatory therapies for autoimmune diseases are at serious risk for PML. Recent reports have demonstrated that viral isolates from PML patients often have distinct changes within the major capsid protein. Our structural-functional approach highlights that these mutations result in abolished engagement of the carbohydrate receptor motif LSTc that is necessary for infection. Viruses with PML-associated mutations are not infectious in glial cells, suggesting that they may play an alternative role in PML pathogenesis.**

Received 3 April 2013 **Accepted** 15 May 2013 **Published** 11 June 2013

Citation Maginnis MS, Ströh 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(3):e00247-13. doi:10.1128/mBio.00247-13.

Editor Michael Imperiale, University of Michigan

Address correspondence to Walter J. Atwood, Walter_Atwood@brown.edu, or Thilo Stehle, thilo.stehle@uni-tuebingen.de.

The human JC polyomavirus (JCPyV) is an icosahedral, nonenveloped double-stranded DNA (dsDNA) virus and a member of the *Polyomaviridae* family (1). JCPyV infects approximately 50% of the population, and the infection is asymptomatic in healthy individuals (2, 3). Viral spread likely occurs via a fecal-oral route, as JCPyV is shed in the urine of healthy individuals (4) and can be detected in untreated wastewater (5–7). The site of initial infection is thought to be the stromal cells of the tonsils (8), followed by a persistent infection in the kidney (9) and in B lymphocytes of the bone marrow (10–12). In healthy individuals, JCPyV remains in the kidney, but in immunosuppressed individuals, JCPyV can spread to the central nervous system (CNS) (10, 13–15) and infect astrocytes and oligodendrocytes (16, 17). Oligodendrocytes produce myelin, and astrocytes are critical to the process of myelination in the CNS (18–20). JCPyV infection of astrocytes and cytolytic destruction of the oligodendrocytes cause the fatal, demyelinating disease progressive multifocal leukoencephalopathy (PML) (21, 22). PML is a devastating disease that can result in fatality within 3 months to 1 year of symptom onset if untreated (23). PML affects approximately 3 to 5% of HIV-1-positive individuals, is considered an AIDS-defining illness, and is one of the most common CNS-related diseases in AIDS (22). Since 2005, the incidence of PML has risen in individuals receiving immunomodulatory therapies for autoimmune diseases (24). In particular, individuals with multiple sclerosis (MS) who are receiving the biological therapy natalizumab have a 1:500 chance of developing PML (25, 26). Natalizumab is an anti-VLA-4 ($\alpha4\beta1$ integrin) antibody that blocks extravasation of VLA-4$^+$ T and B lymphocytes

to the brain, where they normally bind to endothelial cells (27). Therefore, while this treatment prevents the movement of lymphocytes to the brain, thus protecting the brain of an MS patient from attack, the lack of immune surveillance can also result in increased spread of JCPyV to the brain and thus increase the chances of developing PML (25).

The mechanisms of JCPyV spread to the CNS and infection of glial cells are not well understood, although spread is thought to occur via a hematogenous route, possibly involving B lymphocytes (28–30). In addition, it is well documented that JCPyV undergoes certain polymorphic changes within the host that render it neurotropic. The nonpathogenic form of virus that resides in the kidney is referred to as the archetype strain (Cy) and can be detected in the urine of healthy individuals (4, 31–33). JCPyV undergoes rearrangements in the noncoding control region (NCCR), which contains the viral origin of replication and sequences that serve as binding sites for transcription factors necessary for transcription of viral early and late genes (34–38). These rearrangements include duplication of enhancer elements to convert the virus to the neuropathogenic form (34, 39, 40). Viruses found in the cerebral spinal fluid (CSF), brain tissue, and blood but not in the urine contain NCCR rearrangements and are referred to as PML-type strains (22). Mad-1 is the laboratory prototype strain of the PML-type strain that was originally isolated from the brain of a PML patient and contains a canonical 98-bp tandem repeat in the NCCR (35). While NCCR rearrangements are necessary for JCPyV growth in the CNS, the incidence of PML is relatively low, given the rates of seropositivity (22). Thus, it is likely that other viral, cellular, and individual host factors play a role in PML pathogenesis.

Recently, a number of studies have reported that viral isolates from PML patients also contain mutations in the viral capsid VP1 protein. VP1 is a pentameric protein that interacts with neighboring VP1 pentamers through C-terminal extensions linking together 72 pentamers to form the viral capsid. VP1 serves as the viral attachment protein and mediates direct interactions with cell surface receptors (41). Initial studies to define the receptors for JCPyV infection revealed that JCPyV utilizes sialic acid (42, 43) and the 5-HT_{2A} receptor (44). The sialic acid component of the receptor was thought to be an α2,3- or α2,6-linked sialic acid (45). Our laboratory demonstrated that the presence of α2,6-linked sialic acid correlates with JCPyV infection of cells and tissues in the host, including B lymphocytes, kidney, and the glial cells astrocytes and oligodendrocytes (46). Moreover, JCPyV infection of B lymphocytes, kidney, and glial cells is mediated by sialic acid (42, 45, 47, 48). We identified the specific sialic acid receptor motif for the Mad-1 strain of JCPyV as lactoseries tetrasaccharide c (LSTc), which terminates in α2,6-linked sialic acid (41). The high-resolution crystal structure of JCPyV VP1 pentamers in complex with LSTc revealed that the protein specifically binds to the terminal α2,6-linked sialic acid and engages LSTc in a unique L-shaped conformation. Mutation of VP1 residues that contact LSTc leads to a severe defect in viral growth in glial cells (41). Interestingly, viruses isolated from the blood and CSF of individuals with PML exhibit mutations in these VP1 residues, including L54F, S266F, and S268F/Y (49–53), which are located in the sialic acid-binding pocket of VP1 (41). However, these mutations have never been found in the urine of healthy individuals, nor are they present in the urine of individuals with PML, in whom viral isolates with PML-associated mutations are found in the CSF and blood. This suggests that these mutations can arise within the JCPyV-infected individual (50–53). Analysis of viral isolates from urine, blood, and CSF from a single patient infected with a single JCPyV genotype supports this hypothesis. While virus isolated from the blood and CSF of this individual carried PML-associated mutations in VP1 (L54F, N264S, and S266F), these mutations were not present in virus isolated from the urine (53). Mutations in VP1 sialic acid-binding sites arise with high frequency, in as many as 80 to 90% of the viral isolates, indicating that the majority of the isolates from individuals with PML exhibit mutations in one or more of these sites in VP1 (50, 52–55). The role of some of the PML-associated mutations has been analyzed as virus-like particles (VLPs) in the JCPyV genotype 3 background. VLPs with PML-associated mutations have an altered capacity to engage sialic acid-containing receptors, as VLPs with PML-associated mutations bind to glial cells in a neuraminidase-insensitive manner and bind to gangliosides as measured by enzyme-linked immunosorbent assay (ELISA) (50, 52). However, the role of these mutations in JCPyV infection and spread in glial cells has not been addressed.

One hypothesis is that the mutations in VP1 sialic acid-binding sites might render the virus more pathogenic in infected hosts by allowing it to spread more readily to the brain due to reduced nonspecific attachment to sialic acid pseudoreceptors. However, it is also possible that viruses with PML-associated mutations are defective particles produced due to high levels of viral replication or that they are immune escape mutants. Thus, we questioned whether these viruses represent the pathogenic form of the virus that infects glial cells in the CNS. To this end, we utilized a structure-function approach to define the role of these mutations in JCPyV attachment to cellular receptors and infection of glial cells by generating the PML-associated mutations in an infectious viral clone, in pseudoviruses, and in purified VP1 pentamers using the Mad-1 prototype PML strain as the backbone.

RESULTS

JC polyomaviruses with PML-associated mutations are not infectious. Mutations in JCPyV VP1 arise in individuals with PML at a very high frequency. In fact, 80 to 90% of viral isolates from individuals with PML exhibit mutations in one or more of these sites in VP1, and these mutations are never found in JCPyV isolates from individuals without PML. Furthermore, these mutations are found only in the blood and CSF, but not in the urine, of individuals with PML, indicating that the mutations arise within the host (50, 52–55). The most common PML-associated mutations are L54F and S268F, accounting for approximately 50% of PML-associated mutations, while the mutations S266F and S268Y are less frequent but still common (Table 1) (50, 52). Surprisingly, all four mutations target VP1 residues that form contacts with the terminal sialic acid of the specific receptor motif LSTc (Fig. 1A) (41). To define whether PML-associated VP1 mutations affected JCPyV infectivity, we introduced the most common mutations into the JCPyV infectious genomic clone with a Mad-1 VP1. Viral DNA was transfected into the permissive glial SVG-A cell line and analyzed for growth and spread over 22 days in culture. Infected cells were analyzed by indirect immunofluorescence using an antibody directed against the VP1 capsid protein. Throughout 22 days in culture, only the wild-type JCPyV with a Mad-1 VP1 was capable of growth and spread (Fig. 1B). On day 22, viral supernatants were harvested from cells (in Fig. 1B) and used to infect naive SVG-A cells. Wild-type JCPyV produced infectious viral

TABLE 1 Frequency of JCPyV VP1 PML-associated mutations

Residues[a]	% frequency of VP1 mutations in PML patients	Reference
L54F, K59N, D65H, N264T, S266F/L/T, S268F/Y/C	81	49
L54F, K59M/E/N, N264D/T, S266F/L, S268F/Y/C	52	50
L54F, K59E, D65H, N264D, S266F, S268F/Y, Q270H	90	52
L54F, S60P/T, D65H, N264S, S266F/L, S268F, Q270H	81	53

[a] Amino acid numbers are in accordance with the JCPyV Mad-1 strain (1), excluding the methionine at position 1.

particles that were released into the supernatant, while viruses with PML-associated mutations did not, indicating that the mutant viruses could not propagate in glial cells (Fig. 1C).

VP1 pentamers with PML-associated mutations exhibit abolished binding to glial cells. To determine whether the decreased levels of virus growth and infection in the PML-associated mutants correlated with binding to cells, the mutations were introduced into VP1 pentamers. The pentamers serve as a useful tool for studies of JCPyV attachment, entry, and trafficking, as they can bind to cells and traffic through the endosomal compartment to the endoplasmic reticulum with kinetics similar to that of virions (41, 56). Mutated VP1 pentamers were then purified and used to assess binding to SVG-A cells by flow cytometry (Fig. 2). JCPyV wild-type pentamers bind to SVG-A cells, while the VP1 pentamers with PML-associated mutations did not bind to cells, indicating that residues in VP1 that mediate sialic acid binding to the receptor motif LSTc are essential for VP1 engagement of SVG-A cells.

VP1 pentamers with PML-associated mutations are not capable of binding to LSTc. To assess the structural effects of these mutations, we solved crystal structures of three mutant VP1 pentamers: L54F, S268F, and S268Y. In all cases, the crystal lattices are identical to those of the wild-type JCPyV pentamers, with accessible binding sites for LSTc in two of the five VP1 monomers in a given pentamer. In order to analyze whether the mutants can still engage LSTc, crystals of wild-type Mad-1 VP1 pentamers and the L54F, S268F, and S268Y mutants were soaked in LSTc oligosaccharide solution according to the exact procedure used for the complex formation of wild-type VP1 (41), and structures were determined to high resolution (see Table S1 in the supplemental material). Crystal soaking experiments using 5 mM LSTc yielded the JCPyV Mad-1 VP1-LSTc complex structure with two occupied LSTc binding sites per VP1 pentamer (see Fig. S1B) but did not result in additional electron density for carbohydrate moieties in the case of any of the mutated VP1 pentamers. This demonstrates that engagement of LSTc is severely compromised in all three mutants.

We next investigated if any of the VP1 mutants are able to bind LSTc at an increased concentration of 20 mM ligand and an extended soaking time of 2 h. While the L54F and S268Y mutants still do not show any evidence for binding, weak interactions with LSTc can be observed in the S268F mutant. The simulated annealed omit electron density map for S268F clearly shows features of the Neu5Ac-α2,6-Gal-β1,4-GlcNAc portion of LSTc in two LSTc binding sites (see Fig. S1A in the supplemental material). One of these trisaccharides could be built. The structural rearrangements in VP1 upon recognition of LSTc are similar to those seen in the JCPyV Mad-1 VP1-LSTc complex, providing additional evidence for the presence of LSTc (41). N123 interacts via hydrogen bonding with the terminal Neu5Ac and GlcNAc *N*-acetyl group and is critical for the recognition of LSTc in the L-shaped conformation (Fig. 3A). The N123 side chain is rearranged in the S268F structure in order to accommodate the terminal Neu5Ac (Fig. 3B). Consequently, residues 64 to 68 of the clockwise BC2 loop move to prevent clashes with N123 in the new position. Density for these induced-fit movements as well as partially for the native conformation can be observed in the S268F VP1-LSTc complex structures. These findings are consistent with the occupancy of <1.0 of the carbohydrate ligand within the final complex structure. We conclude that the S268F mutant retains some affinity for LSTc, whereas the other two mutants do not.

VP1 pentamers with PML-associated mutations clash with LSTc. We next examined the effects of the three mutations L54F, S268F, and S268Y on the VP1 structure. Residues L54, S266, and S268 mediate direct contacts or are in close proximity to the terminal Neu5Ac within the LSTc binding site of the wild-type JCPyV VP1-LSTc complex (Protein Data Bank [PDB] accession code 3NXD) (41) (Fig. 3). In order to compare the effects of the mutations on the VP1 structure and its ligand binding properties, the L54F, S268F, and S268Y VP1 structures were superposed onto the JCPyV wild-type VP1-LSTc complex structure. In all three VP1 mutant structures, well-defined electron density could be observed for the respective mutated residue, confirming the presence of the mutation.

Analysis of the S268F VP1 structure shows that the Neu5Ac carboxyl group could still be recognized by S266 via a hydrogen bond, but the hydrogen bond with the side chain of residue 268 within the HI loop would be lost. The side chain of F268 and the α2,6-linked Gal are 3.7 Å apart. Although this is a close contact, it would still allow for engagement of LSTc without leading to severe clashes. Consistent with this, interaction with LSTc was observed with reduced affinity for S268F VP1 in the soaking experiment with 20 mM LSTc (Fig. 3B). Based on B-factor analysis of the bound LSTc and surrounding residues, we expect that the occupancy of LSTc is about 0.6 to 0.8 in one accessible binding site (chain C). The second binding site (chain B) has weaker electron density for LSTc, indicating lower occupancy.

In the case of the L54F VP1 mutant structure, the phenyl side chain of F54 points into the Neu5Ac binding pocket. With a distance of only 1.4 Å between the Neu5Ac moiety in the respective location and the F54 side chain, LSTc binding would lead to severe clashes (Fig. 3C). Finally, the S268Y VP1 structure shows that introduction of the larger tyrosine side chain at position 268 would lead to a close contact of 2.9 Å between the Y268 hydroxyl group and the α2,6-linked Gal. This steric interference likely blocks recognition of the LSTc motif (Fig. 3D).

The S266F mutation also occurs frequently in PML patients. We did not succeed in crystallizing the S266F VP1 pentamer, but modeling suggests that replacement of S266 with a bulky phenylalanine would abolish hydrogen bond formation with the Neu5Ac carboxyl group. In addition, all allowed phenylalanine side chain rotamer conformations would clash with the GlcNAc moiety of

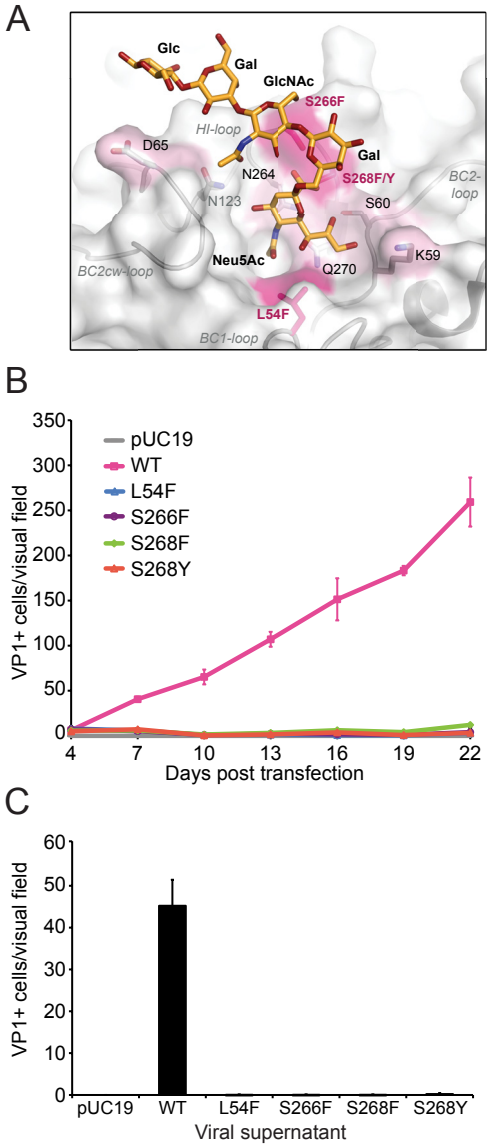


FIG 1 JCPyV with PML-associated mutations are not infectious. (A) PML-associated mutations are highlighted on the wild-type Mad-1 JCPyV VP1 pentamer (surface representation) in complex with LSTc (stick model) (41). Residues in pink indicate mutations used in this study, while those labeled in black represent other PML-associated mutations. (B) Growth of JCPyV VP1 wild-type (WT) and PML-associated mutant viruses. SVG-A cells were transfected with linearized DNA from wild-type and mutant JCPyV. Transfected cells were fixed and stained at day 4 posttransfection and then at 3-day intervals for 22 days by indirect immunofluorescence. Transfected or infected cells were quantified based on nuclear VP1 staining. Each data point represents the average number of infected cells per visual field for 6 fields of view for 3 independent experiments. Error bars indicate standard deviations. (C) Infectivity of supernatants from JCPyV VP1 wild-type and mutant viruses. SVG-A cells were inoculated with supernatants harvested from infected cells at day 22 from panel B. Cells were fixed and stained by indirect immunofluorescence at 72 h postinfection and quantified based on nuclear VP1 staining. The results are presented as the average number of infected cells per visual field for 6 visual fields from 3 individual samples performed in triplicate. Error bars indicate standard deviations.

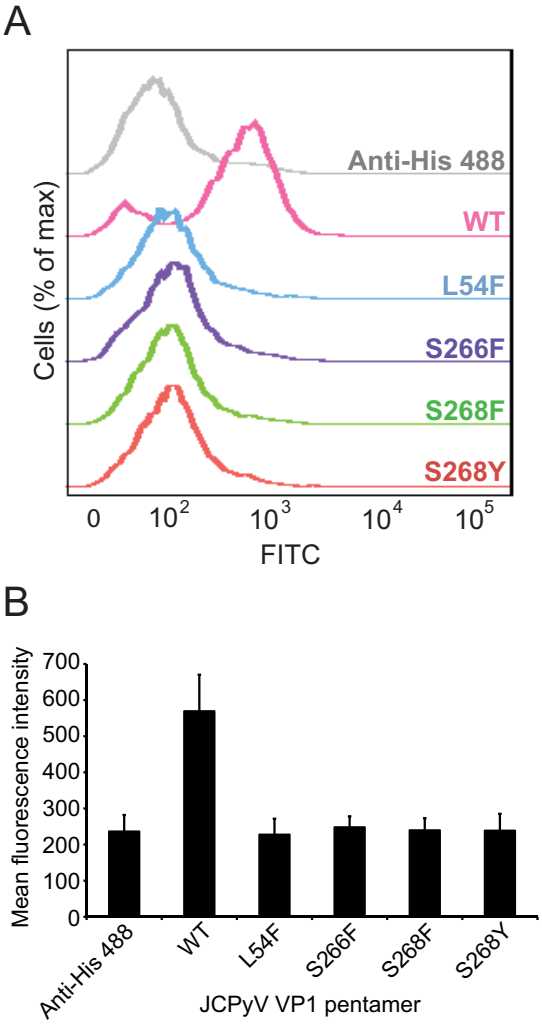


FIG 2 VP1 pentamers of JCPyV Mad-1 with PML-associated mutations exhibit reduced binding to cells. (A) SVG-A cells were incubated with 100 μg/ml of His-tagged wild-type (WT) or mutant pentamers in PBS, washed, and then incubated with a penta-His Alexa Fluor 488 antibody. Pentamer binding was analyzed by flow cytometry. Histograms represent the fluorescence intensity of Alexa 488 for antibody alone (gray) and pentamer samples for 10,000 gated events. (B) Quantitation of binding of VP1 pentamers with PML-associated mutations. Bar graph represents the mean fluorescence intensity of VP1 pentamers binding to SVG-A cells for 3 independent experiments. Error bars indicate standard deviations.

the long leg of the L-shaped LSTc motif. Therefore, binding of LSTc is likely also abolished for S266F VP1. We therefore conclude that the most frequently observed VP1 mutations in PML patients directly alter contacts of VP1 with LSTc, with various effects on binding ranging from substantially reduced interactions for the S268F mutant to clashes for the L54F and S268Y (and likely also the S266F) mutants.

JCPyV pseudoviruses with PML-associated mutations are not infectious. The structural analyses demonstrate that JCPyV pentamers expressing PML mutations are not capable of binding to the functional receptor motif α2,6-linked LSTc (Fig. 3), and the functional data suggest that JCPyV PML mutants cannot bind to or grow in SVG-A cells (Fig. 1 and 2). This led us to hypothesize that JCPyV PML mutants may have low levels of binding and growth that are not easily detected in our established culture model of viral infection. While it is not possible to propagate vi-

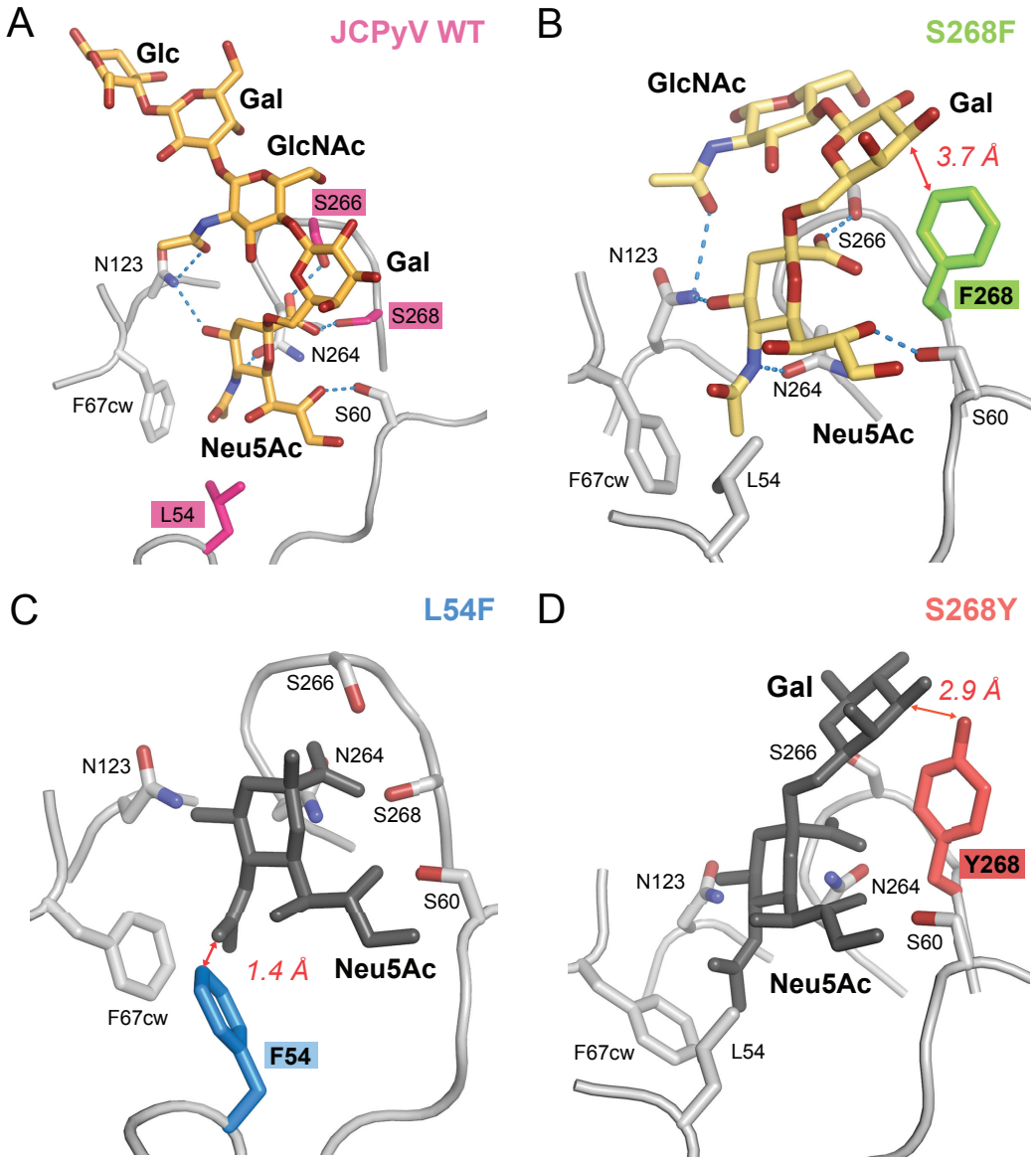


FIG 3 JCPyV VP1 proteins with PML-associated mutations have altered oligosaccharide-binding sites. (A) JCPyV wild-type (WT) VP1 pentamer-LSTc complex. (B) S268F VP1 pentamer-LSTc complex. (C) Unliganded L54F VP1 pentamer. (D) Unliganded S268Y VP1 pentamer. The JCPyV VP1 proteins are shown in cartoon representation. Important residues contributing to ligand binding and specificity are shown in stick representation. Side chains of residues at positions 54, 266, and 268 are highlighted in color. Key hydrogen bonds are shown as blue dashes, and red arrows indicate important distances. Carbohydrate residues are shown in stick representation and colored in orange or light orange (light orange indicates binding with reduced affinity) when present in the complex structures. Carbohydrate moieties depicted in gray are shown only for reference purposes and are obtained by superposing with the JCPyV WT VP1 pentamer-LSTc complex structure.

ruses with growth defects by standard methods, we have established a JCPyV pseudovirus system to help determine the infectivity of mutant viruses. The pseudoviruses are generated by transfecting the viral capsid proteins VP1, VP2, and VP3 into HEK293FT cells together with a reporter plasmid that expresses both green fluorescent protein (GFP) and a secreted form of luciferase (*Gaussia* luciferase). This assay is robust and provides a sensitive method of detecting the infectivity of mutant viruses that do not propagate under traditional culture methods. The PML-associated mutations were introduced into the Mad-1 VP1 and expressed in the pseudovirus system. Wild-type and mutant purified pseudoviruses were tested for infectivity in the human brain cell types SVG-A (glial), SVG-R (glial variant resistant to JCPyV infection), Poj19II (glial), and HFG-T (human fetal glial cells) or the kidney cell line HEK293FT. Infectivity was measured at 72 h postinfection by luciferase (Fig. 4). The JCPyV pseudoviruses with PML-associated mutations were not infectious in any of the cell types tested, including SVG-A cells, compared to the mock control and wild type (Mad-1) (Fig. 4). SVG-R cells are a variant of SVG cells that are resistant to JCPyV infection despite levels of virus binding equivalent to those of SVG-A cells (57). Thus, the SVG-R cells serve as a useful control to demonstrate the level of

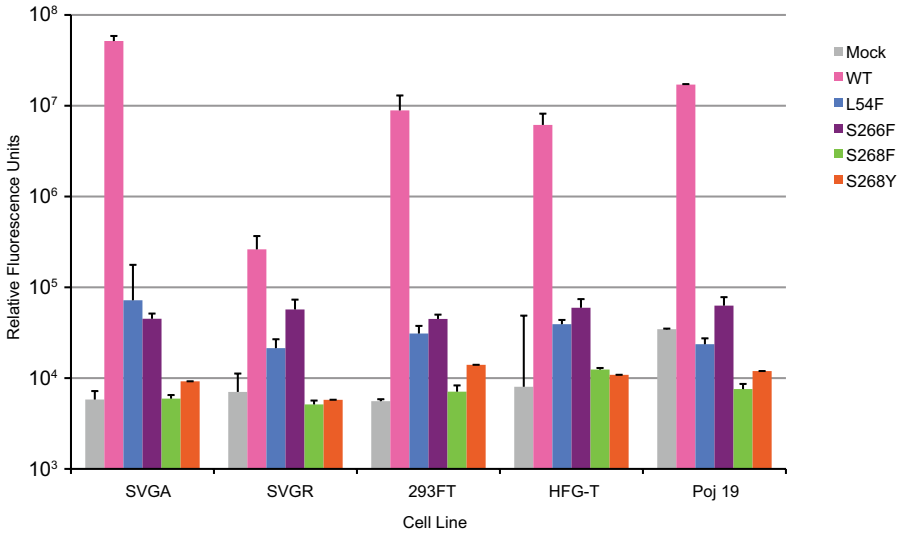


FIG 4 JCPyV pseudoviruses with PML-associated mutations are not infectious. The cell types shown were plated in a 96-well plate O/N. Medium was aspirated, cells were infected with 1×10^7 particles/ml of wild-type (Mad-1) or mutant pseudovirus in incomplete medium without phenol red at 37°C for 1 h, then complete phenol red-free medium was added, and cells were incubated at 37°C for 72 h. Supernatant from infected cells was collected and analyzed for secreted luciferase using a luminometer. The average relative luciferase units for quadruplicate samples are shown in log scale. Error bars represent standard deviations. These data are representative of 3 experiments performed in triplicate.

background luciferase readings in this assay. Furthermore, the cell lines used had variable levels of 5-HT_{2A}R expression (data not shown), indicating that the presence of 5-HT_{2A}R does not influence infection of viruses with PML-associated mutations. These data confirm that LSTc is a critical receptor motif for JCPyV infection and suggest that 5-HT_{2A}R is mediating a postattachment step in the virus life cycle. However, the expression of T antigen does influence infection by pseudoviruses, as the expression plasmid contains a simian virus 40 (SV40) origin of replication. Therefore, HEK293FT cells which have high levels of T antigen expression are readily infected by the pseudovirus, although kidney cells are generally less susceptible to JCPyV infection in the absence of T antigen in traditional culture models of JCPyV infection (48).

JCPyV pseudoviruses with PML-associated mutations do not hemagglutinate human RBCs. The viruses and pseudoviruses with PML-associated mutations were not infectious in a number of cell lines tested. Further, purified pentamers do not bind to SVG-A cells and have significantly reduced binding to LSTc as analyzed by X-ray crystallography. In order to determine if the PML-mutant pseudoviruses were capable of binding to other sialic acid receptors, a hemagglutination assay was performed. Wild-type JCPyV hemagglutinates human type O red blood cells (RBCs) (42). Wild-type and PML-associated mutant JCPyV pseudoviruses of equal pseudovirus particle concentrations were tested for the ability to agglutinate RBCs (Fig. 5). The PML mutant pseudoviruses were not capable of hemagglutination of red blood cells, while the wild-type JCPyV pseudovirus resulted in agglutination at a titer of 3.1×10^6 pseudovirus particles. These data suggest that the viruses isolated from individuals with PML which have mutations in VP1 that are critical for binding to LSTc no longer bind to sialic acid.

DISCUSSION

Mutations in the JCPyV capsid protein VP1 frequently arise in patients with PML. In this study, we examined the effect of three frequently occurring mutations on key properties of the virus, including growth, infectivity, capsid protein structure, and receptor binding. We find that all mutations severely compromise the interaction of the virus with its cognate receptor motif, the sialylated LSTc glycan, due to steric interference. As a result, the mutant viruses no longer hemagglutinate, and they no longer grow or infect glial cells. Taken together, these data suggest that although these viruses with PML-associated mutations are frequently found in individuals with PML (52, 53), they are likely not infectious in glial cells in the CNS due to reduced sialic acid binding.

Although viral isolates with VP1 mutations arise in individuals with PML, it has not been demonstrated whether viruses with PML-associated mutations are pathogenic in cells of the CNS. Individuals with PML who develop viruses with PML-associated mutations have higher levels of viral DNA in the CSF (52), yet it is unclear if these viruses are actually capable of increased spread to the brain or increased replication once they have reached the

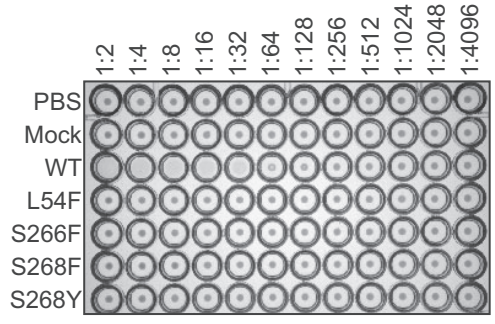


FIG 5 JCPyV pseudoviruses with PML-associated mutations do not bind to sialic acid. Wild-type (WT) and mutant pseudoviruses (1×10^7 particles) were added to U-bottomed 96-well plates containing equal volumes of PBS, and 2-fold serial dilutions were made across the plate. Human type O RBCs were added to each well and incubated at 4°C for 4 h. Data are representative of 2 independent experiments.

brain. Furthermore, these viruses have not been isolated from JCPyV-infected cells of the brains of individuals with PML but rather from the CSF. Given the avid interaction between JCPyV and cell surface sialic acid, it is possible that the JCPyV isolates with PML-associated mutations are found at a high frequency in the CSF due to their loss of sialic acid binding. The wild-type JCPyV, which retains the ability to bind to sialic acid, might remain cell associated, leading to lower levels of the wild-type virus in the CSF (50, 52). Also, it is possible that PML-associated mutations arise in the viral capsid as a mechanism of immune escape. JCPyV establishes a persistent infection in the kidney of healthy individuals with no significant complications, suggesting that the virus is under immune surveillance. To spread hematogenously to the CNS, the virus may undergo mutations in the viral capsid, making these mutant viruses antigenically distinct and capable of evading the host immune responses, resulting in increased spread. VP1 is the most abundant capsid protein of the virus, and increased antigenicity in the receptor-binding region may allow the pathogenic forms of the virus to spread to the CNS, while the immune system attacks the mutant viruses. Alternatively, mutations in VP1 may arise in the CNS of infected individuals as a mechanism of immune escape, resulting in increased spread and infection of glial cells. Thus, it remains to be demonstrated if the viruses with PML-associated mutations are pathogenic in the brain. Our data demonstrate that JCPyV with PML-associated mutations are not infectious in a panel of glial cell lines.

However, we also cannot exclude the hypothesis that JCPyVs with PML-associated mutations are infectious in the brains of individuals with PML. It is plausible that the loss of sialic acid binding through mutation of specific residues in the viral capsid may be necessary for increased spread to the brain and subsequent infection of glial cells in a sialic acid-independent manner (58, 59). It is reasonable to speculate that reduced binding to sialic acid receptors in the periphery leads to increased spread of JCPyV to the brain, as JCPyV binding to both red blood cells and B lymphocytes is mediated via interactions with sialic acid (Fig. 5) (47). This hypothesis is similar to the scenario demonstrated for mouse polyomavirus (MPyV) in which a mutation in the VP1 sialic acid-binding pocket in a position orthologous to position S268 of JCPyV leads to decreased MPyV binding to RBCs and increases viral dissemination and pathogenicity of the virus, resulting in a lethal outcome (58–61). JCPyV with PML-associated mutations might then engage a non-sialic acid receptor on glial cells in the CNS. It is also possible that the host cell factor necessary for infection by JCPyV with PML-associated mutations is not expressed abundantly on the cell lines that we tested. The crystal structure of JCPyV VP1 revealed a groove on the surface of the pentameric capsid protein that is unique among all polyomaviruses crystallized to date (41, 62, 63). This groove could well be involved in the engagement of JCPyV with other receptor structures on the surface of host cells, and such an interaction might be favored if binding to LSTc is significantly reduced or blocked. There are many differences between the *in vitro* infection model and a natural *in vivo* infection that we cannot fully account for in our current tissue culture model of JCPyV infection. JCPyV infects approximately 50% of the population but rarely causes PML and does so only in those who are immunocompromised by HIV infection, AIDS, or immunomodulatory therapies. Therefore, the development of PML must be multifactorial and involve host-specific factors that are influenced in the disease states of HIV infection or immune-mediated diseases such as MS. For instance, the breakdown of the blood-brain barrier and/or use of antiretroviral therapies or natalizumab therapy may cause the upregulation of alternate receptors or host cell factors that allow JCPyV to efficiently attach and enter into glial cells in the CNS. Thus, the tissue culture model of infection may not accurately demonstrate the disease pathogenesis *in vivo*.

Gorelik and colleagues previously demonstrated that VLPs with PML-associated mutations in the genetic background of the JCPyV genotype 3 have reduced binding to kidney cells, red blood cells, and lymphocytic cells but retain the ability to bind to the glial cells SVG-A cells and astrocytes. Additionally, they demonstrated that introduction of these mutations into a VLP leads to altered receptor usage, consistent with an enhanced affinity for gangliosides, including GM1, GM2, and GD3. However, the JCPyV type 3 VLP binds to asialo-GM1, GD1a, GD1b, GD2, and GT1b (52) and has been reported to bind to GM1 and GM2 (Leonid Gorelik, Consortium for Functional Glycomics [CFG], available online at http://www.functionalglycomics.org, according to CFG policy). This affinity of JCPyV type 3 for gangliosides may have influenced the results of the PML-associated mutations in the type 3 VLPs. Additionally, VLPs of JCPyV have been demonstrated to bind to a panel of gangliosides in viral overlay assays, suggesting that VLPs may nonspecifically engage sialic acid-containing gangliosides, particularly those with α2,6-linkages (64). There are 7 known serotypes of JCPyV and 13 distinct geographic subtypes that have been determined (65). JCPyV isolates from genotypes 1 and 2 are more frequently isolated from individuals than are other genotypes, including type 3 (53, 54). In fact, type 3 strains are not commonly isolated from the CSF of PML patients (50). We analyzed the effect of PML-associated mutations in the Mad-1 prototype strain, the prototype PML-type strain for genotype 1, for which the receptors have been identified (41, 44). Additionally, the amino acid sequences of VP1 of genotypes 1 and 3 differ in up to eight residues (depending on the strain), and some of the mutations are located close to the LSTc binding site. We previously demonstrated that Mad-1 VP1 pentamers do not bind to any ganglioside structures in a glycan array screen (41), and we have not been able to demonstrate that expression of gangliosides influences infection of JCPyV Mad-1 in glial cells (Atwood laboratory, unpublished results). Thus, strain-specific differences in receptor engagement could have influenced the results of the study published by Gorelik et al. (52) when the JCPyV type 3 strain was utilized to study the PML-associated mutations.

Structural analysis of Mad-1 and mutant VP1 pentamers reveals that steric hindrance introduced by PML-associated mutations such as L54F or S268Y abolishes engagement of LSTc, and these mutations would also block binding to other sialylated glycan motifs, independent of their location on gangliosides or glycoproteins or their linkage. Although some mutations, such as S268F, lead to only minor clashes, these mutations would nevertheless drastically reduce the affinity of VP1 for LSTc due to reduced contacts and some steric interference. Consequently, such interactions would likely not be functional for LSTc or any other sialylated glycan whether they are in the genotype 1 or genotype 3 background.

Genotype 3 strains differ from Mad-1 in up to eight amino acid positions. Five of these positions are included in our VP1 pentamer construct (R74, S116, G133, L157, and K163). Residues S116 and G133 are located near the LSTc binding pocket but do

not make direct contacts with LSTc, while R74, L157, and K163 are found outside the binding pocket. Additionally, V320, E331, and K344 are found at the C terminus of the full-length VP1 protein and far from the glycan receptor-binding site. Nevertheless, amino acid differences in Mad-1 and genotype 3 that are near the LSTc binding pocket would not change the drastic impact of the mutation L54F or S268F/Y and likely S266F on the interaction with sialylated glycan motifs, as these mutations are located directly in the binding pocket.

Interestingly, engagement of LSTc is likely preserved, perhaps with only moderate decreases in affinity, for less frequently occurring mutations. Reported PML-associated mutations K59M/E/N, S60T, and D65H, which were not addressed in our study, are unlikely to lead to structural changes and would still allow similar contacts with LSTc. In the case of Q270H and N264D/S/T, the hydrogen-bonding network in the Neu5Ac binding pocket would be altered and, thus, these mutations would likely result in an overall reduced affinity for sialylated motifs.

These findings highlight the importance of LSTc as a functional receptor motif. Additionally, we demonstrated that viral isolates with PML-associated mutations have reduced binding to LSTc and other sialic acid motifs, and this reduced binding renders these viruses noninfectious in both kidney and glial cells. Taken together, this work illustrates that engagement of cell surface receptors is an important determinant of tissue tropism and viral pathogenesis for JCPyV infection.

MATERIALS AND METHODS

Cells, viruses, and antibodies. SVG-A cells are a subclone of the human glial cell line SVG transformed with an origin-defective SV40 mutant (66) and were grown in minimum essential medium (MEM) supplemented to contain 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (P/S) (Mediatech, Inc.) in a humidified incubator at 37°C. SVG-R cells are resistant to JCPyV infection (57) and were grown in MEM supplemented to contain 10% FBS. HEK293FT cells are derived from human embryonal kidney cells transformed with the SV40 large T antigen (Invitrogen Life Technologies) and were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented to contain 10% FBS, 0.1 mM nonessential amino acids (NEAA), 6 mM L-glutamine, 1 mM sodium pyruvate, and 500 mg/ml Geneticin. HFG-T cells are human fetal glial cells transformed with an origin-defective SV40 T antigen (67) grown in modified Eagle's medium supplemented to contain 10% FBS. Poj19II cells are human fetal glial cells transformed with a replication-defective JCPyV (68) grown in DMEM supplemented to contain 10% FBS. Generation and propagation of the virus strain Mad-1/SVEΔ were previously described (69). JCPyV infection in SVG-A cells was assessed using PAB597, a hybridoma supernatant that produces a monoclonal antibody against JCPyV VP1 (70) and was generously provided by Ed Harlow. Penta-His Alexa Fluor 488 was used at 20 μg/ml (Qiagen).

Viral growth assay. VP1 mutations were generated in the genomic JCPyV DNA of strain JC12 with a Mad-1 VP1 (68) and subcloned into pUC19 (43). Mutations were introduced by site-directed mutagenesis using an Agilent QuikChange II site-directed mutagenesis kit (Qiagen) according to the manufacturer's instructions. Sequencing was performed at Genewiz Inc. Primers used for mutagenesis were as follows (with mismatched nucleotides shown in boldface), 5′→3′: L54F, GGTGACCCAG ATGAGCA**TTT**TAGGGGTTTTAGTAAGT and ACTTACTA**AAA**CCC CTAAAATGCTCATCTGGGTCACC; S266F, TGTGGCATGTTTACAA ACAGG**TTT**GGTTCCCAGCAG and CTGCTGGGAACCAAACCTGTT TGT**AAA**CATGCCACA; S268F, TTTACAAACAGGTCTGGT**TTC**CAG CAGTGGAGAGG and CCTCTCCACTGCTG**GAA**ACCAGACCTGTTT GTAAA; S268Y, GTTTACAAACAGGTCTGGTT**ATC**AGCAGTGGAGA GGACTC and GAGTCCTCTCCACTGCT**GAT**AACCAGACCTGTTTG TAAAC.

Ten micrograms of purified plasmid DNA was digested with BamHI (Promega) at 37°C for 2 h to separate the JCPyV genomic DNA from the pUC19 backbone plasmid. Digests were performed in triplicate for each sample and verified by agarose gel electrophoresis. SVG-A cells were plated to 40% confluence in 24-well plates (Corning). Cells in medium without antibiotics were transfected with 1 μg of digested DNA using Fugene (Promega) at a 3:2 ratio (Fugene to DNA). Transfected cells were incubated at 37°C overnight (O/N), and medium containing 5% FBS and 2% P/S was added to cells the next day. Cells were incubated at 37°C and fed with 500 ml of medium containing 5% FBS, 1% P/S, and 1% amphotericin B (Mediatech) or fixed for immunofluorescence staining at day 4 and at 3-day intervals thereafter for 22 days. For the infectivity assay, supernatants were collected from the samples of the growth assay at 22 days posttransfection. SVG-A cells at 70% confluence in 24-well plates (Corning) were infected with 150 μl of virus supernatant at 37°C for 1 h; then 1 ml of medium containing 5% FBS, 1% P/S, and 1% amphotericin B was added to cells; and cells were incubated at 37°C for 72 h. Cells were fixed and stained by indirect immunofluorescence.

Indirect immunofluorescence. Cells were washed in phosphate-buffered saline (PBS), fixed in cold methanol (MeOH), and incubated at −20°C. Fixed cells were washed in PBS, permeabilized with 0.5% Triton X-100 (TX-100; USB Corporation) at room temperature (RT) for 5 min, incubated with VP1-specific antibody PAB597 (1:10) in PBS at 37°C for 1 h, washed with PBS, incubated with a goat anti-mouse Alexa Fluor 488 (1:1,000)-conjugated antibody in PBS at 37°C for 1 h, and then washed with PBS. Cells were analyzed for nuclear VP1 staining under a 20× objective using an Eclipse TE2000-U microscope (Nikon).

JCPyV pseudovirus production. Codon optimization of the JCPyV VP1, VP2, and VP3 genes was performed according to the National Cancer Institute Center for Cancer Research Lab of Cellular Oncology Technical Files (http://home.ccr.cancer.gov/LCO/production.asp) to achieve optimal expression in the human-derived cell line 293FT. Genes were synthesized by Blue Heron Biotech, LLC. The VP1 gene was subcloned into the pwP vector in place of the murine polyomavirus (MPyV) VP1 gene. The JCPyV VP2 and VP3 genes were subcloned into the ph2p vectors in place of the MPyV genes (71). The luciferase reporter vector phGluc (72) expresses a secreted form of *Gaussia* luciferase under the control of the EF1α promoter. All plasmids were obtained from AddGene. Site-specific mutations were made using the Agilent QuikChange II site-directed mutagenesis kit prior to subcloning. Primers for generating PML-associated mutations in VP1 (5′→3′) were as follows, with mismatched nucleotides shown in boldface: L54F, TGGGCGACCCCGATG AACAT**TTT**CGCGGATTC and GAATCCGCG**AAA**ATGTTCATCGGG GTCGCCCA; S266F, GGCATGTTCACAAATCGC**TTT**GGCTCACAGC AGTGGAGA and CTCCACTGCTGTGAGCC**AAA**GCGATTTGTGAAC ATGCC; S268F, CACAAATCGCAGTGGC**TTT**CAGCAGTGGAGGG GATT and AATCCCCTCCACTGCTG**AAA**GCCACTGCGATTTGTG; S268Y, CACAAATCGCAGTGGC**TAC**CAGCAGTGGAGGGGATT and AATCCCCTCCACTGCTG**GTA**GCCACTGCGATTTGTG.

Pseudoviruses were produced by transfection of the VP1, VP2, VP3, and phGluc plasmids into 293FT cells using Fugene 6 transfection reagent (Promega) in a 5:1:1:1 ratio. Mock pseudovirus controls were generated by transfecting HEK293FT cells with control plasmid and the phGluc reporter plasmid in a 7:1 ratio. Cells were harvested 48 h posttransfection by scraping and then pelleted and resuspended in buffer A (10 mM Tris, pH 8, 50 mM NaCl, 0.1 mM $CaCl_2$, 0.01% TX-100) with EDTA-free protease inhibitors (Roche Applied Science). Cells were lysed by three rounds of freezing and thawing, sonicated, and treated with 0.25% deoxycholic acid at 37°C for 30 min. The pH was lowered to 6.0, and the lysates were treated with type V neuraminidase (Sigma) at 37°C for 1 h. The pH was then raised to 7.5, $CaCl_2$ was added, and the lysate was treated with DNase I (New England Biolabs). Pseudoviruses were then purified through an iodixanol gradient by centrifugation at 234,000 × *g* in an

SW55 Ti rotor (Beckman) at 16°C for 3.5 h. The band containing pseudovirus or corresponding control was extracted by syringe.

To determine the titers of pseudoviruses for properly encapsidated genomes, they were again treated with DNase I and protected phGluc was extracted using the DNeasy blood and tissue kit (Qiagen). The titer of the packaged genome was determined using absolute quantification and TaqMan quantitative PCR (qPCR; Applied Biosystems) to create a standard curve using serial dilutions of phGluc. The number of copies for the known plasmid was plotted in a scatter plot against the threshold cycle (C_T) value determined for each dilution. A best-fit line was generated, and the trend line equation from regression analysis was used to calculate the relationship between the C_T value of the unknown input template and copies of the packaged pseudovirus plasmid. We determined the volume of pseudovirus particles based on the encapsidated genomes to equal 10^7 viral particles/ml and diluted each sample to use equivalent amounts of pseudovirus to perform experiments.

Pseudovirus luciferase infectivity assay. Cells were plated in a 96-well plate to 70% confluency and infected with equal particle equivalents (1×10^7 particles/ml) of wild-type and mutant pseudovirus in incomplete medium without phenol red, adjusted for equal volume equivalents with the purification reagent Optiprep (33%). Infected cells were incubated at 37°C for 1 h and washed with PBS, complete medium without phenol red was added, and cells were incubated at 37°C for 72 h. Secreted luciferase was quantitated in 20 to 50 μl of cellular supernatants using the BioLux *Gaussia* luciferase assay (New England Biolabs) according to the manufacturer's instructions using an opaque 96-well microplate in a GloMax Multi-Detection System luminometer (Promega) equipped with an autoinjector. Numbers of infected cells were also measured by quantifying GFP-positive cells by fluorescence microscopy using an Eclipse TE2000-U microscope (Nikon).

Generation of pentamers. cDNA coding for amino acids 22 to 289 of the Mad-1 strain of JCPyV VP1 (UniProtKB entry code P03089) was cloned into the pET15b expression vector (EMD Millipore) in frame with an N-terminal hexahistidine tag (His tag) and a thrombin cleavage site as described previously (41). Mutations were introduced by site-directed mutagenesis with the Agilent QuikChange mutagenesis kit according to the manufacturer's instructions using primers listed for generation of mutant viruses for viral growth assay. Sequences were verified by Genewiz. VP1 pentamers were purified as described previously (56). Protein concentrations were determined using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific), and pentamers were concentrated to ~1 mg/ml.

Crystallization. JCPyV VP1 pentamers carrying mutation L54F, S268F, or S268Y were concentrated to 4.5 mg/ml in 20 mM HEPES (pH 7.5), 150 mM NaCl and crystallized at 20°C using the sitting drop vapor diffusion technique and 100 mM HEPES (pH 7.5), 200 mM KSCN, 12% (wt/vol) polyethylene glycol 3350 (PEG 3350) as the reservoir solution. Drops were set up by mixing the protein solution 1:1 with the reservoir solution and cross-seeded by adding 0.2 μl of a microseeding solution obtained from previous JCPyV VP1 crystals. In order to test for interaction with LSTc oligosaccharide, crystals were soaked in the reservoir solution complemented with 5 mM LSTc (Dextra, United Kingdom) for 3 min or with 20 mM LSTc for 2 h. Subsequently, crystals were transferred for 2 s into a harvesting solution supplemented with 30% (vol/vol) glycerol and 5 mM or 20 mM LSTc, respectively, and then flash-frozen in liquid nitrogen.

Data collection and structure determination. Data sets were collected at beamline X06DA at SLS (Villigen, Switzerland). Diffraction data were processed with XDS (73), and structures were solved by molecular replacement with Phaser in CCP4 (74, 75). The native JCPyV Mad-1 VP1 structure (PDB accession number 3NXG) lacking solvent molecules was used as the search model. Rigid body and simulated annealing refinement was carried out with Phenix (76) in order to remove model bias. Crystals of L54F, S268F, and S268Y mutant pentamers have the same space group, and similar unit cell parameters, as do wild-type JCPyV Mad-1 VP1 pentamers, with a single VP1 pentamer forming the asymmetric unit in each case. Mutations were introduced into the respective models, and water molecules were added (excluding the LSTc binding site) using Coot (77). Alternating rounds of model building in Coot and restrained refinement, including 5-fold NCS restraints and the translation-libration-screw (TLS) method (78), were performed with Refmac5 (79). In order to test for the presence of LSTc, simulated annealing F_{obs}-F_{calc} electron density maps in the putative LSTc binding site of the mutated VP1 pentamers soaked with LSTc were compared to the respective omit map of the Mad-1 VP1-LSTc complex. In the case of S268F VP1, which showed some electron density for LSTc, the ligand was built and refined using the CCP4 suite library and user-defined restraints for the α2,6-glycosidic bond. Structure figures were prepared with PyMOL (PyMOL Molecular Graphics System, version 1.3; Schrödinger, LLC).

Flow cytometry. Purified wild-type and mutant JCPyV VP1 pentamers (100 μg/ml) in a 100-μl total volume of PBS were incubated with SVG-A cells in suspension on ice for 2 h with occasional agitation. Cells were washed, pelleted by centrifugation, and suspended in 75 μl of penta-His Alexa Fluor 488 antibody (Qiagen) (20 μg/ml) in PBS on ice for 1 h. Cells were washed, pelleted, resuspended in PBS, and analyzed for pentamer binding using a BD FACSCalibur (Becton, Dickinson and Company) flow cytometer equipped with a 488-nm excitation line. Data were analyzed using BD CellQuestPro (Becton, Dickinson and Company) and FlowJo software (Tree Star, Inc.).

Hemagglutination assay with JC pseudoviruses. Serial 2-fold dilutions of 1×10^7 particles of wild-type or mutant JC pseudoviruses in a volume of 99 μl were prepared in PBS in a U-bottomed 96-well plate (Corning) (dilution range, 1:2 to 1:4,096). A volume equivalent (99 μl) of 1× PBS was used as a negative control. Human type O negative red blood cells (RBCs) that had been washed with Alsever's buffer (0.1 M D-glucose, 0.027 M sodium citrate, 0.07 M NaCl, pH 6.5) were added to each well at a volume equivalent of 99 μl, and plates were gently agitated. Plates were incubated at 4°C for 4 h and imaged using a ChemiDoc XRS imager (Bio-Rad).

Protein structure accession numbers. Coordinates and structure factor amplitudes have been deposited with the RCSB Protein Data Bank (http://www.pdb.org) under accession codes 4JCE (L54F), 4JCF (S268F), and 4JCD (S268Y) of the structures obtained after crystal soaking in harvesting solution supplemented with 20 mM LSTc.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00247-13/-/DCSupplemental.

Table S1, DOCX file, 0.1 MB.

Figure S1, EPS file, 13 MB.

ACKNOWLEDGMENTS

We thank members of the Atwood and Stehle laboratories for critical discussion and review of the manuscript.

Work in the Atwood and Stehle laboratories was supported by PPG 5P01NS065719 (W.J.A. and T.S.) and a Ruth L. Kirschstein National Research Service Award F32NS064870 (M.S.M.) from the National Institute of Neurological Disorders and Stroke.

REFERENCES

1. **Frisque RJ, Bream GL, Cannella MT.** 1984. Human polyomavirus JC virus genome. J. Virol. **51:**458–469.
2. **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.
3. **Egli A, Infanti L, Dumoulin A, Buser A, Samaridis J, Stebler C, Gosert R, Hirsch HH.** 2009. Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. J. Infect. Dis. **199:**837–846.
4. **Yogo Y, Kitamura T, Sugimoto C, Ueki T, Aso Y, Hara K, Taguchi F.** 1990. Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J. Virol. **64:**3139–3143.
5. **McQuaig SM, Scott TM, Lukasik JO, Paul JH, Harwood VJ.** 2009.

Quantification of human polyomaviruses JC virus and BK virus by TaqMan quantitative PCR and comparison to other water quality indicators in water and fecal samples. Appl. Environ. Microbiol. **75:**3379–3388.
6. **Hamza IA, Jurzik L, Stang A, Sure K, Uberla K, Wilhelm M.** 2009. Detection of human viruses in rivers of a densely-populated area in Germany using a virus adsorption elution method optimized for PCR analyses. Water Res. **43:**2657–2668.
7. **Ahmed W, Wan C, Goonetilleke A, Gardner T.** 2010. Evaluating sewage-associated JCV and BKV polyomaviruses for sourcing human fecal pollution in a coastal river in Southeast Queensland, Australia. J. Environ. Qual. **39:**1743–1750.
8. **Monaco MC, Jensen PN, Hou J, Durham LC, Major EO.** 1998. Detection of JC virus DNA in human tonsil tissue: evidence for site of initial viral infection. J. Virol. **72:**9918–9923.
9. **Dörries K.** 1998. Molecular biology and pathogenesis of human polyomavirus infections. Dev. Biol. Stand. **94:**71–79.
10. **Houff SA, Major EO, Katz DA, Kufta CV, Sever JL, Pittaluga S, Roberts JR, Gitt J, Saini N, Lux W.** 1988. Involvement of JC virus-infected mononuclear cells from the bone marrow and spleen in the pathogenesis of progressive multifocal leukoencephalopathy. N. Engl J. Med. **318:**301–305.
11. **Monaco MC, Atwood WJ, Gravell M, Tornatore CS, Major EO.** 1996. JC virus infection of hematopoietic progenitor cells, primary B lymphocytes, and tonsillar stromal cells: implications for viral latency. J. Virol. **70:**7004–7012.
12. **Tan CS, Dezube BJ, Bhargava P, Autissier P, Wüthrich C, Miller J, Koralnik IJ.** 2009. Detection of JC virus DNA and proteins in the bone marrow of HIV-positive and HIV-negative patients: implications for viral latency and neurotropic transformation. J. Infect. Dis. **199:**881–888.
13. **Dörries K, Vogel E, Günther S, Czub S.** 1994. Infection of human polyomaviruses JC and BK in peripheral blood leukocytes from immunocompetent individuals. Virology **198:**59–70.
14. **Dubois V, Lafon ME, Ragnaud JM, Pellegrin JL, Damasio F, Baudouin C, Michaud V, Fleury HJ.** 1996. Detection of JC virus DNA in the peripheral blood leukocytes of HIV-infected patients. AIDS **10:**353–358.
15. **Dubois V, Dutronc H, Lafon ME, Poinsot V, Pellegrin JL, Ragnaud JM, Ferrer AM, Fleury HJ.** 1997. Latency and reactivation of JC virus in peripheral blood of human immunodeficiency virus type 1-infected patients. J. Clin. Microbiol. **35:**2288–2292.
16. **Silverman L, Rubinstein LJ.** 1965. Electron microscopic observations on a case of progressive multifocal leukoencephalopathy. Acta Neuropathol. **5:**215–224.
17. **Zurhein G, Chou SM.** 1965. Particles resembling Papova viruses in human cerebral demyelinating disease. Science **148:**1477–1479.
18. **Spiegel I, Peles E.** 2006. A new player in CNS myelination. Neuron **49:**777–778.
19. **Sorensen A, Moffat K, Thomson C, Barnett SC.** 2008. Astrocytes, but not olfactory ensheathing cells or Schwann cells, promote myelination of CNS axons *in vitro*. Glia **56:**750–763.
20. **Bradl M, Lassmann H.** 2010. Oligodendrocytes: biology and pathology. Acta Neuropathol. **119:**37–53.
21. **Astrom KE, Mancall EL, Richardson EP, Jr.** 1958. Progressive multifocal leuko-encephalopathy; a hitherto unrecognized complication of chronic lymphatic leukaemia and Hodgkin's disease. Brain **81:**93–111.
22. **Ferenczy MW, Marshall LJ, Nelson CD, Atwood WJ, Nath A, Khalili K, Major EO.** 2012. Molecular biology, epidemiology, and pathogenesis of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. Clin. Microbiol. Rev. **25:**471–506.
23. **Brew BJ, Davies NW, Cinque P, Clifford DB, Nath A.** 2010. Progressive multifocal leukoencephalopathy and other forms of JC virus disease. Nat. Rev. Neurol. **6:**667–679.
24. **Carson KR, Focosi D, Major EO, Petrini M, Richey EA, West DP, Bennett CL.** 2009. Monoclonal antibody-associated progressive multifocal leucoencephalopathy in patients treated with rituximab, natalizumab, and efalizumab: a review from the Research on Adverse Drug Events and Reports (RADAR) Project. Lancet Oncol. **10:**816–824.
25. **Bloomgren G, Richman S, Hotermans C, Subramanyam M, Goelz S, Natarajan A, Lee S, Plavina T, Scanlon JV, Sandrock A, Bozic C.** 2012. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N. Engl. J. Med. **366:**1870–1880.
26. **Hellwig K, Gold R.** 2011. Progressive multifocal leukoencephalopathy and natalizumab. J. Neurol. **258:**1920–1928.
27. **Kawamoto E, Nakahashi S, Okamoto T, Imai H, Shimaoka M.** 2012. Anti-integrin therapy for multiple sclerosis. J. Autoimmune Dis. 2012: 357101. http://dx.doi.org/10.1155/2012/357101.
28. **Tornatore C, Berger JR, Houff SA, Curfman B, Meyers K, Winfield D, Major EO.** 1992. Detection of JC virus DNA in peripheral lymphocytes from patients with and without progressive multifocal leukoencephalopathy. Ann. Neurol. **31:**454–462.
29. **Atwood WJ, Amemiya K, Traub R, Harms J, Major EO.** 1992. Interaction of the human polyomavirus, JCV, with human B-lymphocytes. Virology **190:**716–723.
30. **Chapagain ML, Nerurkar VR.** 2010. Human polyomavirus JC (JCV) infection of human B lymphocytes: a possible mechanism for JCV transmigration across the blood-brain barrier. J. Infect. Dis. **202:**184–191.
31. **Daniel AM, Swenson JJ, Mayreddy RP, Khalili K, Frisque RJ.** 1996. Sequences within the early and late promoters of archetype JC virus restrict viral DNA replication and infectivity. Virology **216:**90–101.
32. **Agostini HT, Ryschkewitsch CF, Stoner GL.** 1996. Genotype profile of human polyomavirus JC excreted in urine of immunocompetent individuals. J. Clin. Microbiol. **34:**159–164.
33. **Yogo Y, Zhong S, Shibuya A, Kitamura T, Homma Y.** 2008. Transcriptional control region rearrangements associated with the evolution of JC polyomavirus. Virology **380:**118–123.
34. **Sock E, Renner K, Feist D, Leger H, Wegner M.** 1996. Functional comparison of PML-type and archetype strains of JC virus. J. Virol. **70:**1512–1520.
35. **Frisque RJ.** 1983. Nucleotide sequence of the region encompassing the JC virus origin of DNA replication. J. Virol. **46:**170–176.
36. **Martin JD, King DM, Slauch JM, Frisque RJ.** 1985. Differences in regulatory sequences of naturally occurring JC virus variants. J. Virol. **53:**306–311.
37. **Kenney S, Natarajan V, Salzman NP.** 1986. Mapping 5′ termini of JC virus late RNA. J. Virol. **58:**216–219.
38. **Kenney S, Natarajan V, Selzer G, Salzman NP.** 1986. Mapping 5′ termini of JC virus early RNAs. J. Virol. **58:**651–654.
39. **Krebs CJ, McAvoy MT, Kumar G.** 1995. The JC virus minimal core promoter is glial cell specific in vivo. J. Virol. **69:**2434–2442.
40. **Frisque RJ.** 1983. Regulatory sequences and virus-cell interactions of JC virus. Prog. Clin. Biol. Res. **105:**41–59.
41. **Neu U, Maginnis MS, Palma AS, Ströh 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.
42. **Liu CK, Wei G, Atwood WJ.** 1998. Infection of glial cells by the human polyomavirus JC is mediated by an N-linked glycoprotein containing terminal alpha(2–6)-linked sialic acids. J. Virol. **72:**4643–4649.
43. **Gee GV, Tsomaia N, Mierke DF, Atwood WJ.** 2004. Modeling a sialic acid binding pocket in the external loops of JC virus VP1. J. Biol. Chem. **279:**49172–49176.
44. **Elphick GF, Querbes W, Jordan JA, Gee GV, Eash S, Manley K, Dugan A, Stanifer M, Bhatnagar A, Kroeze WK, Roth BL, Atwood WJ.** 2004. The human polyomavirus, JCV, uses serotonin receptors to infect cells. Science **306:**1380–1383.
45. **Dugan AS, Gasparovic ML, Atwood WJ.** 2008. Direct correlation between sialic acid binding and infection of cells by two human polyomaviruses (JC virus and BK virus). J. Virol. **82:**2560–2564.
46. **Eash S, Tavares R, Stopa EG, Robbins SH, Brossay L, Atwood WJ.** 2004. Differential distribution of the JC virus receptor-type sialic acid in normal human tissues. Am. J. Pathol. **164:**419–428.
47. **Wei G, Liu CK, Atwood WJ.** 2000. JC virus binds to primary human glial cells, tonsillar stromal cells, and B-lymphocytes, but not to T lymphocytes. J. Neurovirol. **6:**127–136.
48. **Maginnis MS, Haley SA, Gee GV, Atwood WJ.** 2010. Role of N-linked glycosylation of the 5-HT2A receptor in JC virus infection. J. Virol. **84:**9677–9684.
49. **Zheng HY, Takasaka T, Noda K, Kanazawa A, Mori H, Kabuki T, Joh K, Oh-ishi T, Ikegaya H, Nagashima K, Hall WW, Kitamura T, Yogo Y.** 2005. New sequence polymorphisms in the outer loops of the JC polyomavirus major capsid protein (VP1) possibly associated with progressive multifocal leukoencephalopathy. J. Gen. Virol. **86:**2035–2045.
50. **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. http://dx.doi.org/10.1371/journal.pgen.1000368.
51. **Delbue S, Branchetti E, Bertolacci S, Tavazzi E, Marchioni E, Maserati**

R, Minnucci G, Tremolada S, Vago G, Ferrante P. 2009. JC virus VP1 loop-specific polymorphisms are associated with favorable prognosis for progressive multifocal leukoencephalopathy. J. Neurovirol. **15:**51–56.
52. **Gorelik L, Reid C, Testa M, Brickelmaier M, Bossolasco S, Pazzi A, Bestetti A, Carmillo P, Wilson E, McAuliffe M, Tonkin C, Carulli JP, Lugovskoy A, Lazzarin A, Sunyaev S, Simon K, Cinque P.** 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.
53. **Reid CE, Li H, Sur G, Carmillo P, Bushnell S, Tizard R, McAuliffe M, Tonkin C, Simon K, Goelz S, Cinque P, Gorelik L, Carulli JP.** 2011. Sequencing and analysis of JC virus DNA from natalizumab-treated PML patients. J. Infect. Dis. **204:**237–244.
54. **Sala M, Vartanian JP, Kousignian P, Delfraissy JF, Taoufik Y, Wain-Hobson S, Gasnault J.** 2001. Progressive multifocal leukoencephalopathy in human immunodeficiency virus type 1-infected patients: absence of correlation between JC virus neurovirulence and polymorphisms in the transcriptional control region and the major capsid protein loci. J. Gen. Virol. **82:**899–907.
55. **Zheng HY, Ikegaya H, Takasaka T, Matsushima-Ohno T, Sakurai M, Kanazawa I, Kishida S, Nagashima K, Kitamura T, Yogo Y.** 2005. Characterization of the VP1 loop mutations widespread among JC polyomavirus isolates associated with progressive multifocal leukoencephalopathy. Biochem. Biophys. Res. Commun. **333:**996–1002.
56. **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.
57. **Gee GV, Manley K, Atwood WJ.** 2003. Derivation of a JC virus-resistant human glial cell line: implications for the identification of host cell factors that determine viral tropism. Virology **314:**101–109.
58. **Bauer PH, Bronson RT, Fung SC, Freund R, Stehle T, Harrison SC, Benjamin TL.** 1995. Genetic and structural analysis of a virulence determinant in polyomavirus VP1. J. Virol. **69:**7925–7931.
59. **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.
60. **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.
61. **Dubensky TW, Freund R, Dawe CJ, Benjamin TL.** 1991. Polyomavirus replication in mice: influences of VP1 type and route of inoculation. J. Virol. **65:**342–349.
62. **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.
63. **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.
64. **Komagome R, Sawa H, Suzuki T, Suzuki Y, Tanaka S, Atwood WJ, Nagashima K.** 2002. Oligosaccharides as receptors for JC virus. J. Virol. **76:**12992–13000.
65. **Cubitt CL, Cui X, Agostini HT, Nerurkar VR, Scheirich I, Yanagihara R, Ryschkewitsch CF, Stoner GL.** 2001. Predicted amino acid sequences for 100 JCV strains. J. Neurovirol. **7:**339–344.
66. **Major EO, Miller AE, Mourrain P, Traub RG, de Widt E, Sever J.** 1985. Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc. Natl. Acad. Sci. U. S. A. **82:**1257–1261.
67. **Schweighardt B, Atwood WJ.** 2001. HIV type 1 infection of human astrocytes is restricted by inefficient viral entry. AIDS Res. Hum. Retroviruses **17:**1133–1142.
68. **Chen BJ, Atwood WJ.** 2002. Construction of a novel JCV/SV40 hybrid virus (JCSV) reveals a role for the JCV capsid in viral tropism. Virology **300:**282–290.
69. **Vacante DA, Traub R, Major EO.** 1989. Extension of JC virus host range to monkey cells by insertion of a simian virus 40 enhancer into the JC virus regulatory region. Virology **170:**353–361.
70. **Atwood WJ, Wang L, Durham LC, Amemiya K, Traub RG, Major EO.** 1995. Evaluation of the role of cytokine activation in the multiplication of JC virus (JCV) in human fetal glial cells. J. Neurovirol. **1:**40–49.
71. **Tolstov YL, Pastrana DV, Feng H, Becker JC, Jenkins FJ, Moschos S, Chang Y, Buck CB, Moore PS.** 2009. Human Merkel cell polyomavirus infection II. MCV is a common human infection that can be detected by conformational capsid epitope immunoassays. Int. J. Cancer **125:** 1250–1256.
72. **Pastrana DV, Tolstov YL, Becker JC, Moore PS, Chang Y, Buck CB.** 2009. Quantitation of human seroresponsiveness to Merkel cell polyomavirus. PLoS Pathog. **5:**e1000578. http://dx.doi.org/10.1371/journal.ppat.1000578.
73. **Kabsch W.** 2010. XDS. Acta Crystallogr. D Biol. Crystallogr. **66:**125–132.
74. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40:** 658–674.
75. **Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, 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.
76. **Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66:**213–221.
77. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66:**486–501.
78. **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.
79. **Murshudov GN, Vagin AA, Dodson EJ.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. **53:**240–255.

Cell Host & Microbe
Article

Structure-Function Analysis of the Human JC Polyomavirus Establishes the LSTc Pentasaccharide as a Functional Receptor Motif

Ursula Neu,[1,5] Melissa S. Maginnis,[2,5] Angelina S. Palma,[3,6] Luisa J. Ströh,[1] Christian D.S. Nelson,[2] Ten Feizi,[3] Walter J. Atwood,[2,*] and Thilo Stehle[1,4,*]
[1]Interfaculty Institute for Biochemistry, University of Tübingen, D-72076 Tübingen, Germany
[2]Department of Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, RI 02912, USA
[3]Glycosciences Laboratory, Faculty of Medicine, Imperial College London, Harrow, Middlesex HA1 3UJ, UK
[4]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA
[5]These authors contributed equally to this work
[6]Present address: REQUIMTE, CQFB, Faculty of Science and Technology/UNL, 2829-516 Caparica, Portugal
*Correspondence: walter_atwood@brown.edu (W.J.A.), thilo.stehle@uni-tuebingen.de (T.S.)
DOI 10.1016/j.chom.2010.09.004

SUMMARY

The human JC polyomavirus (JCV) causes a fatal demyelinating disease, progressive multifocal leukoencephalopathy (PML), in immunocompromised individuals. Current treatment options for PML are inadequate. Sialylated oligosaccharides and the serotonin receptor are known to be necessary for JCV entry, but the molecular interactions underlying JCV attachment remain unknown. Using glycan array screening and viral infectivity assays, we identify a linear sialylated pentasaccharide with the sequence NeuNAc-α2,6-Gal-β1,4-GlcNAc-β1,3-Gal-β1,4-Glc (LSTc) present on host glycoproteins and glycolipids as a specific JCV recognition motif. The crystal structure of the JCV capsid protein VP1 was solved alone and in complex with LSTc. It reveals extensive interactions with the terminal sialic acid of the LSTc motif and specific recognition of an extended conformation of LSTc. Mutations in the JCV oligosaccharide-binding sites abolish cell attachment, viral spread, and infectivity, further validating the importance of this interaction. Our findings provide a powerful platform for the development of antiviral compounds.

INTRODUCTION

The human JC polyomavirus (JCV) is a member of the Polyomaviridae family, which also includes simian virus 40 (SV40), murine polyomavirus (Polyoma), and the human BK virus (BKV). JCV is a significant human pathogen for which approximately 50%–80% of individuals are seropositive (Kean et al., 2009; Knowles et al., 2003). JCV establishes a persistent, mostly asymptomatic infection in the kidney (Dorries, 1998). However, the virus can become reactivated in immunosuppressed hosts, leading to enhanced viral replication and infection of glial cells, including astrocytes and the myelin-producing oligodendrocytes, in the central nervous system (CNS). JCV infection leads to cytolytic destruction of oligodendroglia and causes the fatal disease progressive multifocal leukoencephalopathy (PML) (Khalili and White, 2006; Seth et al., 2003; Silverman and Rubinstein, 1965; Zurhein and Chou, 1965). PML is commonly associated with human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) (Cinque et al., 2009). However, the incidence of PML has increased recently in patients receiving immunosuppressive therapies for autoimmune diseases (Carson et al., 2009b; Major, 2009; Van Assche et al., 2005). The prognosis of PML is bleak, as it usually proves fatal within 1 year of symptom onset. Currently there are few treatments for PML (Kishida, 2007).

JCV has a nonenveloped, icosahedral capsid that encloses a circular double-stranded DNA genome (Shah et al., 1996). The major component of the capsid is viral protein 1 (VP1). Structural studies of SV40 and Polyoma virions revealed that polyomavirus capsids consist of 72 VP1 pentamers that are linked via C-terminal extensions of VP1 (Liddington et al., 1991; Stehle et al., 1994). VP1 attaches to receptors on host cells and thereby initiates infection. In crystal structures of Polyoma and SV40 receptor complexes, the receptor binding site is formed by VP1 loops at the outer surface of the capsid (Neu et al., 2008; Stehle et al., 1994). JCV is internalized into cells via clathrin-dependent endocytosis (Pho et al., 2000) and traffics through endosomes and caveosomes (Querbes et al., 2006) to the endoplasmic reticulum (ER). Viral uncoating is thought to begin in the ER, followed by translocation to the cytosol and transport of the genome to the nucleus for viral replication. Previous studies have shown that sialylated oligosaccharides (Dugan et al., 2008; Komagome et al., 2002; Liu et al., 1998b) and the serotonin receptor 5-HT_{2A}R (Elphick et al., 2004) are necessary for JCV entry. The predominant sialic acid in humans is α-5-*N*-acetyl neuraminic acid (NeuNAc), which caps many different oligosaccharide structures (Varki, 2001).

In this study, we demonstrate that JCV binds specifically to the oligosaccharide lactoseries tetrasaccharide c (LSTc), which contains an α2,6-linked NeuNAc, using a glycan array screen. The relevance of LSTc as a functional receptor for JCV is confirmed through infectivity assays. We determined the crystal structure of JCV VP1 in complex with LSTc, and based on the structural analysis performed mutagenesis experiments that

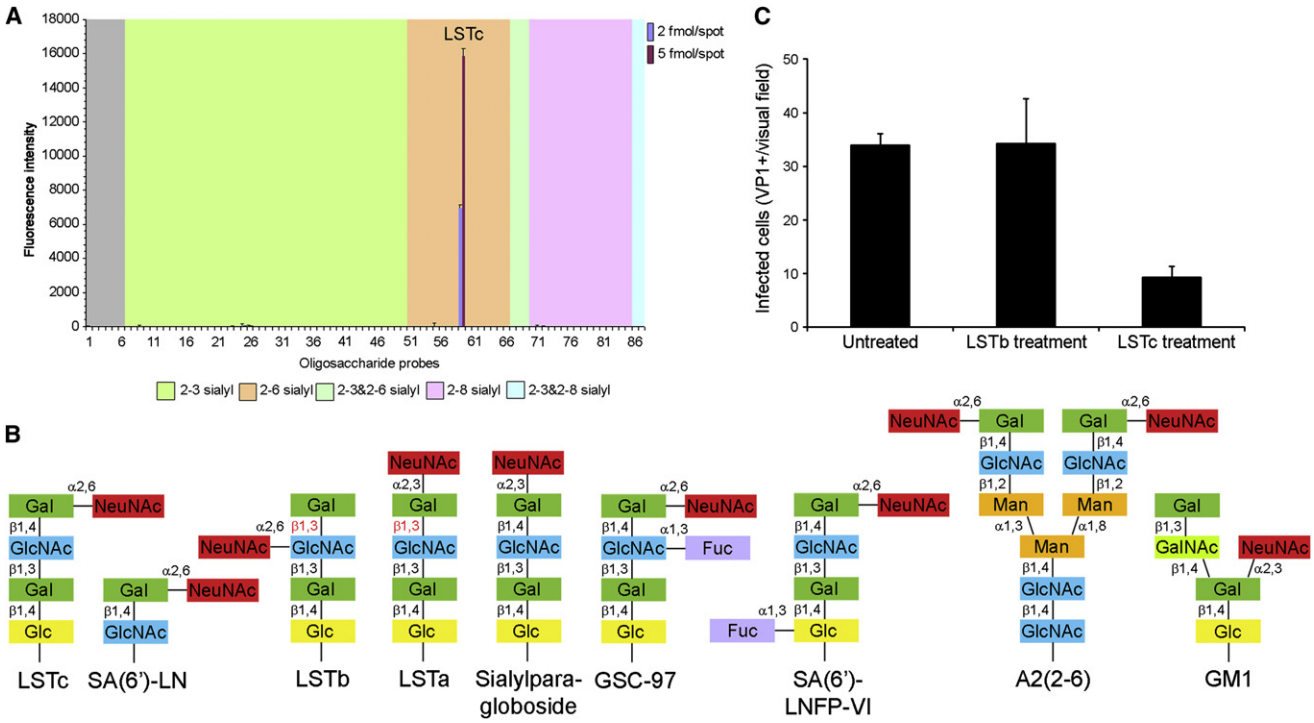


Figure 1. LSTc Is a Functional, Specific Receptor Motif for JCV

(A) Glycan microarray analysis of JCV VP1 showing highly selective binding to LSTc. Numerical scores for the binding intensity are shown as means of fluorescence intensities of duplicate spots at 2 (in blue) and 5 (in red) fmol/spot. Error bars represent half of the difference between the two values.

(B) Structures of selected glycans present on the glycan microarray. The oligosaccharide sequence of the LSTc probe is shown, as well as those of similar compounds that were not bound.

(C) LSTc inhibits JCV Infection. JC virus was preincubated with LSTb or LSTc, and complexes were added to SVG-A cells for infection. Infected cells were quantified based on nuclear VP1 staining. The data represent the average number of infected cells per visual field for eight fields of view from an experiment performed in triplicate. Error bars indicate SD of triplicate samples. *$p < 0.05$.

probe the specificity of JCV binding to LSTc. These studies establish the mechanism of JCV cell attachment and provide a foundation for strategies to intervene in the attachment process.

RESULTS

JCV Specifically Engages the Oligosaccharide LSTc

Whereas it is clearly established that JCV infection depends on interaction between VP1 and sialic acid, the nature of the oligosaccharide receptor has remained elusive. To approach this problem, we expressed recombinant JCV VP1 pentamers (Neu et al., 2008; Stehle and Harrison, 1997). The purified pentamers were analyzed on glycan microarrays containing 81 lipid-linked sialylated oligosaccharides differing in the sialyl linkage, backbone sequences, chain lengths, and branching patterns. The microarrays represented carbohydrates found on *N*- and *O*-linked glycoproteins as well as glycolipids (see Table S1 available online). JCV VP1 bound specifically to α2,6-linked LSTc (Figure 1A). LSTc is a linear pentasaccharide with the sequence NeuNAc-α2,6-Gal-β1,4-GlcNAc-β1,3-Gal-β1,4-Glc. This sequence represents the type that occurs in the peripheral region of glycans of glycoproteins and glycolipids. The glycan array contained several sequences that closely resembled LSTc, but VP1 did not bind to these (Figures 1A and 1B).

To determine whether JCV engages LSTc to mediate infection of host cells, JCV was preincubated with LSTc, and the mixture was added to SVG-A cells, a glial cell line permissive for JCV. The JCV/LSTc mixture significantly reduced JCV infection, while incubation with JCV/LSTb had no effect on infection (Figure 1C). LSTb is identical to LSTc in molecular weight and composition but features a branching α2,6-linked sialic acid (Figure 1B). Thus, the specific engagement of an LSTc-like sequence on host cells is critical for JCV infection.

Overall Structure of the JCV VP1-LSTc Complex

We solved the crystal structure of a JCV VP1 pentamer in complex with LSTc at 2.0 Å resolution (Figure 2A, Table 1). The final structure contains amino acids 25–85 and 100–287 for all five VP1 chains. Each VP1 monomer adopts the antiparallel β sandwich fold iconic for viral capsid proteins. Two four-stranded β sheets pack against each other. Their strands are linked by loops in a way that gives rise to a jelly roll topology, with the two β sheets consisting of strands B, I, D, G and C, H, E, F, respectively.

LSTc binds on top of the JCV VP1 pentamer, on the outer surface of the virion (Figures 2A and 2B). Two of the five possible binding sites on VP1 contain LSTc, the other three are occluded by crystal contacts. In the context of cell attachment by an entire

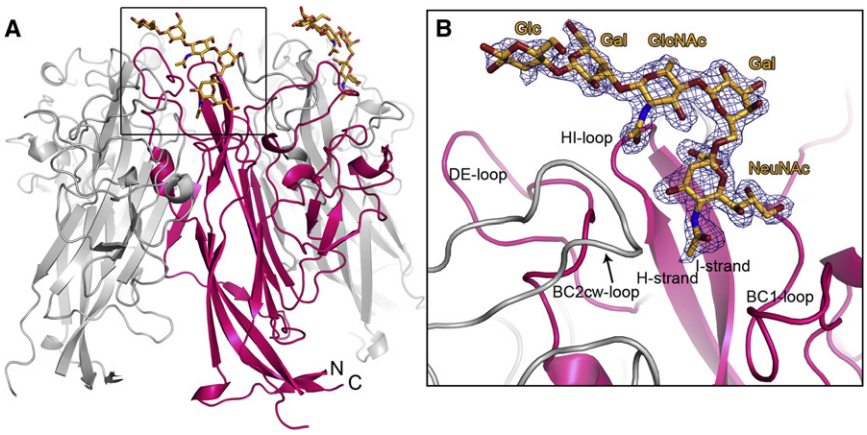


Figure 2. Structure of JCV VP1 in Complex with LSTc

(A) Structure of the JCV VP1 pentamer in complex with LSTc. The protein is shown in cartoon representation, with one VP1 monomer highlighted in pink and the other monomers depicted in gray. The LSTc oligosaccharide is drawn as a stick model and colored according to atom type (nitrogens in blue, oxygens in red, and carbons in orange).

(B) Close-up view of the LSTc binding site. JCV VP1 and LSTc are drawn as in (A). A composite annealed omit difference density map of LSTc is shown contoured at 3.0 σ for 2.0 Å around LSTc.

virion, however, all five binding sites on a given pentamer could be occupied by LSTc, which is not unlikely, given the relatively low affinity of a single protein-carbohydrate interaction. The binding site is formed by residues from the BC, DE, and HI loops from one monomer as well as the BC loop from its clockwise neighbor (Figure 2B). These loops also engage receptors in SV40 and Polyoma VP1, and they are the only parts of VP1 that differ significantly among the three viruses (Neu et al., 2008; Stehle and Harrison, 1997). The extensive BC loop can be subdivided into two parts facing in different directions (BC1 and BC2). LSTc binds where the BC1 loop of one monomer approaches the tip of the BC2 loop of its clockwise neighbor, and it interacts with both. For reasons of clarity, residues and loops of the main binding VP1 monomer will not be specifically labeled, while those of its clockwise and counterclockwise neighbors will be designated cw and ccw, respectively, herein.

Table 1. Crystallographic Statistics of JCV VP1 and JCV VP1-LSTc Structures

	JCV VP1-LSTc Complex		JCV VP1	
Data Collection				
Unit cell (Å)	149.7, 96.3, 128.2		149.6, 96.0, 128.6	
Unit cell (°)	90, 110.3, 90		90, 110.4, 90	
Space group	C2		C2	
Resolution (Å)	35–2.0	2.05 –2.00	20–1.95	2.00–1.95
Total reflections	437,707	30,603	419,669	15,867
Unique reflections	114,938	8,302	119,853	8,429
I/σI	9.6	2.7	10.9	2.5
R_{merge} (%)	11.7	48.3	9.3	32.4
Completeness (%)	99.5	97.5	96.5	91.7
Structure Refinement				
Protein atoms	10,200		10,135	
Average B factor (Å^2)	16.6		15.7	
Carbohydrate atoms	136		-	
Average B factor (Å^2)	24.3		-	
Solvent atoms	980		1,054	
Average B factor (Å^2)	25.4		26.4	
R_{work} (%)	18.9		18.0	
R_{free} (%)	21.2		20.8	
Rmsd bond lengths [Å]	0.006		0.007	
Rmsd bond angles (°)	0.9		1.0	

Interactions of JCV VP1 with LSTc

The bound LSTc molecule resembles the letter "L," with the shorter leg formed by NeuNAc and its α2,6-linkage to Gal, and the longer leg consisting of the GlcNAc-β1,3-Gal-β1,4-Glc trisaccharide (Figure 2B). In solution, the glycosidic bonds of LSTc can rotate freely due to its unbranched sequence, and the molecule can assume different conformations (Breg et al., 1989). However, as the NeuNAc and GlcNAc moieties of LSTc are attached to positions 6 and 1 of the same Gal residue, all conformations feature a kink. JCV VP1 interacts selectively with one of these conformations, contacting sugar moieties in both legs of the L-shaped ligand. The observed conformation of LSTc as well as its orientation with respect to the VP1 surface differs from previously published homology models of JCV interactions with oligosaccharides (Gee et al., 2004; Sunyaev et al., 2009). These homology models were based on interactions observed in the structure of Polyoma VP1 in complex with an α2,3-sialylated oligosaccharide (Stehle and Harrison, 1997). Although NeuNAc is bound in a similar location at the outer surface of both proteins, the orientation of NeuNAc with respect to the protein surface is different, and the interactions with the oligosaccharide structures are not conserved.

Most of the interactions between JCV and LSTc involve the three longer functional groups projecting from the terminal NeuNAc residue (Figures 3A and 3B). The NeuNAc carboxylate faces toward the HI loop and is recognized by parallel hydrogen bonds to the side chains of S266 and S268, as well as a water-mediated hydrogen bond to S266. The extended NeuNAc glycerol chain docks into a shallow depression on the protein surface that is shaped by residues in the BC and HI loops. The bottom of the depression is formed by the side chain of Q270. Hydrogen bonds to the amide and hydroxyl groups of S60, located at the rim of the depression, further anchor the glycerol chain. The methyl moiety of the *N*-acetyl group inserts into a hydrophobic cavity at the intersection between the HI, BC1, and BC2cw loops. The

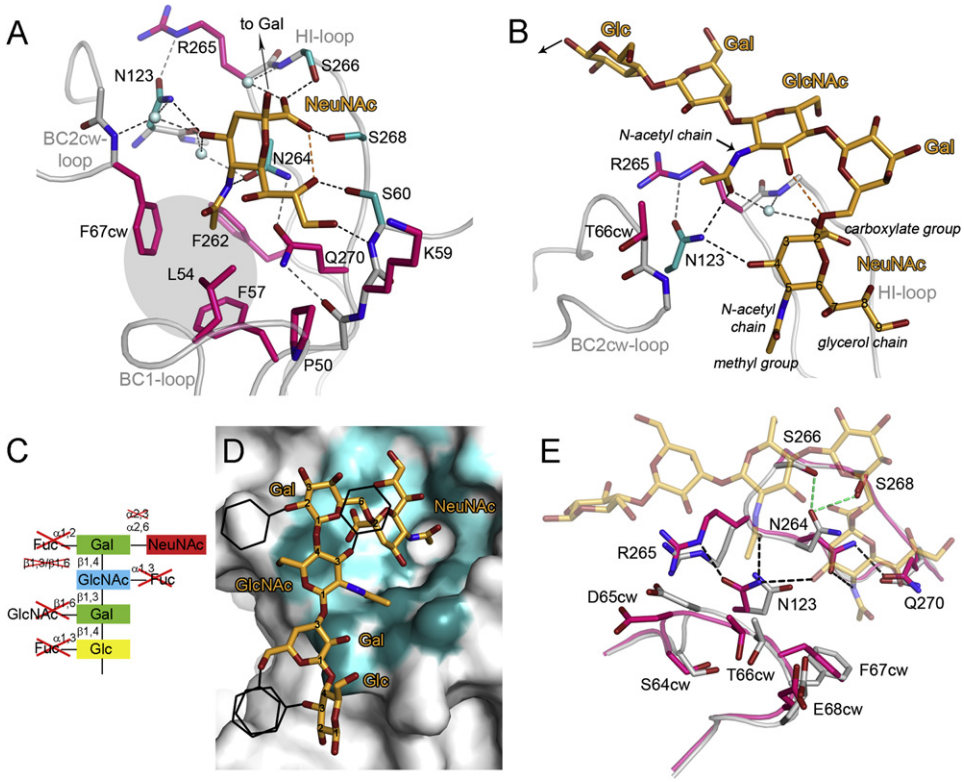


Figure 3. Interactions between JCV VP1 and LSTc

(A) Interactions between JCV VP1 and the terminal NeuNAc of LSTc. JCV VP1 is shown as a cartoon, with side chains interacting with LSTc in stick representation. Waters are represented with spheres. Residues forming direct hydrogen bonds to NeuNAc are colored teal and residues forming van der Waals contacts or water-mediated hydrogen bonds are colored pink. Direct hydrogen bonds between JCV VP1 and NeuNAc are shown as black dashed lines, and water-mediated hydrogen bonds or bonds between protein atoms are colored gray. Intramolecular hydrogen bonds within the oligosaccharide are orange.

(B) Interactions between JCV VP1 and other parts of LSTc.

(C) The cartoon represents structural features of oligosaccharides that are required for JCV binding. These were extracted from our glycan microarray data. Crossed-out sugar residues would produce steric clashes.

(D) Structural basis for JCV VP1 specificity for LSTc. JCV is shown in surface representation, with residues interacting with LSTc colored according to their change in surface accessibility upon LSTc binding (gray <1 Å^2 change, light teal 1–10 Å^2 change, dark teal >10 Å^2 change). The branching substitutions at LSTc that abolish binding are indicated as black hexagons, indicating where they would clash with protein or LSTc atoms.

(E) Structural changes in JCV VP1 upon LSTc binding. The structures of unliganded (gray) and liganded (pink) JCV VP1 were superposed using the β sandwich core residues. Hydrogen bonds only present in unliganded VP1 are indicated with green dashes, and those only present in the complex are colored black.

N-acetyl group is hydrogen bonded to the side chain of N264. Finally, the O4 hydroxyl group of NeuNAc forms a hydrogen bond with the side chain of N123. The O4 hydroxyl group also makes two water-mediated hydrogen bonds to N123 and F67cw.

Specific contacts between JCV VP1 and the long leg of the L-shaped LSTc mostly involve the GlcNAc residue. The side chain of N123 forms hydrogen bonds to both the carbonyl of the GlcNAc *N*-acetyl group and the side chain of R265. In turn, the carbonyl group of the *N*-acetyl chain forms water-mediated hydrogen bonds with S266 and the carboxyl group of the NeuNAc residue. The methyl group of the GlcNAc *N*-acetyl chain makes van der Waals interactions with the methyl group of T66cw. In addition, there is a hydrogen bond within LSTc linking GlcNAc with the glycosidic oxygen of Gal.

Interestingly, 6-sialyl-lactosamine, which like LSTc has a NeuNAc-α2,6-Gal-β1,4-GlcNAc motif, was not recognized by JCV VP1. Since the oligosaccharide is directly linked to a lipid tag (Table S1), productive engagement of this compound might be prevented because the GlcNAc group, which contributes contacts in the case of LSTc, is in close proximity to the lipid here. Thus, for reasons of accessibility, JCV requires at least the LSTc pentasaccharide as a binding motif (Figure 3C). The glycan array tested the effect of additional, branching residues to this minimal binding motif. Attachment of a fucose to O3 of GlcNAc and O3 of Glc, or attachment of a branching oligosaccharide chain to O6 of Gal (probe 64, Figure 1A, Table S1), would lead to clashes with either JCV VP1 or other LSTc residues (Figure 3D). Attachment of a fucose to O2 of the first Gal does not occur in parallel with sialylation of that group. However, the presence of a GlcNAc residue instead of the terminal Glc of LSTc, as present in most glycans carrying LSTc-like sequences, is unlikely to alter the conformation of the oligosaccharide.

There are no hydrogen bonds between JCV VP1 and the remaining three sugar moieties in LSTc. However, the Glc and two Gal residues contribute to the surface area buried upon

complex formation (Figure 3D) and are well-ordered in the structure, indicating low flexibility due to van der Waals interactions with VP1. These interactions are clearly relevant for affinity and specific binding of JCV, as demonstrated by the glycan microarray.

JCV VP1 Undergoes Induced-Fit Movements upon Binding LSTc

To determine whether JCV VP1 undergoes conformational changes upon ligand binding, we solved the structure of unliganded JCV VP1 at 1.95 Å resolution (Table 1). Comparison with liganded VP1 shows that residues in the binding site rearrange upon ligand binding (Figure 3E). In the unliganded structure, the N264 side chain is oriented toward the side chains of S266 and S268 and forms hydrogen bonds with both residues. Upon ligand binding, N264 swings around to interact with Q270 and the *N*-acetyl chain of NeuNAc. In the unliganded structure, the "binding" conformation of N264 is already present as a second conformation, suggesting a dynamic equilibrium between two conformations that is shifted by ligand binding.

More drastic changes are seen at residue N123, whose hydrogen bonds are critical for recognizing the L-shaped binding conformation of LSTc. In the absence of LSTc, the N123 side chain points closer toward the NeuNAc binding site, and would clash with NeuNAc. To accommodate NeuNAc, the N123 side chain moves. Its altered position leads to clashes with mainchain atoms in the BC2cw loop, and so residues 64–68 of the BC2cw loop move to accommodate the new conformation of N123. Induced-fit movements were not observed in receptor interactions of SV40 and Polyoma (Neu et al., 2008; Stehle and Harrison, 1997). The induced fit of JCV VP1 may facilitate binding to LSTc in a specific conformation.

Carbohydrate Binding Sites in VP1 Are Necessary for JCV Attachment, Infection, and Spread

The crystal structure of VP1 in complex with LSTc revealed numerous contacts between the protein and LSTc. To test whether binding of LSTc is required for JCV infection, we introduced point mutations in the oligosaccharide-binding sites of VP1 into an infectious JCV clone (highlighted in Figure 4A). One set of mutants was designed to abolish the interaction with NeuNAc, either by removal of hydrogen bonds (S266A/S268A) or by steric hindrance (S266N/S268N and L54Y). We also probed the importance of N123, which undergoes induced-fit movements to specifically recognize the L-shaped conformation of LSTc, by mutating it to alanine. In addition, we mutated A126 and R265 in a unique groove on the VP1 surface (Figure 4A), which is not present in Polyoma and SV40 (Neu et al., 2008; Stehle et al., 1994) and may be relevant for interactions with JCV cell surface receptors. In addition, R265 forms van der Waals contacts with the long leg of LSTc. Mutants were analyzed for infectious spread in long-term growth assays (Figures 4B and 4D). On the final day of the growth assay, supernatants were harvested and used to inoculate SVG-A cells for an infectivity assay (Figures 4C and 4E). Mutation of residues in the NeuNAc binding site (L54Y, S266A/S268A, and S266N/S268N) completely abolished JCV spread and infectivity (Figures 4B and 4C). Interactions mediated by each of these residues must therefore be required for a productive interaction with sialic acid, which in turn is essential for JCV growth. N123A had delayed viral spread in culture compared to wild-type (Figure 4D) and exhibited a 50% reduction in infectivity (Figure 4E), demonstrating the important role of N123 in accommodating LSTc. Mutation of R265, which forms van der Waals contacts with LSTc, did not affect growth and infectvity but resulted in reduced binding (Figure 4F). Virus with a mutation in the groove (A126Q) was capable of infection, and grew to levels equivalent to wild-type JCV. This indicates that, while the groove is unique to JCV VP1, residue A126 likely does not contribute to receptor binding. Virus carrying the more drastic A126Y mutation was unable to grow. We envisage that this mutation introduces local structural changes that affect residues critical for sialic acid binding.

The residues targeted for mutation are far from the VP1 regions that mediate capsid assembly or interact with minor capsid proteins. In order to verify that the mutants do not alter or destablize the structure of VP1, we introduced the mutations into the pentamer expression construct and expressed mutant proteins in *E. coli*. Like wild-type, all mutant proteins were purified as pentamers, indicating that they are properly folded (Figure S1A). Circular dichroism spectroscopy experiments confirm that secondary structure compositions of wild-type and mutant proteins are identical (Figure S1B).

To determine whether the decrease in viral growth and infectivity exhibited by mutant viruses was attributable to virus binding, we analyzed wild-type and mutant JCV VP1 pentamers in a flow cytometry-based cell-binding assay (Figure 4F). Purified pentamers were incubated with SVG-A cells, and binding was detected using an antibody specific for their His tag. Flow cytometry analysis revealed that mutants that do not support infection (L54Y, A126Y, S266A/S268A and S266N/S268N) or do so at reduced levels (N123A) also have reduced binding to cells in comparison to wild-type VP1. Mutants with wild-type-like growth properties (A126Q, R265A) were capable of binding to host cells. While A126Q bound similar to wild-type, R265A exhibits reduced binding. As the R265 side chain forms van der Waals contacts with LSTc (Figures 3 and 4), its mutation likely results in a modest reduction of affinity that can be compensated by avidity in the context of the virion and thus does not affect viral growth and infectivity. These data demonstrate that the specific contacts between JCV VP1 and LSTc observed in the crystal structure play a critical role in JCV attachment and infection.

Structural Basis for the Binding Specificity of JCV

The crystal structure provides a platform for understanding the binding properties of JCV VP1 to host cells. Comparison with the structure of SV40 VP1 in complex with its receptor GM1, which carries an α2,3-linked NeuNAc at a branching point (Figure 1B), reveals that JCV and SV40 bind terminal NeuNAc at the same location in the VP1 pentamer (Neu et al., 2008). The specific interactions with the functional groups of NeuNAc are conserved (Figure 5), highlighting their importance.

However, the JCV and SV40 VP1 proteins differ at the cavity into which the methyl group of the terminal NeuNAc inserts. In SV40 VP1, this cavity is quite large, polar at its rim, and likely water filled, while the cavity is smaller and more hydrophobic in JCV. This reflects the different host organisms of the two viruses.

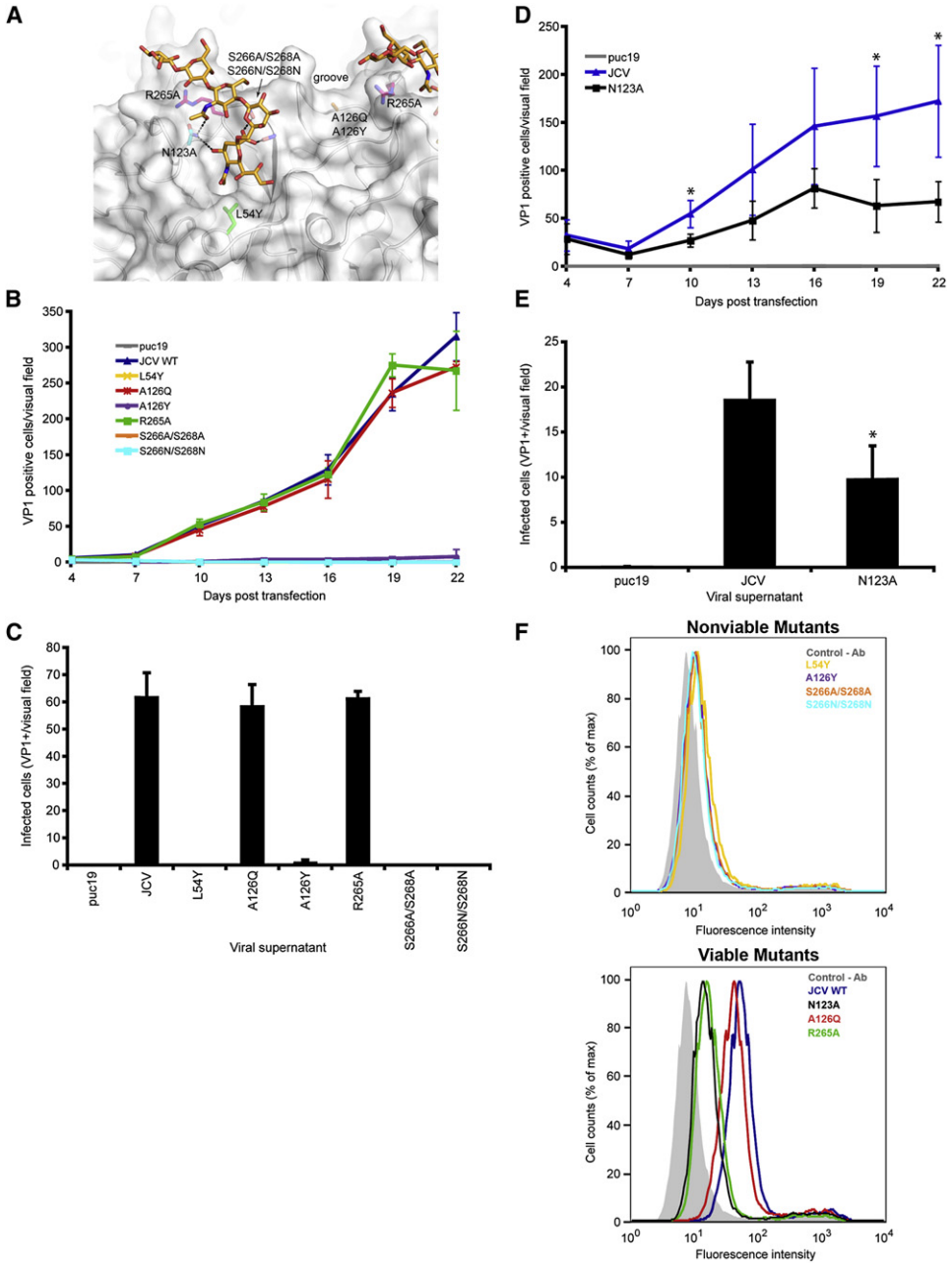


Figure 4. Growth, Infectivity, and Binding of JCV VP1 Mutants

(A) Point mutations introduced into JCV VP1. LSTc is shown in stick representation.

(B) Growth of JCV VP1 wild-type and mutant viruses. SVG-A cells were transfected with linearized DNA from JCV VP1 wild-type and mutant constructs. Transfected cells were fixed and stained at day 4 posttransfection, then at 3 day intervals for 22 days by indirect immunofluorescence. Transfected or infected cells were quantified based on nuclear VP1 staining. Each data point represents the average number of infected cells per visual field for ten fields of view for three independent experiments. Error bars indicate SD.

(C) Infectivity of supernatants from JCV VP1 wild-type and mutant viruses. SVG-A cells were inoculated with supernatants harvested from infected cells at day 22 from (B). Cells were fixed and stained by indirect immunofluorescence at 72 hr postinfection and quantified based on nuclear VP1 staining. The results are presented as the average number of infected cells per visual field for ten visual fields from three individual samples performed in triplicate. Error bars indicate SD.

(D) Growth of JCV VP1 wild-type and N123A. N123A was analyzed for viral growth as in (B). Each data point represents the average of number of infected cells per visual field for ten fields of view for three independent experiments. Error bars indicate SD. *$p < 0.05$.

(E) Infectivity of supernatants from JCV VP1 wild-type and N123A. N123A was analyzed for infectivity as in (C). Error bars indicate SD. *$p < 0.05$.

(F) Cell-binding analysis of JCV wild-type and mutant pentamers. SVG-A cells were incubated with His-tagged wild-type or mutant pentamers and a Penta His Alexa Fluor 488 antibody. Cells were fixed and pentamer binding was analyzed by flow cytometry. Histograms represent the fluorescence intensity of Alexa 488 for antibody alone (filled) and pentamer samples (open) for 10,000 gated events. Data are grouped into two histograms based on mutants that propagate (bottom) or do not propagate (top) in SVG-A cells.

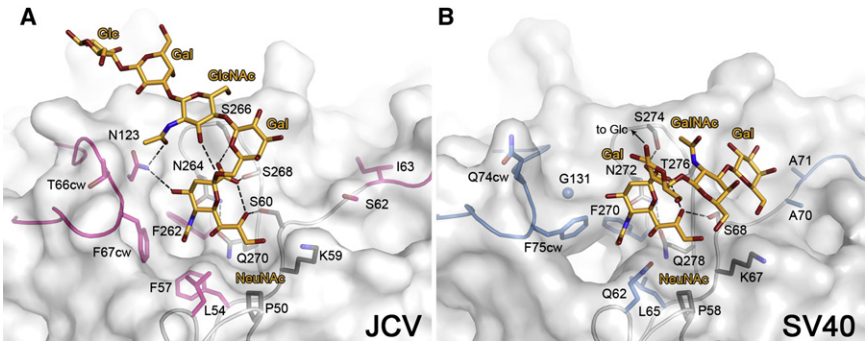


Figure 5. Comparison of Oligosaccharide Binding Sites of JCV and SV40 VP1

(A) JCV VP1 in complex with LSTc. (B) SV40 VP1 in complex with GM1. The Glc in GM1 does not contact the protein and was omitted for clarity. The proteins are shown in surface representation, with the BC and HI loops also indicated in cartoon representation. Residues contributing to ligand binding or specificity are shown in stick representation. They are colored gray when they are in the same conformation in the two proteins. Residues that are not conserved and assume different conformations are colored pink for JCV and blue for SV40. A blue sphere indicates the Cα position of G131 in SV40. The carbohydrate ligands are shown as orange sticks. Key hydrogen bonds are shown as black dashes.

In simians, the natural hosts of SV40, the predominant sialic acid is not NeuNAc but *N*-glycolyl neuraminic acid (NeuNGc), which carries the larger and more polar CH_2-OH group instead of the methyl group of NeuNAc (Varki, 2001). SV40 binds to NeuNGc-GM1 better than to NeuNAc-GM1 (Campanero-Rhodes et al., 2007), which is consistent with its larger, more polar cavity. However, NeuNAc is the most abundant sialic acid in humans, and the smaller and more hydrophobic cavity of JCV is consistent with its preference of NeuNAc-LSTc over NeuNGc-LSTc. Thus, the structure of the oligosaccharide binding site has consequences for the species specificity of the virus.

The position of NeuNAc in the larger context of an oligosaccharide is clearly a major determinant of specificity (Figures 3C and 3D). Amino acid N123 is key to recognizing the observed L-shaped conformation of LSTc, since it makes hydrogen bonds with both the GlcNAc and terminal NeuNAc sugar residues (Figures 3A and 3B). Different linkages of NeuNAc, Gal, and GlcNAc would not produce the precise arrangement of NeuNAc and GlcNAc that is required for complex formation. As a result, sialylparagloboside, which carries α2,3-linked NeuNAc but is otherwise identical in sequence to LSTc, is not recognized (Figure 1). The residue corresponding to N123 in SV40 is a glycine, which would not be able to mediate these contacts (Figure 5), explaining why SV40 does not attach to sequences terminating in NeuNAc-α2,6-Gal-β1,4-GlcNAc (Campanero-Rhodes et al., 2007; Neu et al., 2008).

Our data therefore provide strong support for a mode of binding that requires both a terminal NeuNAc and the extended L-shaped binding conformation of LSTc in the context of a long, linear chain.

DISCUSSION

We demonstrate here by glycan array screening that recombinant JCV VP1 pentamers selectively bind the LSTc oligosaccharide. The observed interaction is of physiologic relevance, as JCV infection can be blocked specifically by incubation with soluble LSTc. In order to understand the molecular basis for this recognition, we determined the structure of the JCV capsid protein, VP1, in complex with LSTc at 2.0 Å resolution. The structure shows that JCV engages both the terminal sialic acid and the unique L-shaped conformation of LSTc, which results in additional contacts of the GlcNAc moiety with VP1 and likely is a key determinant of its interaction with JCV VP1. Mutagenesis of residues involved in sialic acid binding results in viruses that are unable to grow and infect due to a binding defect.

Our finding that glycans terminating in the LSTc motif serve as main receptors for JCV is consistent with biochemical assays involving natively glycosylated cells (Liu et al., 1998b) (Figure 1B). Homologs of LSTc, in which a GlcNAc replaces the terminal Glc, are present in the human body as the termini of long oligosaccharide chains of the "i" antigen type (Feizi, 1985). All of the glycans terminating in GlcNAc-LSTc contain repeating units of [3-Gal-β1,4-GlcNAc-β1-]$_n$ (polyLacNAc repeats), and fulfill the steric condition of a long carbohydrate that is unbranched at its tip. These compounds should therefore bind JCV. The i antigen is expressed on a high proportion of human peripheral lymphocytes (Feizi et al., 1980). Glycan profiling by mass spectrometry has identified long glycans consistent with the LSTc motif in the *N*-glycan fractions of several human tissues including kidney, lung, and spleen (http://www.functionalglycomics.org/), but tissues of the entire CNS were not analyzed. Lectin staining of human tissues revealed that α2,6-linked NeuNAc is present on the surface of B lymphocytes of the tonsils and spleen, on kidney and lung cells, as well as on astrocytes and oligodendrocytes (Eash et al., 2004). Only very low-level lectin staining was seen for neurons and T lymphocytes. Notably, neurons, which express the serotonin receptor, are not susceptible to infection with JCV, most likely because they lack the oligosaccharide receptor. As the expression pattern of LSTc-carrying glycans corresponds to sites of JCV persistence (the kidney), and pathogenesis (the glial cells of the CNS), it is likely that the observed specificity of JCV for LSTc contributes to targeting the virus to these sites.

Previous studies report that JCV was also able to use α2,3-linked sialic acids as receptors (Dugan et al., 2008). It is possible that our glycan microarrays performed with pentamers failed to detect low-affinity interactions of JCV with α2,3-linked compounds because pentamers bind with much lower avidity than virions. However, given the strong signal observed for LSTc, such interactions would be very weak. Inspection of the structure shows that linear α2,3-linked compounds would not be able to interact favorably with JCV VP1. However, branched compounds such as GM1 and GM2 would be able to bind, albeit with significantly lower affinity compared to LSTc because of their much smaller contact areas. Interestingly, unpublished

results by Leonid Gorelik and the Consortium for Functional Glycomics (CFG), which are available online at http://www.functionalglycomics.org/ according to CFG policy, confirm these predictions. In glycan array screens of JCV VLPs of Mad-1 (which has the same VP1 as the strain used in our studies), strong binding was observed to LSTc-containing sequences, and weak binding was observed for GM2, GM1, and similar short, branched, α2,3-linked sequences. The weaker interactions with GM1 and GM2 would not appear to be functional because SV40, which does use GM1 as a receptor, does not compete with JCV for binding sites (Liu et al., 1998a).

In contrast to the related polyomaviruses SV40, Polyoma, and BKV, JCV does not enter cells by cholesterol-dependent endocytosis but instead uses clathrin-mediated endocytosis. It is enticing to speculate that this difference is reflected by their carbohydrate receptor specificities. SV40, Polyoma, and BKV all use gangliosides (Low et al., 2006; Tsai et al., 2003), which in the case of SV40 initiate entry by inducing curvature of the plasma membrane (Ewers et al., 2010). To induce curvature upon binding, the oligosaccharide portion of the receptor must be relatively rigid, which is the case for short, branched gangliosides (Acquotti et al., 1991), but not for long, flexible polyLacNAc extensions. Thus, the few amino acid changes between JCV and related polyomaviruses might not only determine receptor specificity but also target the virus to a different entry pathway.

Statistical analyses of JCV sequences demonstrate that 52% of PML patients carry JCV with one mutation in VP1 (L54F, K59M/E/N, N264D/T, S266F/L, or S268F/Y/C) that likely arose from positive selection (Sunyaev et al., 2009). Residues L54, N264, S266, and S268 mediate direct contacts with NeuNAc. Mutation of any of these four residues therefore adversely affects sialic acid binding (Figures 3A and 4). Consistently, the L54F and S268F mutants have reduced hemagglutination activity (Sunyaev et al., 2009). Although residue K59 also contacts NeuNAc, its replacement with M, E, or N would still allow for similar contacts. Not all of the mutations can therefore be attributed to a selective pressure on the virus to alter its receptor binding site.

It is interesting to speculate about how mutations that should block LSTc binding would be selected for in individuals with PML. It is conceivable that JCV featuring a mutated oligosaccharide binding site can spread more rapidly to the CNS from its site of persistence in the kidney because of diminished attachment to carbohydrates. This scenario would be similar to a point mutation in Polyoma VP1 that abolishes binding to nonfunctional pseudoreceptors, thereby increasing viral spread and rendering the virus more pathogenic (Freund et al., 1991). However, a defect in sialic acid binding would abolish engagement of both LSTc-carrying glycans and branched α2,3-linked compounds. Unpublished glycan array screening data by Leonid Gorelik and the CFG confirms this (http://www.functionalglycomics.org/). One also cannot ascertain whether JCV develops mutations in VP1 while replicating in remote sites of infection such as the kidney or if they develop in the brain. Moreover, it is not clear if patients carrying mutant JCV also carry wild-type JCV. Finally, viruses with the mutations L54Y, S266A/S268A, and S266N/S268N in the sialic acid binding site were unable to infect cells and propagate in a long-term growth assay because of a cell binding defect (Figure 4). Since we mutated the residues identified in PML patients, albeit to different amino acids, we would predict that the mutations present in PML patients would not propagate efficiently in glial cells.

In recent years, the incidence of PML has increased beyond the HIV/AIDS population to include patients undergoing immunosuppressive therapies for immunological disorders such as multiple sclerosis (Linda et al., 2009; Wenning et al., 2009), Crohn's disease (Van Assche et al., 2005), rheumatoid arthritis, and systemic lupus erythematosus (Carson et al., 2009a). Treatment options for PML are currently inadequate (Kishida, 2007), especially given the rapid devastation of this disease. We expect that knowledge about the structure of JCV VP1 and its interaction with receptors can serve as a platform from which such treatment strategies can be developed.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

cDNA coding for amino acids 22–289 of the Mad-1 strain of JCV VP1 was cloned into the pET15b expression vector (Novagen) in frame with an N-terminal hexahistidine tag (His tag) and a thrombin cleavage site. The N-terminal truncation was made because presence of these amino acids had inhibited crystallization of the homologous Polyoma VP1, while the C-terminal truncation was made to prevent capsid assembly in crystallization setups (Stehle and Harrison, 1997). Mutations were introduced using Phusion mutagenesis (NEB) according to the manufacturer's instructions and confirmed by sequencing (Table S2). Proteins were overexpressed in *E. coli* BL21(DE3) and purified using nickel affinity chromatography. For glycan array screening, flow cytometry, and circular dichroism, the proteins were additionally purified by gel filtration. For crystallization, the His tag was removed by incubation with thrombin prior to gel filtration. After cleavage, the nonnative amino acid sequence GSHM remains at the N terminus of VP1.

Glycan Array Screening

The microarray was composed of 87 sequence-defined lipid-linked oligosaccharide probes: 81 sialyl-terminating probes and six neutral probes as negative controls (Table S1). The probes were robotically printed in duplicate on nitrocellulose-coated glass slides at 2 and 5 fmol per spot using a noncontact instrument (Palma et al., 2006). The protein was precomplexed with mouse monoclonal anti-poly-histidine (Ab1) and biotinylated anti-mouse IgG antibodies (Ab2) (both from Sigma) in a ratio of 4:2:1 (by weight). In brief, the JCV VP1-His tagged protein-antibody precomplexes were prepared by preincubating Ab1 with Ab2 for 15 min at ambient temperature, followed by addition of VP1 and incubation for a further 15 min on ice. The VP1-antibody complexes were diluted in 5 mM HEPES (pH 7.4), 150 mM NaCl, 0.3% (v/v) Blocker Casein (Pierce), 0.3% (w/v) bovine serum albumin (Sigma), 2.5 mM DTT, and 5 mM $CaCl_2$ to give a final VP1 concentration of 150 μg/ml, and overlaid onto the arrays at 20°C for 2 hr. Binding was detected with Alexa Fluor 647-labeled streptavidin (Molecular Probes) and imaging (Palma et al., 2006), and data analysis (Stoll and Feizi, 2009) was as described.

Crystallization and Data Collection

JCV VP1 was concentrated to 4.5 mg/mL and crystallized by sitting drop vapor diffusion at 20°C. The reservoir solution contained 12% (w/v) PEG 3,350, 0.1 M HEPES (pH 7.5), and 0.2 M KSCN. Drops were set up by mixing 1 μl protein solution with 1 μl reservoir solution and 0.2 μl microseeding stock from previously obtained crystals. Crystals were harvested into a solution containing 10% PEG 3,350, 0.1 M HEPES (pH 7.5), and 0.2 M KSCN and cryoprotected by soaking for 10 s in harvesting solution containing 30% (v/v) glycerol and flash freezing in liquid nitrogen. For complex formation, crystals were soaked for 3 min in harvesting solution supplemented with 5 mM LSTc oligosaccharide. The cryoprotecting solution also contained the same concentration of LSTc oligosaccharide.

Structure Determination

Data reduction was performed with XDS (Kabsch, 2010), and the structure was solved by molecular replacement in Phaser (McCoy et al., 2007) using a search

model generated from the SV40 VP1 pentamer structure (3BWQ). Refinement was then performed with Phenix and Refmac5 (Adams et al., 2010; Murshudov et al., 1997), and model building was done in Coot (Emsley and Cowtan, 2004). Five-fold noncrystallographic symmetry restraints were used throughout the refinement for most of the protein. In data from crystals soaked in LSTc, the ligand was located in $2F_o–F_c$ and $F_o–F_c$ weighted electron density maps. The ligand was refined using restraints from the CCP4 library, with the exception of the α2,6-glycosidic bond, which had to be user defined. Waters were incorporated using Coot and ARP/wARP. The final models have good stereochemistry and agree very well with the experimental data (Table 1).

Cells, Viruses, and Antibodies

SVG-A cells are a subclone of the human glial cell line SVG transformed with an origin-defective SV40 mutant (Major et al., 1985). SVG-A cells were grown in Minimum Essential Medium supplemented to contain 10% fetal bovine serum (FBS) and penicillin/streptomycin (P/S) (Mediatech, Inc.) in a humidified incubator at 37°C. MAB597 is a hybridoma supernatant that produces a monoclonal antibody against JCV VP1 (Atwood et al., 1995) and was generously provided by Ed Harlow. Penta His Alexa Fluor 488 was used at 10 μg/mL (QIAGEN). Generation and propagation of the Mad-1/SVE strain of JCV were performed as previously described (Vacante et al., 1989).

Viral Growth and Reinfection Assays

VP1 mutations were generated in the genomic JCV DNA of strain JC12 (Chen and Atwood, 2002), subcloned in puc19 (Gee et al., 2004) with QuikChange (Stratagene) according to the manufacturer's instructions (Table S2), and confirmed by sequencing. Of plasmid DNA, 10 μg was digested with BamHI (Promega) for 2 hr at 37°C to separate the JCV genomic DNA from the puc19 backbone plasmid. Digests were performed in triplicate for each sample. SVG-A cells were plated to 40% confluence on 18 mm round glass coverslips (Thermo Scientific) in 12-well plates (Costar). Cells in medium without antibiotics were transfected with 2 μg of digested DNA using Fugene (Roche) at 3:2 ratio (Fugene:DNA). Transfected cells were incubated at 37°C O/N, and medium containing 5% FCS and 2% P/S was added to cells the next day. Cells were incubated at 37°C and fed with 500 μl of medium containing 5% FBS, 1% P/S, and 1% amphotericin B (Mediatech) or harvested for immunofluorescence at day 4 and at 3 day intervals thereafter for 22 days. For reinfection assays, supernatants were collected at 22 days posttransfection. SVG-A cells at 70% confluence on 22 mm square glass coverslips (Fisher) in 6-well plates (Costar) were infected with 150 μl of virus supernatant at 37°C for 1 hr, then medium containing 5% FBS, 1% P/S, 1% amphotericin B was added to cells, and cells were incubated at 37°C for 72 hr. Cells were fixed and stained by indirect immunofluorescence.

Indirect Immunofluorescence

Cells were washed in PBS, fixed in cold MeOH, and incubated at −20°C. Cells were washed in PBS, permeabilized with 0.5% TX-100 (USB Corporation) at RT for 15 min, blocked with 10% goat serum (MP Biomedicals)/PBS at RT for 30 min, incubated with VP1-specific antibody MAB597 (1:10) in PBS at 37°C for 1 hr, washed with PBS, incubated with a goat-anti-mouse Alexa Fluor 488 (1:1000) in PBS at 37°C for 1 hr, then washed with PBS. Cells were then mounted on slides using VectaShield with Dapi (Vector Laboratories Inc.). Cells were analyzed for nuclear VP1 staining under a 20× objective using an Eclipse 800 or Eclipse TE2000-U microscope (Nikon) equipped with an ORCA-ER digital camera (Hamamatsu).

LSTc Inhibition Assay

JCV (Mad-1/SVE) was pretreated with 5 mM of LSTb or LSTc (V Labs, Inc.) (diluted in sterile diH_2O) in media containing 2% FCS on ice for 1 hr. SVG-A cells in 96-well plates were prechilled at 4°C for 30 min. JCV-LST complexes were added to cells and incubated at 4°C for 1 hr. Cells were washed with PBS twice; media containing 10% FCS, 1% P/S, and 1% amphotericin B was added; and cells were incubated at 37°C for 72 hr. Cells were fixed and stained by indirect immunofluorescence as described above.

Flow Cytometry

Wild-type and mutant JCV VP1 pentamers (100 μl) in PBS were incubated with SVG-A cells in suspension on ice for 2 hr with occasional agitation. Cells were washed, pelleted by centrifugation, and suspended in 100 μl of Penta His Alexa Fluor 488 antibody (QIAGEN) (10 μg/mL) in PBS on ice for 1 hr. Cells were washed, pelleted, and fixed in 1% paraformaldehyde and analyzed for pentamer binding using a BD FACSCaliber (Benton, Dickinson, and Company) flow cytometer equipped with a 488 nm excitation line. Data were analyzed using BD CellQuestPro (Benton, Dickinson, and Company) and FlowJo (Tree Star, Inc) software.

Statistical Analysis

Means for triplicate samples were compared using an unpaired Student's t test (Microsoft Excel). P values <0.05 were considered statistically significant.

ACCESSION NUMBERS

Coordinates and structure factor amplitudes were deposited with the RCSB data bank (http://www.rcsb.org/) under accession codes 3NXD and 3NXG. Structure figures were prepared with PyMol (Schrödinger Inc.).

SUPPLEMENTAL INFORMATION

Supplemental Information includes one figure and two tables and can be found with this article online at doi:10.1016/j.chom.2010.09.004.

ACKNOWLEDGMENTS

We thank members of the Stehle, Atwood, and Feizi laboratories for critical discussion. We also thank the staff at beamlines X06SA and X06DA of the Swiss Light Source (Villigen, CH) for assistance with data collection. This work was funded by the National Institutes of Health (NIH) (Institute of Neurological Disorders and Stroke, NINDS) PPG 1P01NS065719-01A1 (W.J.A. and T.S.), the Deutsche Forschungsgemeinschaft SFB-685 (T.S.), the U.K. Research Councils Basic Technology Grant GR/S79268 "Glycoarrays" (T.F.), Engineering and Physical Research Councils Translational Grant EP/G037604/1 (T.F.), the NCI Alliance of Glycobiologists for Detection of Cancer and Cancer Risk U01 CA128416 (T.F.), the Fundação para a Ciência e Tecnologia SFRH/BPD/26515/2006 (A.S.P.), and Ruth L. Kirchstein National Research Service Awards F32NS064870 (M.S.M.) and F32NS070687 (C.D.S.N.).

Received: April 19, 2010
Revised: July 16, 2010
Accepted: August 17, 2010
Published: October 20, 2010

REFERENCES

Acquotti, D., Fronza, G., Ragg, E., and Sonnino, S. (1991). Three dimensional structure of GD1b and GD1b-monolactone gangliosides in dimethylsulphoxide: a nuclear Overhauser effect investigation supported by molecular dynamics calculations. Chem. Phys. Lipids *59*, 107–125.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Atwood, W.J., Wang, L., Durham, L.C., Amemiya, K., Traub, R.G., and Major, E.O. (1995). Evaluation of the role of cytokine activation in the multiplication of JC virus (JCV) in human fetal glial cells. J. Neurovirol. *1*, 40–49.

Breg, J., Kroon-Batenburg, L.M., Strecker, G., Montreuil, J., and Vliegenthart, J.F. (1989). Conformational analysis of the sialyl alpha(2-3/6)N-acetyllactosamine structural element occurring in glycoproteins, by two-dimensional NOE 1H-NMR spectroscopy in combination with energy calculations by hard-sphere exo-anomeric and molecular mechanics force-field with hydrogen-bonding potential. Eur. J. Biochem. *178*, 727–739.

Campanero-Rhodes, M.A., Smith, A., Chai, W., Sonnino, S., Mauri, L., Childs, R.A., Zhang, Y., Ewers, H., Helenius, A., Imberty, A., et al. (2007). N-glycolyl GM1 ganglioside as a receptor for simian virus 40. J. Virol. *81*, 12846–12858.

Carson, K.R., Evens, A.M., Richey, E.A., Habermann, T.M., Focosi, D., Seymour, J.F., Laubach, J., Bawn, S.D., Gordon, L.I., Winter, J.N., et al. (2009a). Progressive multifocal leukoencephalopathy after rituximab therapy in HIV-negative patients: a report of 57 cases from the Research on Adverse Drug Events and Reports project. Blood *113*, 4834–4840.

Carson, K.R., Focosi, D., Major, E.O., Petrini, M., Richey, E.A., West, D.P., and Bennett, C.L. (2009b). Monoclonal antibody-associated progressive multifocal leucoencephalopathy in patients treated with rituximab, natalizumab, and efalizumab: a review from the Research on Adverse Drug Events and Reports (RADAR) Project. Lancet Oncol. *10*, 816–824.

Chen, B.J., and Atwood, W.J. (2002). Construction of a novel JCV/SV40 hybrid virus (JCSV) reveals a role for the JCV capsid in viral tropism. Virology *300*, 282–290.

Cinque, P., Koralnik, I.J., Gerevini, S., Miro, J.M., and Price, R.W. (2009). Progressive multifocal leukoencephalopathy in HIV-1 infection. Lancet Infect. Dis. *9*, 625–636.

Dorries, K. (1998). Molecular biology and pathogenesis of human polyomavirus infections. Dev. Biol. Stand. *94*, 71–79.

Dugan, A.S., Gasparovic, M.L., and Atwood, W.J. (2008). Direct correlation between sialic acid binding and infection of cells by two human polyomaviruses (JC virus and BK virus). J. Virol. *82*, 2560–2564.

Eash, S., Tavares, R., Stopa, E.G., Robbins, S.H., Brossay, L., and Atwood, W.J. (2004). Differential distribution of the JC virus receptor-type sialic acid in normal human tissues. Am. J. Pathol. *164*, 419–428.

Elphick, G.F., Querbes, W., Jordan, J.A., Gee, G.V., Eash, S., Manley, K., Dugan, A., Stanifer, M., Bhatnagar, A., Kroeze, W.K., et al. (2004). The human polyomavirus, JCV, uses serotonin receptors to infect cells. Science *306*, 1380–1383.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Ewers, H., Romer, W., Smith, A.E., Bacia, K., Dmitrieff, S., Chai, W., Mancini, R., Kartenbeck, J., Chambon, V., Berland, L., et al. (2010). GM1 structure determines SV40-induced membrane invagination and infection. Nat. Cell. Biol. *12*, 11–18, Supp. 11–12.

Feizi, T. (1985). Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. Nature *314*, 53–57.

Feizi, T., Kapadia, A., and Yount, W.J. (1980). I and i antigens of human peripheral blood lymphocytes cocap with receptors for concanavalin A. Proc. Natl. Acad. Sci. USA *77*, 376–380.

Freund, R., Garcea, R.L., Sahli, R., and Benjamin, T.L. (1991). A single-amino-acid substitution in polyomavirus VP1 correlates with plaque size and hemagglutination behavior. J. Virol. *65*, 350–355.

Gee, G.V., Tsomaia, N., Mierke, D.F., and Atwood, W.J. (2004). Modeling a sialic acid binding pocket in the external loops of JC virus VP1. J. Biol. Chem. *279*, 49172–49176.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. *66*, 133–144.

Kean, J.M., Rao, S., Wang, M., and Garcea, R.L. (2009). Seroepidemiology of human polyomaviruses. PLoS Pathog. *5*, e1000363. 10.1371/journal.ppat.1000363.

Khalili, K., and White, M.K. (2006). Human demyelinating disease and the polyomavirus JCV. Mult. Scler. *12*, 133–142.

Kishida, S. (2007). Progressive multifocal leukoencephalopathy—epidemiology, clinical pictures, diagnosis and therapy. Brain Nerve *59*, 125–137.

Knowles, W.A., Pipkin, P., Andrews, N., Vyse, A., Minor, P., Brown, D.W., and Miller, E. (2003). Population-based study of antibody to the human polyomaviruses BKV and JCV and the simian polyomavirus SV40. J. Med. Virol. *71*, 115–123.

Komagome, R., Sawa, H., Suzuki, T., Suzuki, Y., Tanaka, S., Atwood, W.J., and Nagashima, K. (2002). Oligosaccharides as receptors for JC virus. J. Virol. *76*, 12992–13000.

Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L., and Harrison, S.C. (1991). Structure of simian virus 40 at 3.8-Å resolution. Nature *354*, 278–284.

Linda, H., von Heijne, A., Major, E.O., Ryschkewitsch, C., Berg, J., Olsson, T., and Martin, C. (2009). Progressive multifocal leukoencephalopathy after natalizumab monotherapy. N. Engl. J. Med. *361*, 1081–1087.

Liu, C.K., Hope, A.P., and Atwood, W.J. (1998a). The human polyomavirus, JCV, does not share receptor specificity with SV40 on human glial cells. J. Neurovirol. *4*, 49–58.

Liu, C.K., Wei, G., and Atwood, W.J. (1998b). Infection of glial cells by the human polyomavirus JC is mediated by an N-linked glycoprotein containing terminal alpha(2-6)-linked sialic acids. J. Virol. *72*, 4643–4649.

Low, J.A., Magnuson, B., Tsai, B., and Imperiale, M.J. (2006). Identification of gangliosides GD1b and GT1b as receptors for BK virus. J. Virol. *80*, 1361–1366.

Major, E.O. (2009). Reemergence of PML in natalizumab-treated patients—new cases, same concerns. N. Engl. J. Med. *361*, 1041–1043.

Major, E.O., Miller, A.E., Mourrain, P., Traub, R.G., de Widt, E., and Sever, J. (1985). Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc. Natl. Acad. Sci. USA *82*, 1257–1261.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. *40*, 658–674.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Neu, U., Woellner, K., Gauglitz, G., and Stehle, T. (2008). Structural basis of GM1 ganglioside recognition by simian virus 40. Proc. Natl. Acad. Sci. USA *105*, 5219–5224.

Palma, A.S., Feizi, T., Zhang, Y., Stoll, M.S., Lawson, A.M., Diaz-Rodriguez, E., Campanero-Rhodes, M.A., Costa, J., Gordon, S., Brown, G.D., et al. (2006). Ligands for the beta-glucan receptor, Dectin-1, assigned using "designer" microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. J. Biol. Chem. *281*, 5771–5779.

Pho, M.T., Ashok, A., and Atwood, W.J. (2000). JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. J. Virol. *74*, 2288–2292.

Querbes, W., O'Hara, B.A., Williams, G., and Atwood, W.J. (2006). Invasion of host cells by JC virus identifies a novel role for caveolae in endosomal sorting of noncaveolar ligands. J. Virol. *80*, 9402–9413.

Seth, P., Diaz, F., and Major, E.O. (2003). Advances in the biology of JC virus and induction of progressive multifocal leukoencephalopathy. J. Neurovirol. *9*, 236–246.

Shah, K.V., Fields, B.N., Knipe, D.M., and Howley, P.M. (1996), Third Edition. Fields Virology, *Volume 2* (Philadelphia: Lippincott-Raven Publishers).

Silverman, L., and Rubinstein, L.J. (1965). Electron microscopic observations on a case of progressive multifocal leukoencephalopathy. Acta Neuropathol. *5*, 215–224.

Stehle, T., and Harrison, S.C. (1997). High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. EMBO J. *16*, 5139–5148.

Stehle, T., Yan, Y., Benjamin, T.L., and Harrison, S.C. (1994). Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. Nature *369*, 160–163.

Stoll, M.S., and Feizi, T. (2009). Software tools for storing, processing and displaying carbohydrate microarray data. Paper presented at Beilstein Symposium on Glyco-Bioinformatics (Potsdam, Germany: Beilstein Institute for the Advancement of Chemical Sciences).

Sunyaev, S.R., Lugovskoy, A., Simon, K., and Gorelik, L. (2009). Adaptive mutations in the JC virus protein capsid are associated with progressive multifocal leukoencephalopathy (PML). PLoS Genet. *5*, e1000368. 10.1371/journal.pgen.1000368.

Tsai, B., Gilbert, J.M., Stehle, T., Lencer, W., Benjamin, T.L., and Rapoport, T.A. (2003). Gangliosides are receptors for murine polyoma virus and SV40. EMBO J. *22*, 4346–4355.

Vacante, D.A., Traub, R., and Major, E.O. (1989). Extension of JC virus host range to monkey cells by insertion of a simian virus 40 enhancer into the JC virus regulatory region. Virology *170*, 353–361.

Van Assche, G., Van Ranst, M., Sciot, R., Dubois, B., Vermeire, S., Noman, M., Verbeeck, J., Geboes, K., Robberecht, W., and Rutgeerts, P. (2005). Progressive multifocal leukoencephalopathy after natalizumab therapy for Crohn's disease. N. Engl. J. Med. *353*, 362–368.

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.

Wenning, W., Haghikia, A., Laubenberger, J., Clifford, D.B., Behrens, P.F., Chan, A., and Gold, R. (2009). Treatment of progressive multifocal leukoencephalopathy associated with natalizumab. N. Engl. J. Med. *361*, 1075–1080.

Zurhein, G., and Chou, S.M. (1965). Particles resembling Papova viruses in human cerebral demyelinating disease. Science *148*, 1477–1479.

⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 46, pp. 39254–39262, November 9, 2012

The Archaeal Proteasome Is Regulated by a Network of AAA ATPases*[S]

Received for publication, May 31, 2012, and in revised form, September 2, 2012 Published, JBC Papers in Press, September 19, 2012, DOI 10.1074/jbc.M112.386458

Dara Forouzan[‡1], Moritz Ammelburg[‡1], Cedric F. Hobel[‡2], Luisa J. Ströh[§], Nicole Sessler[¶], Jörg Martin[‡], and Andrei N. Lupas[‡3]

From the [‡]Department of Protein Evolution, Max Planck Institute for Developmental Biology, D-72076 Tübingen and the [§]Interfaculty Institute of Biochemistry and the [¶]Proteome Center, Interfaculty Institute of Cell Biology, University of Tübingen, D-72076 Tübingen, Germany

Background: The AAA ATPases of the PAN/Rpt1–6 group regulate access of substrates to the 20S proteasome.
Results: Two groups of AAA proteins, CDC48 and AMA, function as novel proteasomal ATPases in archaea.
Conclusion: This network of regulatory ATPases increases the capacity of proteasomal protein degradation in archaea.
Significance: Diversification at the level of the regulatory ATPase provides a contrast to the fully differentiated 26S proteasome of eukaryotes.

The proteasome is the central machinery for targeted protein degradation in archaea, Actinobacteria, and eukaryotes. In its basic form, it consists of a regulatory ATPase complex and a proteolytic core particle. The interaction between the two is governed by an HbY*X* motif (where Hb is a hydrophobic residue, Y is tyrosine, and *X* is any amino acid) at the C terminus of the ATPase subunits, which stimulates gate opening of the proteasomal α-subunits. In archaea, the proteasome-interacting motif is not only found in canonical proteasome-activating nucleotidases of the PAN/ARC/Rpt group, which are absent in major archaeal lineages, but also in proteins of the CDC48/p97/VAT and AMA groups, suggesting a regulatory network of proteasomal ATPases. Indeed, *Thermoplasma acidophilum*, which lacks PAN, encodes one CDC48 protein that interacts with the 20S proteasome and activates the degradation of model substrates. In contrast, *Methanosarcina mazei* contains seven AAA proteins, five of which, both PAN proteins, two out of three CDC48 proteins, and the AMA protein, function as proteasomal gatekeepers. The prevalent presence of multiple, distinct proteasomal ATPases in archaea thus results in a network of regulatory ATPases that may widen the substrate spectrum of proteasomal protein degradation.

Dynamic regulation of the proteome through protein degradation is a central metabolic strategy for organisms in a changing environment. In archaea, Actinobacteria, and eukaryotes, the proteasome plays a crucial role in protein degradation as part of stress responses and for regulatory purposes (1). Removal of misfolded proteins reduces the burden of potentially toxic proteins (2), and targeted destruction by the proteasome, especially in eukaryotes, is involved in processes such as cell cycle progression (3) and the cellular immune response (4). The proteasome is composed of a regulatory particle, which binds to and unfolds substrate proteins, and a catalytic particle, which fragments the substrates into peptides of 8–11 amino acids in length (5). An additional layer of regulation is provided by the covalent attachment of degradation tags to substrates, SAMP[4] in archaea (6), Pup in Actinobacteria (7), and ubiquitin in eukaryotes (8), to direct them to the proteasome.

The catalytic 20S proteasome core particle (CP) forms a cylindrical assembly consisting of four stacked heptameric rings with an $\alpha_7\beta_7\beta_7\alpha_7$ stoichiometry (9). This arrangement sequesters the active sites to a central compartment formed by the β-subunits, limiting access to only those proteins that present a degradation signal. The catalytic nucleophile, a threonine residue (10), is the N-terminal residue of β-subunits after autoprocessing of a leader peptide (11). The α-subunits guide substrates to the proteolytic chamber through a central channel, which is occluded in the resting state by their flexible N-terminal ends (12). The regulatory AAA ATPases Rpt1–6 (in eukaryotes), PAN (in archaea), and ARC/Mpa (in Actinobacteria) recognize, unfold, and translocate substrates into the CP (13–15). Members of this orthologous group share a common architecture, which consists of an N-terminal domain, involved in substrate binding and interaction with the degradation tag or with subunits of the regulatory particle (16–18), and an AAA unfoldase module. At their C terminus, the ATPases possess an HbY*X* motif (where Hb is a hydrophobic residue, Y is tyrosine, and *X* is any amino acid) (19) that can penetrate into a binding pocket of the α-subunits, thereby stabilizing the open gate conformation of their N-terminal ends (20). The functional impor-

* This work was supported by institutional funds of the Max Planck Society.
⌘ *Author's Choice*—Final version full access.
[S] This article contains supplemental Figs. S1–S3 and Table 1.
[1] Both authors contributed equally to this work.
[2] Present address: Novozymes A/S, Bagsvaerd 2880, Denmark.
[3] To whom correspondence should be addressed: Dept. of Protein Evolution, Max Planck Institute for Developmental Biology, Spemannstrasse 35, D-72076 Tübingen, Germany. Tel.: 49-7071-601-340; Fax: 49-7071-601-349; E-mail: andrei.lupas@tuebingen.mpg.de.

[4] The abbreviations used are: SAMP, small archaeal modifier protein; CP, core particle; PAN, proteasome-activating nucleotidase; ATPγS, adenosine-5′-(γ-thio)-triphosphate; Mm, *M. mazei*; Ta, *T. acidophilum*; Pup, prokaryotic ubiquitin-like protein; Rpt, regulatory particle triple-A ATPases; ARC, AAA ATPase forming ring-shaped complexes; Mpa, mycobacterial proteasome ATPase; VAT, VCP-like ATPase of Thermoplasma acidophilum; AMA, AAA protein of methanogenic archaea and Archaeoglobus; Mca, (7-Methoxycoumarin-4-yl)acetyl; Dpa, N-3-(2,4-Dinitrophenyl)-L-2,3-diaminopropionyl; Dnp, 2,4-dinitrophenyl.

tance of the HbY*X* motif is reflected by the ability of 7-residue C-terminal peptides, isolated or fused to the 11S/PA26 non-ATPase activator, to mimic the biochemical effects of full-length PAN (19–22).

Although preparations of the 26S proteasome from eukaryotes are routinely obtained via fractionation of whole cell lysate, in archaea, the preparation of proteasome ATPase complexes has been notoriously difficult. So far, there is no description of a fractionation approach, and the heterologous complex consisting of PAN from *Methanocaldococcus jannaschii* and the core particle from *Thermoplasma acidophilum*, an organism that lacks PAN (23), is the most widely studied archaeal model system and the only one that could be imaged to date (24).

Although the core particle degrades peptides and intrinsically disordered proteins in the absence of regulatory particles (5, 25), the Rpt subunits are essential for *Saccharomyces cerevisiae* (26), and Mpa is required by *Mycobacterium tuberculosis* in an infectious context (27), illustrating that the proteasomal ATPases perform a crucial function, in particular for the unfolding and degradation of (mis)-folded polypeptides under stress conditions. Nonetheless, we find that PAN ATPases are absent in a number of archaea (28), which raises the question of how substrate proteins are made available to the proteasome in these organisms.

Detecting a proteasome-interacting motif in the AAA ATPase CDC48 of *T. acidophilum* prompted us to perform a systematic analysis of archaeal AAA proteins, which uncovered a network of ATPases with a common HbY*X* motif including CDC48 and AMA proteins. For two model organisms, we provide evidence that these ATPases indeed physically interact with their cognate core particle and show that they stimulate proteasome activity in proteolytic assays, establishing CDC48 and AMA proteins as regulators of the proteasome in archaea.

EXPERIMENTAL PROCEDURES

Bioinformatics—Homologs of archaeal AAA proteins were identified with HHsenser (29) searching the non-redundant database of archaeal proteins (National Center for Biotechnology Information (NCBI), nr_arc) with the AAA+ module of AMA from *Methanosarcina mazei* (gi 21226406, Mm_$0304_{119-372}$). Assignment to orthologous groups of full-length sequences was based on cluster analyses using CLANS (30). *p* values for clustering were selected interactively to achieve formation of orthologous groups. Groups of AAA proteins were distinguished from other members of the AAA+ superfamily using different *p* value cutoffs and relying on our classification of AAA+ proteins (31). Members of orthologous groups were verified by testing for concordant domain composition with HHpred (32) and MUSCLE (33). The presence or absence of genes was mapped onto the archaeal species tree with iToL (34). C-terminal peptides comprising the last seven residues of AAA proteins were extracted from full-length sequences. Assignment of the HbY*X* motif (19) was based on the presence of a small or hydrophobic residue in third last and a Phe or Tyr residue in penultimate position.

Cloning—Ta20Sα (Ta1288, gi 16081896), Ta20Sβ (Ta0612, gi 16081708), TaCDC48ΔC (TaCDC48$_{1-733}$ lacking the last 12 residues), and TaCDC48-L745W (containing W466F, W541Y, and L745W mutations) genes were synthesized by GenScript. MmPAN-A (Mm1006, gi 20905437), MmPAN-B (Mm0789, gi 20905207), Mm20Sα (Mm2620, gi 21228722), Mm20Sβ (Mm0694, gi 21226796), and TaCDC48 (Ta0840, gi 16081896) were obtained as gifts from W. Baumeister and P. Zwickl. MmCDC48-A (Mm0248, gi 20904601), MmCDC48-B (Mm0447, gi 20904821), MmCDC48-C (Mm1256, gi 20905716), MmAMA (Mm0304, gi 20904664), and Mm0854 (gi 20905268) ORFs were amplified from genomic DNA of *M. mazei* strain OCM88 (ATCC number: BAA159) by PCR. GFPssrA fragment was PCR-amplified from pEGFP-N1 plasmid (Clontech) using a reverse primer containing the ssrA tag (AANDENYALAA) sequence. Proteasomal α-subunit DNA fragments were cloned into pET30b expression vector (Novagen); ATPases and GFPssrA were cloned as N-terminally hexahistidine-tagged proteins into pET28b; and proteasomal β-subunits were cloned as C-terminally hexahistidine-tagged proteins into pET22b.

Protein Production and Purification—Plasmids were transformed into *Escherichia coli* C41(DE3) RIL expression strain. Plasmids encoding the proteasomal α- and β-subunits were co-transformed to assemble the CP proteasome directly inside *E. coli* cells. Expression was achieved by growing single colonies in LB medium, supplemented with the appropriate antibiotics at 37 °C until an optical density of ~0.6 was reached followed by induction with 1 mM isopropyl-1-thio-β-D-galactopyranoside and continued culturing overnight at 20 °C. Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 10 mM imidazole, 5 mM β-mercaptoethanol, 0.4 mg/ml lysozyme, 40 μg/ml DNase, and except for the proteasome, EDTA-free protease inhibitor mixture (Roche Applied Science)) and lysed by French press treatment. After centrifugation, lysates were applied consecutively to nickel-nitrilotriacetic acid metal affinity (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 10 mM-1 M imidazole gradient, 5 mM β-mercaptoethanol), mono-Q anion-exchange (50 mM Tris-HCl, pH 7.5, 50 mM-1 M NaCl gradient, 10 mM $MgCl_2$, 5% glycerol, 1 mM dithiothreitol (DTT)), and Superose 6 gel filtration (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, and 1 mM DTT) chromatography columns (GE Healthcare). MmAMA accumulated in inclusion bodies and was purified under denaturing conditions using 4 M urea containing buffers. In parallel, the quality of protein preparations was controlled by electron microscopy, activity assays, and mass spectrometry.

Peptidase and Protease Assays—Proteolytic assays made use of three different fluorescent substrates, namely: quenched LFP (24) (Mca-AKVYPYPME-Dpa(Dnp)-amide; GenScript; excitation, 322 nm; emission, 398 nm) in the assays using 7-residue peptides and quenched BODIPY-TR X casein (35) (Invitrogen; excitation, 590 nm; emission, 645 nm) and GFPssrA (excitation, 488 nm; emission, 507 nm) in the assays using full-length ATPases. Peptides were synthesized by GenScript. LFP assays were performed with 0.6 nM CP, 250 μM 7-residue peptide, and 10 μM LFP in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, and 1 mM DTT, for the duration of 120 min at 59 °C for *T. acidophilum* proteins and 37 °C for *M. mazei* ones. Full-length protein assays were done with 100 nM CP, 400 nM ATPase, 1 μM substrate, 20

mM ATP, and 10 mM $MgCl_2$. All measurements are the average of three independent experiments and were carried out using a Synergy H4 plate reader (BioTek). For the tryptophan fluorescence shift assay, proteins were excited at 280 nm, and emission spectra of the proteins with or without ATPγS (Sigma-Aldrich) were collected from 330 nm to 380 nm.

Surface Plasmon Resonance—Surface Plasmon Resonance (SPR) experiments were performed at 25 °C with a Biacore 2000 instrument (Biacore) and a data collection rate of 1 Hz. MmCP was immobilized via amine coupling on a CM5 sensor chip (GE Healthcare). The surface of the upstream flow cell was subjected to the same coupling reaction in the absence of protein and was used as a reference. MmPAN-A, MmCDC48-B, and Mm0854 were supplemented with a 100-fold molar excess of ATPγS and then serially diluted into running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, and 0.005% (v/v) surfactant P20). Analyte samples with five different concentrations (1000 – 62.5 nM) were injected over the control and experimental surfaces at a flow rate of 10 μl/min. A blank sample, containing only the running buffer supplemented with ATPγS in the respective concentration, was injected under the same conditions to allow for double referencing. Data analysis was carried out with the BIAevaluation software (version 4.1.1).

Pulldowns—100 μg of bait proteins (Ta20S, TaCDC48, and TaCDC48ΔC) were bound to cobalt-nitrilotriacetic acid immobilized beads (GE Healthcare), mixed with *T. acidophilum* cellular lysate, and incubated for 1 h at 59 °C. The samples were then applied to Mobicol spin columns (MoBiTec) and washed three times with washing buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 40 mM imidazole, 5 mM β-mercaptoethanol). Bound proteins were eluted with elution buffer (50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 500 mM imidazole, 5 mM β-mercaptoethanol), and captured binding partners were identified with mass spectrometry.

Nano LC-MS/MS Analysis—Tryptic digestion of the samples and analyses of the peptides by LC-MS/MS and subsequent data processing were done as described previously with slight modifications (36). Analyses of the peptides were performed on a Proxeon Easy-nLC system (Proxeon Biosystems) coupled to an LTQ-Orbitrap-XL (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source (Proxeon Biosystems). The injected peptide mixtures were subsequently eluted with a 47-min segmented gradient of 2–80% HPLC solvent B (80% acetonitrile in 0.5% acetic acid) at a flow rate of 200 nl/min. Mass spectra were analyzed using the software suite MaxQuant, version 1.0.14.3 (37). The data were searched against a target-decoy *M. mazei* database containing 5645 forward protein sequences and 262 frequently observed contaminants. Trypsin was set as protease in which two missed cleavage sites were allowed. Acetylation at the N terminus and oxidation of methionine were set as variable modifications. Carbamidomethylation of cysteine was set as fixed modification. Initial precursor mass tolerance was set to 7 ppm at the precursor ion and 0.5 Da at the fragment ion level. False discovery rates were set to 1% at peptide and protein group level.

RESULTS

Prediction of Novel Proteasomal ATPases in the Archaeal Kingdom—AAA proteins form a distinct group within the AAA+ superfamily, whose unifying activity is the energy-dependent remodeling of proteins and nucleic acids (31, 38). In a cluster analysis of archaeal AAA proteins (supplemental Fig. S1), we obtained separate clusters for Vps4 (39), membrane-bound AAA (MBA) (40), an uncharacterized group exemplified by open reading frame number 854 in the *M. mazei* genome (Mm0854), AMA (41, 42), the proteasome-activating nucleotidase PAN (14), and CDC48 (43). Comparative analysis across 81 sequenced archaeal genomes showed that CDC48 is the sole subgroup found in all archaea and also has the highest degree of paralogy (Fig. 1*A*). The proteasome-activating nucleotidase PAN, however, is absent from major archaeal groups, including all Thermoplasmata, Thermoproteales, and the deep-branching Thaumarchaeota and Korarchaeota (supplemental Fig. S2).

The C-terminal Hb*YX* motif required for interaction with the 20S proteasome is not only present in the known proteasomal regulators of the PAN group, but also in all AMA proteins and in many CDC48 proteins. All archaeal organisms sequenced to date have at least one AAA protein with an Hb*YX* motif, supporting the notion that ATPase-mediated proteasomal regulation is an essential function (supplemental Fig. S3). Thus, organisms lacking PAN encode at least one CDC48 protein with a canonical Hb*YX* motif.

The number of putative proteasomal ATPases varies from one in Thermoplasmata to five in some Methanosarcinales (Fig. 1*B*). We therefore selected AAA proteins and cognate proteasome core particles of *T. acidophilum* and *M. mazei* to explore the idea that archaea employ multiple AAA proteins to regulate proteasomal activity. *T. acidophilum* encodes only one CDC48 protein and no PAN ortholog, and its proteasomal gatekeeper is still elusive despite the fact that the core particle (TaCP) has served as a long-standing model system together with the heterologous PAN protein of *M. jannaschii* (19–22, 24). In contrast, *M. mazei* contains seven AAA proteins, five of which have the proteasome-interacting motif, including two CDC48 proteins and a member of the AMA group (Fig. 1).

CDC48 Is the Proteasomal ATPase of T. acidophilum—Because the C-terminal peptides of PAN and Rpt ATPases have been shown to mimic efficiently the activation of the proteasome core particle by the full-length proteins (19), we used the C-terminal peptide of TaCDC48, whose last three residues are Leu-Tyr-Leu, to stimulate the hydrolysis of the fluorogenic nonapeptide LFP by TaCP. Although the proteasome can degrade LFP on its own, we observed a 2-fold higher fluorescence signal in the presence of C-terminal peptide as compared with TaCP alone (Fig. 2*A*), providing initial support for our hypothesis that CDC48 is the proteasomal gatekeeper in *T. acidophilum*.

In the next step, we tested the ability of full-length TaCDC48 to activate the *T. acidophilum* proteasome, using BODIPY-casein as a substrate. TaCP alone is able to degrade casein, and TaCDC48 could stimulate this activity 2.5-fold, as judged by the gain in fluorescence intensity of BODIPY fluorophores (Fig. 2*B*). This stimulation was entirely dependent on the Hb*YX* motif as a C-terminal deletion mutant of TaCDC48 lacking the

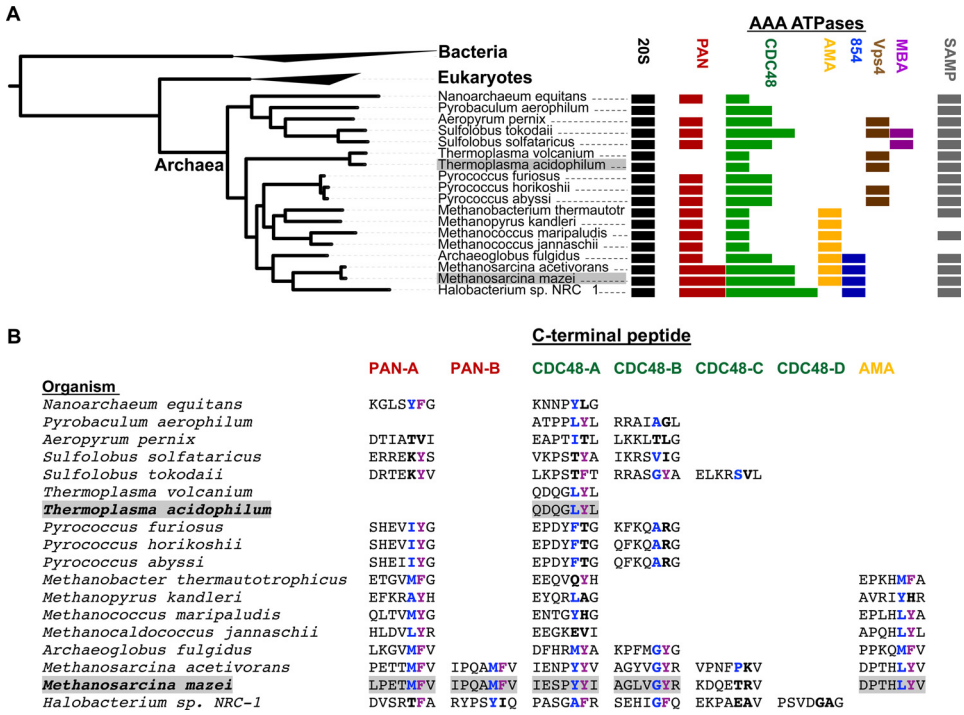


FIGURE 1. **Archaeal AAA proteins and the prediction of novel proteasome regulators.** *A*, the distribution of AAA ATPases is mapped onto a tree of selected archaeal species. Analysis of the C termini of archaeal AAA proteins indicates the presence of the HbY*X* motif in CDC48 and AMA, as well as in PAN proteins. CDC48 is ubiquitous and shows the highest degree of paralogy, whereas PAN is absent in Thermoproteales and Thermoplasmata and in deep-branching Korarchaeota and Thaumarchaeota, suggesting subgroup-specific loss rather than lateral gene transfer. AMA most likely originated in the ancestor of Archaeoglobales and methanogens, and was lost in Halobacteria. For comparison, the occurrence of the 20S proteasome and SAMP proteins is shown. *Panel A* was generated with iToL (34) (*MBA,* membrane-bound AAA). *B*, C-terminal sequences of PAN, CDC48, and AMA proteins. The species and sequences are shown in the same order as in *panel A* and illustrate the presence of the HbY*X* motif (hydrophobic/small residue in *blue*, Tyr/Phe in *magenta*). Species lacking PAN contain at least one CDC48 ortholog with the HbY*X* motif. All AMA proteins contain the HbY*X* motif. The number of putative proteasomal ATPases ranges from one to five, suggesting a regulatory network in many archaea. *T. acidophilum* and *M. mazei*, which were selected for experimental study, are *highlighted*.

last 12 residues (TaCDC48ΔC) was unable to activate the proteasome to any measurable extent (Fig. 2*B*).

Finally, we attempted to stimulate the degradation of GFPssrA using full-length TaCDC48. GFPssrA is an established model substrate for PAN and proteasome from *M. jannaschii,* as well as for Mpa and proteasome from the actinobacterium *M. tuberculosis* (44, 45). However, we did not detect its degradation by the TaCDC48-TaCP system (Fig. 2*C*). It has previously been reported that GFPssrA is unfolded by TaCDC48 at high magnesium concentrations and can then be degraded by a highly active, gateless proteasome mutant, in which the 12 N-terminal residues of α-subunits have been removed (46). We attribute the need for a gateless proteasome in that experiment to the fact that the TaCDC48 construct contained a C-terminal His tag and could not activate the proteasome via its C-terminal HbY*X* motif. We propose that this uncoupling and the use of a constitutively active proteasome accounts for the different outcome of the experiments. As we will show in the next section, the ability to degrade GFPssrA appears to be limited to a subset of active ATPase-proteasome tandems.

To probe the physical interaction of TaCDC48 and TaCP, we changed the last residue of the ATPase from leucine to a tryptophan in a tryptophan-free mutant background of TaCDC48, TaCDC48-L745W, in analogy to an experiment conducted by Goldberg and colleagues (19). In fluorescence emission spectra, upon the addition of ATPγS and in the presence of tryptophan-free TaCP, the C-terminal tryptophan of TaCDC48-L745W experiences a blue shift and a decrease in fluorescence intensity (Fig. 2*D*). This is the expected outcome if binding of the exposed C-terminal residues into the binding pocket of the proteasomal α-subunit moved the peptide tail from an aqueous to a more hydrophobic environment. To collect further evidence for an interaction of TaCDC48 and TaCP, we performed pull-down experiments with either TaCDC48 or TaCP as bait in cellular lysate of *T. acidophilum*. Both CDC48 and TaCP return each other as one of the most significant binding partners, but the C-terminal deletion mutant TaCDC48ΔC does not retrieve TaCP (Table 1, supplemental Table 1). In summary, we find that TaCDC48 directly interacts with and stimulates TaCP and that its C-terminal residues are crucial for this activity.

A Network of Five AAA Proteins Regulates the Proteasome of M. mazei—In contrast to *T. acidophilum*, *M. mazei* encodes seven AAA proteins, of which we predict five to function as proteasomal ATPases. This is the largest number of ATPases in the putative regulatory network among sequenced archaea. To collect evidence for such a network, we tested whether the 7-residue peptides corresponding to the C termini of the ATPases could stimulate LFP peptidase activity of the *M. mazei* 20S proteasome (MmCP). In agreement with our prediction, we found that peptides from the five AAA proteins that have the HbY*X* motif increased the hydrolysis of the fluorogenic LFP. These are the two PAN proteins (35), two out of three CDC48 proteins, and the AMA protein (Fig. 3*A*). The C-terminal peptides of CDC48-C and of Mm0854, which lack the motif,

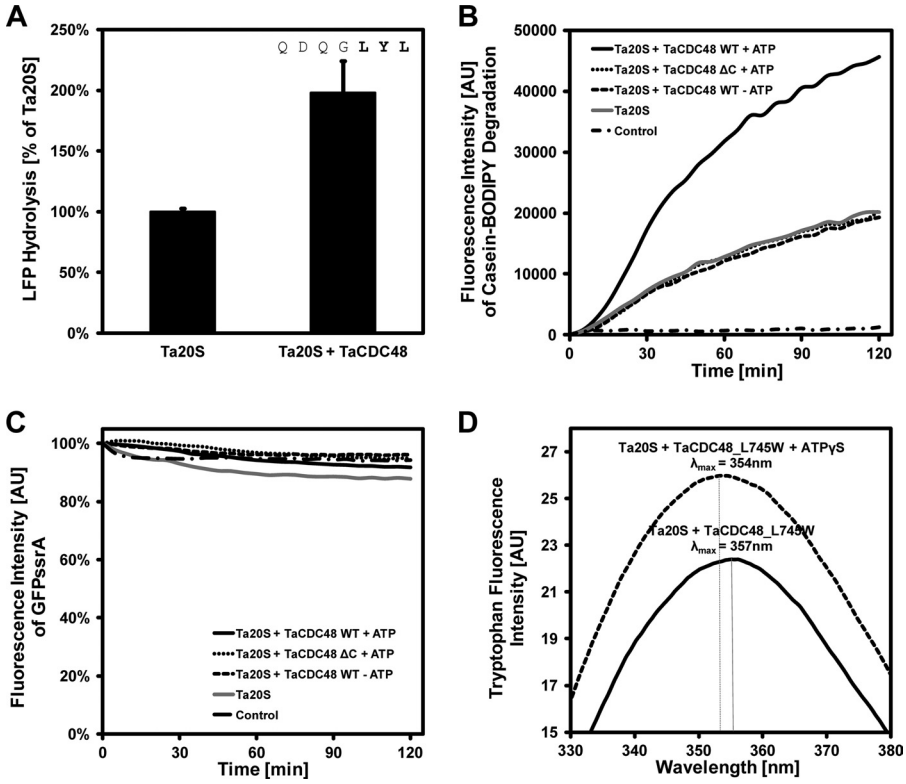


FIGURE 2. **CDC48 binds to and activates the proteasome of *T. acidophilum*.** *A*, the addition of the C-terminal peptide of TaCDC48 results in a 2-fold increase in proteasomal degradation of the fluorogenic nonapeptide LFP, measured as gain in fluorescence intensity. *Error bars* indicate S.D. *B*, full-length TaCDC48 enhances the degradation of BODIPY-labeled casein by TaCP in the presence of ATP. A C-terminal deletion mutant of TaCDC48 (TaCDC48ΔC) lacking the last 12 residues of TaCDC48 does not stimulate casein degradation in comparison with TaCP alone. *C*, eGFP-ssrA is not degraded by TaCP in the presence of TaCDC48 and ATP. *AU*, absorbance units. *D*, in the presence of TaCP, the fluorescence emission spectrum of a C-terminal tryptophan variant of TaCDC48 (TaCDC48-L745W) moves from an aqueous to a more hydrophobic environment upon the addition of ATPγS, indicated by a blue shift of the emission maximum and a decreased intensity.

TABLE 1

TaCDC48 interacts with TaCP

Table 1 summarizes the identification of the α- and β-subunits of TaCP and TaCDC48 in pulldown experiments using three different bait proteins (TaCP, TaCDC48, TaCDC48ΔC) in cellular lysate of *T. acidophilum*. The complete list of identified proteins is shown in supplemental Table 1.

	Bait proteins[a]			
	Control	TaCP	TaCDC48	TaCDC48ΔC
TaCDC48	0	13	117	120
α-Subunit	10	46	17	8
β-Subunit	2	20	8	2

[a] Number represents the number of identified peptides.

had no effect. The degree of stimulation ranged from 1.5-fold for CDC48-B to ~9-fold for AMA. We also found that a PAN-A variant, in which the naturally occurring Phe of the HbY*X* motif was mutated to Tyr, was a much stronger activator than wild type, in line with results showing that the hydroxyl group of Tyr, which is the most frequent residue in this position, stabilizes the open gate conformation more efficiently (21).

Next, we extended our analysis to the seven full-length ATPases, using BODIPY-casein as a substrate in protease assays. The five ATPases with HbY*X* motifs stimulated proteasome activity from 2-fold for CDC48-A to 2.5-fold for PAN-A (Fig. 3*B*). Again, CDC48-C and Mm0854, which do not contain the HbY*X* motif, did not show any effect. Our results show that full-length proteins stimulate the proteasome, albeit not as strongly as isolated peptides. Although peptides are able to bind independently and in maximal number, in full-length proteins, only a subset of C-terminal tails can be expected to bind at a given time due to steric constraints (symmetry mismatch).

Remarkably, of the five ATPases able to activate the proteasome, only PAN-A was able to stimulate the degradation of GFPssrA (Fig. 3*C*). This suggests that the ATPases may also have differing substrate specificities in a cellular context, possibly mediated in part by their different N-terminal substrate recognition domains. Nevertheless, it is worth noting that differences are apparent even in orthologous ATPases with highly similar N-domains, as seen by the inability of PAN-B to mediate GFPssrA degradation.

To probe the physical interaction between the ATPases and MmCP, we selected three for analysis by surface plasmon resonance (SPR), based on the degree of stimulation and ease of handling (PAN-A, CDC48-B, and Mm0854). Mm0854 showed only a background level of binding to MmCP, which correlates with the absence of an HbY*X* motif in this protein and its inability to stimulate MmCP in degradation assays. PAN-A and CDC48-B, in contrast, showed a clear concentration-dependent response, providing evidence that a direct interaction of ATPase and CP is responsible for their ability to activate the proteasome (Fig. 3*D*).

DISCUSSION

We have defined a regulatory network of proteasomal ATPases across the archaeal kingdom and validated their interactions with the proteasome through *in vitro* studies with pro-

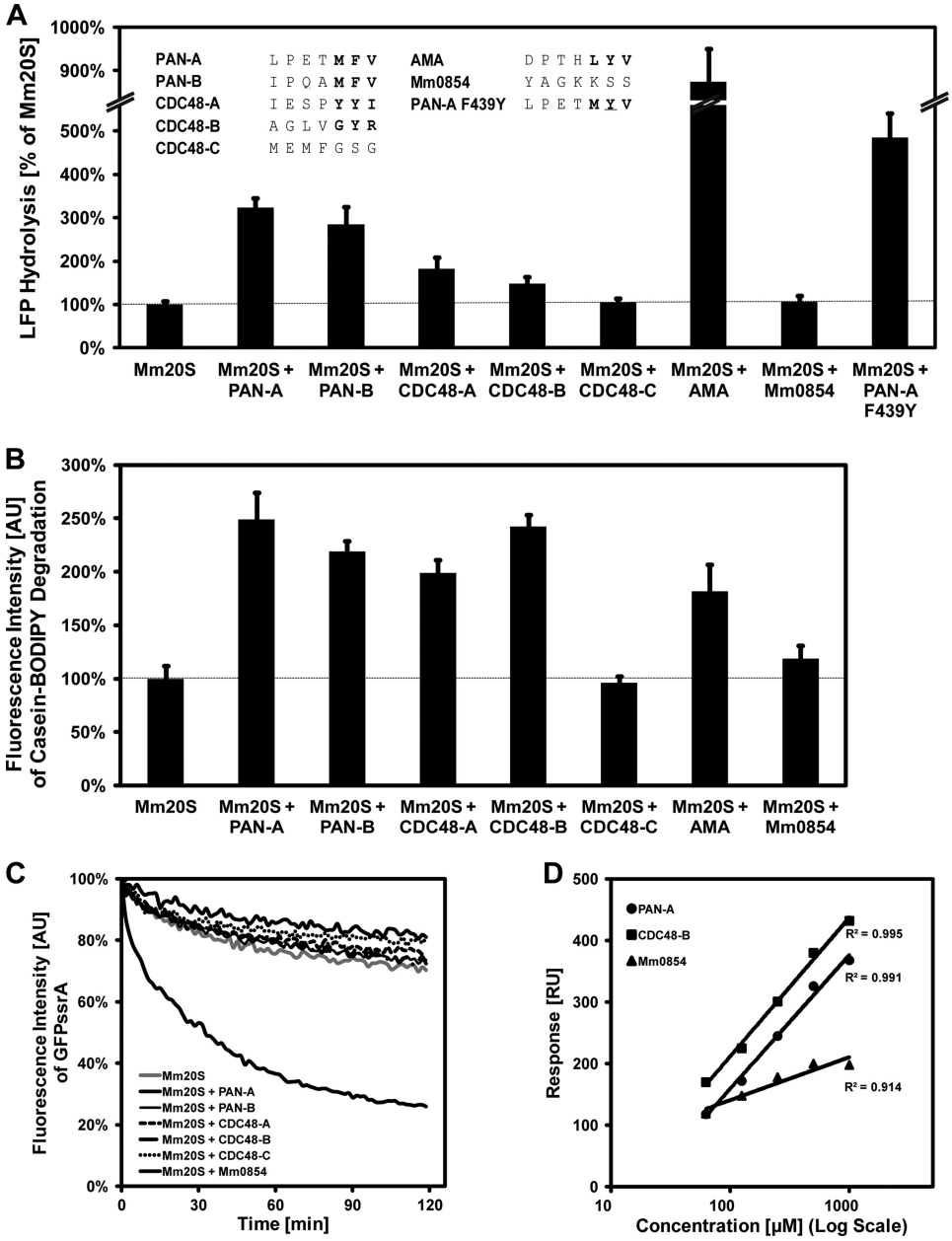


FIGURE 3. **A network of AAA ATPases regulates the proteasome of *M. mazei*.** *A*, C-terminal peptides of AAA proteins from *M. mazei* containing the HbY*X* motif enhance LFP hydrolysis. The PAN-A F439Y mutant is a stronger activator of MmCP than PAN-A with the wild-type HbY*X* motif. *Error bars* indicate S.D. *B*, full-length ATPases with the HbY*X* motif stimulate degradation of BODIPY-labeled casein by MmCP. *Error bars* indicate S.D. *AU*, absorbance units. *C*, PAN-A is the only AAA protein of *M. mazei* that mediates degradation of eGFP-ssrA by MmCP. *D*, SPR sensorgrams show that MmPAN-A and MmCDC48-B physically interact with MmCP in a concentration-dependent manner, whereas Mm0854 shows a much weaker interaction. *RU*, response units.

teins from two archaeal organisms that mark extremes in the number of genomically encoded AAA ATPases. This network features, in addition to known proteasome-activating nucleotidases of the PAN type, two groups of AAA proteins, AMA and CDC48, that hitherto have not been seen directly associated with the 20S proteasome (Fig. 4).

Archaea-wide Significance of the Network Model—The interaction of AAA ATPases with the proteasome is determined by the presence of HbY*X* motifs at the very C terminus of ATPase subunits. However, not only AAA ATPases contain this motif. We find the HbY*X* motif (pattern: [LVIMGAPFYW][FY]x>) at the C terminus of 63 proteins (of 1484 proteins coding genes) in the *T. acidophilum* genome, one of which is CDC48. For *M. mazei*, we detect this motif in 133 (of 3368) proteins, five of which are the AAA proteins that are part of the regulatory network we describe. We assume that only very few of these proteins form ring-shaped complexes and offer this interaction motif in a sufficiently high local concentration and extended conformation to stimulate gate opening of the core particle. Furthermore, there are other interaction partners of the proteasome with this motif, for instance, the proteasome assembly chaperones PbaA and PbaB, which are also conserved in archaea (47), and there might be more remaining to be found. Nevertheless, our analysis specifically and comprehensively tracks this motif in ATP-dependent proteasome regulators of the family of AAA unfoldases.

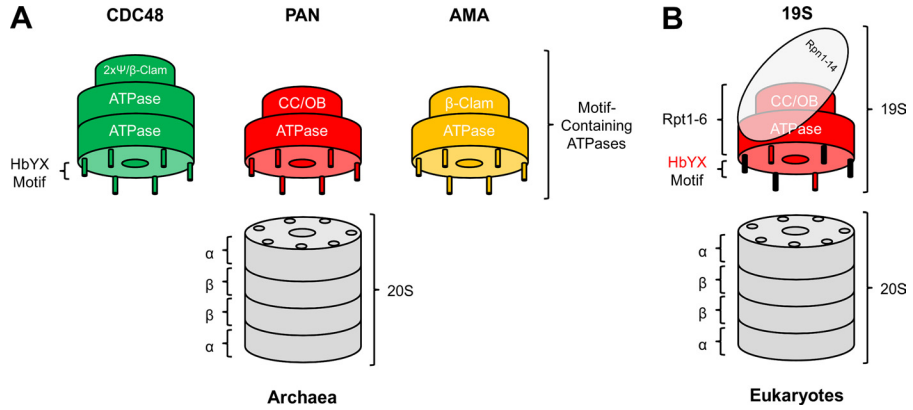


FIGURE 4. **The network of proteasomal AAA ATPases.** *A*, in archaea, PAN, CDC48, and AMA proteins constitute a regulatory network of proteasomal ATPases. Interaction is governed by the C-terminal HbY*X* motif; different N-domain folds in the ATPases are denoted. *2xpsi,* double-psi barrel; *CC/OB*, coiled-coil and oligosaccaride-binding domain. *B*, in eukaryotes, the heterooligomeric PAN-like Rpt1–6 subunits function within the 19S particle as the regulatory ATPase of the proteasome. The HbY*X* motifs of Rpt2 and Rpt5 are important for gate opening.

The conservation of residues forming the binding pocket for the HbY*X* motif (invariant residues proline 17 and glycine 19) in the proteasome α-subunits indicate that all archaeal proteasomes are most likely activated by the same mechanism. Although we observe peculiar or highly divergent C termini that do not allow an unambiguous decision regarding the presence of the motif, all archaea encode at least one AAA protein that contains the HbY*X* motif in canonical form, emphasizing the importance of proteasomal regulation by AAA proteins. Even if PAN orthologs are absent in a genome, there is at least one member of the CDC48 family with a clear HbY*X* motif, defining these members as proteasomal ATPases in the context of our degradation and interaction assays. Likewise, as all members of the AMA group contain the HbY*X* motif, we conclude that gatekeeping of the proteasome is the main cellular function of this group, thereby increasing the repertoire of proteasome regulators.

Functional Implications—Components of the archaeal proteasome machinery generally improve survival rate under stress conditions (48) and are essential for growth (49). Functional diversification of the core particle through gene duplication is known for *Haloferax volcanii*, in which a second type of α-subunit (and PAN-B) is up-regulated during starvation (50), and for *Pyrococcus furiosus*, where the β1-subunit is up-regulated under heat stress, resulting in more thermostable PfCP particles (51). Accordingly, for the dynamic proteome of a cell living in changing environments, an adaptable network of regulators of the central protein destruction machinery conceivably confers advantages. The general capabilities of the proteasomal system for targeted degradation would be increased through the participation of different N-terminal substrate recognition domains of the regulatory ATPases. As a consequence, a greater variety of substrates could be recognized with higher affinity and specificity. In addition to a more sophisticated regulation of proteolysis, the presence of multiple proteasomal ATPases may also generate some redundancy. *H. volcanii*, for example, tolerates a gene knock-out of both of its PAN proteins under normal growth conditions (6), potentially because it still has two CDC48 proteins with an HbY*X* motif.

Currently, it is unclear how this regulatory network of ATPases is integrated with the SAMP-based tagging system, whose conjugates are increased (SAMP1) or decreased (SAMP2) in a genetic background depleted of the PAN proteins (6). It remains to be studied whether motif-containing CDC48 proteins participate in the SAMPylation system and whether they are able to compensate for a deletion of PAN proteins. Importantly, there is no correlation between the presence of the SAMP proteins and specific proteasomal ATPases (Fig. 1*A*).

Evolutionary Implications—Comparative analysis of the content of putative proteasomal ATPases in archaeal genomes suggests a complex ancestral state, which is likely to include two CDC48 proteins and one PAN protein. The pattern of conservation suggests that PAN has been lost in certain lineages rather than laterally transferred (23), whereas AMA most likely originated in a common ancestor of methanogens and Archaeoglobales. The number of putative proteasomal ATPases varies in extant archaea between one and five, pointing to clade-specific solutions for the functional task of targeted protein degradation. Nevertheless, the approximate number of three proteasomal ATPases in the ancestor of all archaea (one PAN, two CDC48 proteins) suggests that it would have already relied on a network of AAA proteins to regulate its central proteolytic machinery.

The employment of different AAA proteins as regulators of the proteasome in the majority of archaea contrasts with the situation in eukaryotes, where the heterooligomeric PAN orthologs Rpt1–6 act exclusively as ATPase regulators (the regulators Blm10 and 11S are not ATPases). Although AMA proteins are not found in eukaryotes, CDC48 (p97) is conserved and fulfills functions in a plethora of pathways. However, there is only evidence for a direct interaction of CDC48 with the 19S regulatory particle, rather than the 20S core particle (52), serving as a gearbox that segregates ubiquitylated from unmodified substrate proteins (53). Therefore, the fully differentiated 26S proteasome is, with the notable exception of immunological responses in vertebrates, basically fixed in its composition. Diversity in the ubiquitin-proteasome system of eukaryotes mainly takes place at the level of E3 ligases, which confer specificity in the tagging of substrates. Their radiating number, ranging from ~80 in yeast to ~600 in humans (54), enables specific involvement of the proteasome in the regulation of numerous cellular processes. Despite the fact that E3-like

ligases for SAMPylation in archaea have not been discovered yet, a network of regulatory ATPases constitutes an additional strategy to expand the capacity of proteasomal protein degradation.

Acknowledgments—We thank Heinz Schwarz for electron microscopy, Peter Zwickl and Wolfgang Baumeister for plasmids, Thilo Stehle for support with SPR, Boris Macek for support with MS, and Volkmar Braun, Jörn Engelmann, and Remco Sprangers for helpful discussions.

REFERENCES

1. Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92,** 367–380
2. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins. *Science* **286,** 1888–1893
3. Pagano, M., Tam, S. W., Theodoras, A. M., Beer-Romero, P., Del Sal, G., Chau, V., Yew, P. R., Draetta, G. F., and Rolfe, M. (1995) Role of the ubiquitin-proteasome pathway in regulating abundance of the cyclin-dependent kinase inhibitor p27. *Science* **269,** 682–685
4. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78,** 761–771
5. Kisselev, A. F., Akopian, T. N., and Goldberg, A. L. (1998) Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J. Biol. Chem.* **273,** 1982–1989
6. Humbard, M. A., Miranda, H. V., Lim, J. M., Krause, D. J., Pritz, J. R., Zhou, G., Chen, S., Wells, L., and Maupin-Furlow, J. A. (2010) Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463,** 54–60
7. Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P., and Darwin, K. H. (2008) Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* **322,** 1104–1107
8. Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* **67,** 425–479
9. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Ä resolution. *Science* **268,** 533–539
10. Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., and Baumeister, W. (1995) Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science* **268,** 579–582
11. Seemuller, E., Lupas, A., and Baumeister, W. (1996) Autocatalytic processing of the 20S proteasome. *Nature* **382,** 468–471
12. Rabl, J., Smith, D. M., Yu, Y., Chang, S. C., Goldberg, A. L., and Cheng, Y. (2008) Mechanism of gate opening in the 20S proteasome by the proteasomal atpases. *Mol. Cell* **30,** 360–368
13. Rubin, D. M., Glickman, M. H., Larsen, C. N., Dhruvakumar, S., and Finley, D. (1998) Active site mutants in the six regulatory particle ATPases reveal multiple roles for ATP in the proteasome. *EMBO J.* **17,** 4909–4919
14. Zwickl, P., Ng, D., Woo, K. M., Klenk, H. P., and Goldberg, A. L. (1999) An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes. *J. Biol. Chem.* **274,** 26008–26014
15. Wolf, S., Nagy, I., Lupas, A., Pfeifer, G., Cejka, Z., Müller, S. A., Engel, A., De Mot, R., and Baumeister, W. (1998) Characterization of ARC, a divergent member of the AAA ATPase family from *Rhodococcus erythropolis*. *J. Mol. Biol.* **277,** 13–25
16. Wang, T., Darwin, K. H., and Li, H. (2010) Binding-induced folding of prokaryotic ubiquitin-like protein on the *Mycobacterium* proteasomal ATPase targets substrates for degradation. *Nat. Struct. Mol. Biol.* **17,** 1352–1357
17. Imkamp, F., Striebel, F., Sutter, M., Ozcelik, D., Zimmermann, N., Sander, P., and Weber-Ban, E. (2010) Dop functions as a depupylase in the prokaryotic ubiquitin-like modification pathway. *EMBO Rep.* **11,** 791–797
18. Lander, G. C., Estrin, E., Matyskiela, M. E., Bashore, C., Nogales, E., and Martin, A. (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* **482,** 186–191
19. Smith, D. M., Chang, S. C., Park, S., Finley, D., Cheng, Y., and Goldberg, A. L. (2007) Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's α ring opens the gate for substrate entry. *Mol. Cell* **27,** 731–744
20. Religa, T. L., Sprangers, R., and Kay, L. E. (2010) Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* **328,** 98–102
21. Stadtmueller, B. M., Ferrell, K., Whitby, F. G., Heroux, A., Robinson, H., Myszka, D. G., and Hill, C. P. (2010) Structural models for interactions between the 20S proteasome and its PAN/19S activators. *J. Biol. Chem.* **285,** 13–17
22. Yu, Y., Smith, D. M., Kim, H. M., Rodriguez, V., Goldberg, A. L., and Cheng, Y. (2010) Interactions of PAN's C-termini with archaeal 20S proteasome and implications for the eukaryotic proteasome-ATPase interactions. *EMBO J.* **29,** 692–702
23. Ruepp, A., Rockel, B., Gutsche, I., Baumeister, W., and Lupas, A. N. (2001) The chaperones of the archaeon *Thermoplasma acidophilum*. *J. Struct. Biol.* **135,** 126–138
24. Smith, D. M., Kafri, G., Cheng, Y., Ng, D., Walz, T., and Goldberg, A. L. (2005) ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. *Mol. Cell* **20,** 687–698
25. Liu, C. W., Corboy, M. J., DeMartino, G. N., and Thomas, P. J. (2003) Endoproteolytic activity of the proteasome. *Science* **299,** 408–411
26. Bajorek, M., Finley, D., and Glickman, M. H. (2003) Proteasome disassembly and downregulation is correlated with viability during stationary phase. *Curr. Biol.* **13,** 1140–1144
27. Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N., and Nathan, C. F. (2003) The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* **302,** 1963–1966
28. Maupin-Furlow, J. (2012) Proteasomes and protein conjugation across domains of life. *Nat. Rev. Microbiol.* **10,** 100–111
29. Söding, J., Remmert, M., Biegert, A., and Lupas, A. N. (2006) HHsenser: exhaustive transitive profile search using HMM-HMM comparison. *Nucleic Acids Res.* **34,** W374–378
30. Frickey, T., and Lupas, A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20,** 3702–3704
31. Ammelburg, M., Frickey, T., and Lupas, A. N. (2006) Classification of AAA+ proteins. *J. Struct. Biol.* **156,** 2–11
32. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33,** W244–248
33. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797
34. Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* **39,** W475–478
35. Medalia, N., Sharon, M., Martinez-Arias, R., Mihalache, O., Robinson, C. V., Medalia, O., and Zwickl, P. (2006) Functional and structural characterization of the *Methanosarcina mazei* proteasome and PAN complexes. *J. Struct. Biol.* **156,** 84–92
36. Borchert, N., Dieterich, C., Krug, K., Schütz, W., Jung, S., Nordheim, A., Sommer, R. J., and Macek, B. (2010) Proteogenomics of *Pristionchus pacificus* reveals distinct proteome structure of nematode models. *Genome Res.* **20,** 837–846
37. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **4,** 698–705
38. Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9,** 27–43
39. Makarova, K. S., Yutin, N., Bell, S. D., and Koonin, E. V. (2010) Evolution of diverse cell division and vesicle formation systems in *Archaea*. *Nat. Rev.*

Microbiol. **8,** 731–741
40. Serek-Heuberger, J., Hobel, C. F., Dunin-Horkawicz, S., Rockel, B., Martin, J., and Lupas, A. N. (2009) Two unique membrane-bound AAA proteins from *Sulfolobus solfataricus*. *Biochem. Soc. Trans.* **37,** 118–122
41. Djuranovic, S., Rockel, B., Lupas, A. N., and Martin, J. (2006) Characterization of AMA, a new AAA protein from *Archaeoglobus* and methanogenic archaea. *J. Struct. Biol.* **156,** 130–138
42. Summer, H., Bruderer, R., and Weber-Ban, E. (2006) Characterization of a new AAA+ protein from archaea. *J. Struct. Biol.* **156,** 120–129
43. Pamnani, V., Tamura, T., Lupas, A., Peters, J., Cejka, Z., Ashraf, W., and Baumeister, W. (1997) Cloning, sequencing, and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum*. *FEBS Lett.* **404,** 263–268
44. Striebel, F., Hunkeler, M., Summer, H., and Weber-Ban, E. (2010) The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *EMBO J.* **29,** 1262–1271
45. Zhang, F., Wu, Z., Zhang, P., Tian, G., Finley, D., and Shi, Y. (2009) Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol. Cell* **34,** 485–496
46. Gerega, A., Rockel, B., Peters, J., Tamura, T., Baumeister, W., and Zwickl, P. (2005) VAT, the thermoplasma homolog of mammalian p97/VCP, is an N domain-regulated protein unfoldase. *J. Biol. Chem.* **280,** 42856–42862
47. Kusmierczyk, A. R., Kunjappu, M. J., Kim, R. Y., and Hochstrasser, M. (2011) A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nat. Struct. Mol. Biol.* **18,** 622–629
48. Ruepp, A., Eckerskorn, C., Bogyo, M., and Baumeister, W. (1998) Proteasome function is dispensable under normal but not under heat shock conditions in *Thermoplasma acidophilum*. *FEBS Lett.* **425,** 87–90
49. Zhou, G., Kowalczyk, D., Humbard, M. A., Rohatgi, S., and Maupin-Furlow, J. A. (2008) Proteasomal components required for cell growth and stress responses in the haloarchaeon *Haloferax volcanii*. *J. Bacteriol.* **190,** 8096–8105
50. Reuter, C. J., Kaczowka, S. J., and Maupin-Furlow, J. A. (2004) Differential regulation of the PanA and PanB proteasome-activating nucleotidase and 20S proteasomal proteins of the haloarchaeon *Haloferax volcanii*. *J. Bacteriol.* **186,** 7763–7772
51. Madding, L. S., Michel, J. K., Shockley, K. R., Conners, S. B., Epting, K. L., Johnson, M. R., and Kelly, R. M. (2007) Role of the β1 subunit in the function and stability of the 20S proteasome in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **189,** 583–590
52. Isakov, E., and Stanhill, A. (2011) Stalled proteasomes are directly relieved by P97 recruitment. *J. Biol. Chem.* **286,** 30274–30283
53. Jentsch, S., and Rumpf, S. (2007) Cdc48 (p97): a "molecular gearbox" in the ubiquitin pathway? *Trends Biochem. Sci.* **32,** 6–11
54. Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S., and Joazeiro, C. A. (2008) Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3,** e1487