Rules and Exceptions: Sialic Acid Variants and Their Role in Determining Viral Tropism

Thilo Stehle,[a,b] Zaigham M. Khan[a]

Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany[a]; Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[b]

Sialic acids decorate the surfaces of most mammalian cells and are used by many viruses as attachment receptors. In contrast to other mammals, humans cannot synthesize a version of sialic acid known as *N*-glycolyl neuraminic acid. This difference is exploited by some viruses to establish tropism. Here we compare recently determined structures of closely related animal and human polyomaviruses and examine their strategies for engaging specific sialic acid variants.

SIALIC ACID VARIANTS AND THEIR ROLE IN VIRAL ATTACHMENT

To initiate an infectious cycle, a virus must first attach to one or several receptors, or attachment factors, at the host cell surface. In addition to adhering the virus to its target cell, these initial interactions often serve as crucial determinants of host range and tissue tropism. An increasing number of viruses, including many human pathogens, are being found to engage cell surface glycan receptors (reviewed in reference 1). Recent advances in glycan microarray technology (2) and structural analysis techniques (3) have helped to identify glycan receptors and have enabled a better understanding of how these glycans interact with viruses. However, the interrelationships between host-specific glycan structures and viral host and cell tropism remain poorly understood, in part due to lack of knowledge about the expression profiles, distributions, densities, and anchoring of glycan structures on cell surfaces.

Many viruses bind glycans terminating in sialic acids, acidic sugars with a nine-carbon backbone that decorate all cell surfaces and most secreted proteins of vertebrates and "higher" invertebrates (1). While the most common sialic acid in humans is 5-*N*-acetyl neuraminic acid (Neu5Ac), various modifications (acetylation, methylation, and sulfation) yield more than 50 different sialic acid variants. Of particular importance for host and tissue tropism is 5-*N*-glycolyl neuraminic acid (Neu5Gc), a sialic acid variant in which the *N*-acetyl group carries an additional hydroxyl group attached to the methyl carbon (Fig. 1A and B). Unlike other mammals, humans cannot synthesize Neu5Gc due to a gene defect that inactivates the enzyme that hydroxylates Neu5Ac, CMP *N*-acetyl neuraminic acid hydroxylase (CMAH). This defect leads to differences in many sialic acid-mediated biological processes in humans compared with those in other species (4). Furthermore, the absence of Neu5Gc, and the resulting excess of Neu5Ac, is thought to affect the interactions of various pathogens with human tissues by conferring resistance to Neu5Gc-binding pathogens and sensitivity to Neu5Ac-binding pathogens (4).

The role of Neu5Gc in receptor engagement and in defining viral tropism is only beginning to emerge, and structural analyses defining the differences in specificity for Neu5Ac and Neu5Gc are limited to only a few examples of viral (5, 6) and bacterial (7) pathogens. Recent structure-function analyses of human and animal polyomaviruses illustrate the complicated role of Neu5Gc as a component of receptors. Polyomaviruses constitute a rapidly expanding family of small, nonenveloped double-stranded DNA (dsDNA) viruses (8). The major capsid protein of these viruses, VP1, is organized into 72 pentamers on the virion surface. Each pentamer possesses five equivalent binding sites on the outermost rim of VP1 that can engage glycan receptors terminating in sialic acid (1) (Fig. 1C). Below, we compare the abilities of two pairs of closely related human and monkey polyomaviruses to interact with Neu5Ac- and Neu5Gc-based receptors (Fig. 2).

SIMIAN VIRUS 40 AND HUMAN BK POLYOMAVIRUS

Simian virus 40 (SV40) uses the GM1 ganglioside, a branched glycan carrying α2,3-linked sialic acid on one of its two branches, as its cellular receptor (9). Glycan microarray screening revealed that the GM1 ganglioside terminating in Neu5Gc (GM1-Gc) is a much better ligand for SV40 than its Neu5Ac-carrying counterpart (GM1-Ac) (10), in line with the presence of Neu5Gc in simian species. While structural data are available only for SV40 VP1 pentamers bound to GM1-Ac (11), they nevertheless provide a framework for understanding the enhanced affinity of the virus for GM1-Gc. SV40 engages the Neu5Ac moiety of GM1-Ac in a shallow pocket, with its *N*-acetyl group facing into a large cavity lined with polar and hydrophobic residues. The cavity is much larger than would be required to accommodate the methyl group of the *N*-acetyl chain, and it could easily accept the hydroxymethyl group of an *N*-glycolyl chain (11). As shown in Fig. 2A, the predicted conformations of the CH_2-OH group of Neu5Gc would result in additional water-mediated hydrogen bonds with surrounding polar side chains of residues R169 and Q187 as well as the backbone carbonyl of H61, thus providing a structural basis for the observed stronger interaction between SV40 and GM1-Gc than between SV40 and GM1-Ac.

Phylogenetically, SV40 is especially closely related to the human polyomaviruses BK polyomavirus (BKPyV) and JC polyomavirus (JCPyV) (8). Both BKPyV and JCPyV are widely distrib-
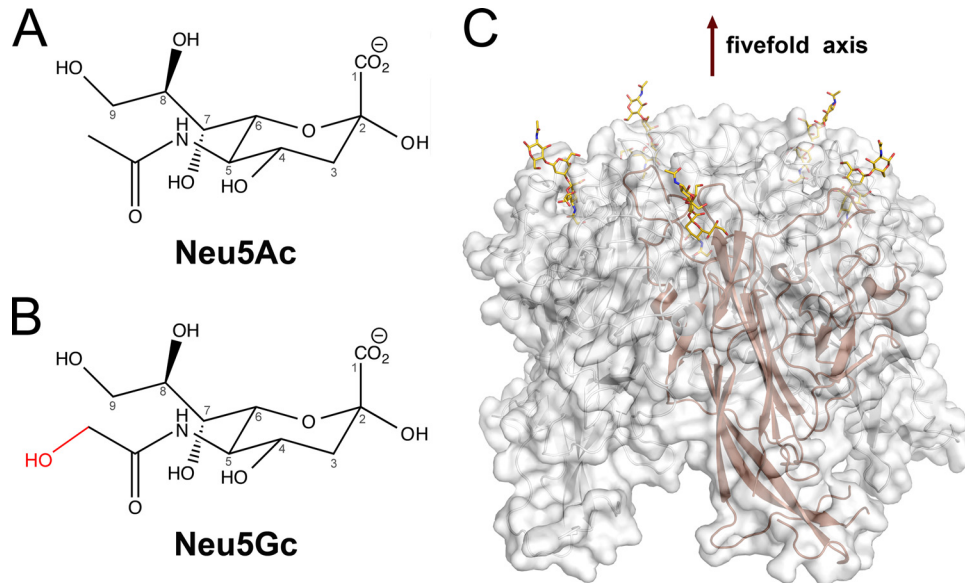
FIG 1 Sialic acid variants engaged by polyomaviruses. (A and B) Chemical structures of two prominent sialic acids, Neu5Ac and Neu5Gc. Numbering of the sialic acid structures begins at the carboxylate carbon and continues along the chain. The extra hydroxyl group that is present in Neu5Gc but absent in Neu5Ac is shown in red. Sialic acids can be attached to many other monosaccharides via α-glycosidic linkages at the C2 and C8 positions. (C) Structure of HPyV9 VP1 pentamer in complex with five sialyloligosaccharide receptors (6). The pentamer is shown in cartoon and surface representation. Sialyloligosaccharide is drawn as a stick model, and colored gold (carbons), blue (nitrogens), and red (oxygens). This structure is shown as an example. All structurally known sialic acid-binding polyomaviruses engage terminal sialic acid at a similar location.

uted in the human population but cause disease only in immunocompromised individuals. Structural analyses of JCPyV and BKPyV VP1 pentamers in complex with sialylated receptors have shown that both viruses, like SV40, engage Neu5Ac in a highly conserved location and with nearly identical contacts (1, 12). In fact, key contacts between protein residues and Neu5Ac atoms are conserved in all three viruses (compare Fig. 4A through C in reference 12). The only substantial difference between SV40 and the two human viruses is the absence of the large cavity in BKPyV (Fig. 2B) and also in JCPyV (12). Furthermore, the large hydrophobic side chain of F75 (whose orientation is conserved in BKPyV and JCPyV) would clash with the glycolyl hydroxyl group. Therefore, neither BKPyV nor JCPyV could accommodate glycan receptors terminating in Neu5Gc. The inability of BKPyV to recognize Neu5Gc is also supported by experiments in which a single-point mutation of a residue outside the sialic acid-binding region switched the specificity of BKPyV from α2,8-α2,3-di-Neu5Ac b-series gangliosides to the SV40 receptor GM1. The mutated virus grows well in human cells carrying GM1-Ac but does not grow in simian cells, which carry GM1-Gc glycans (12). Thus, the receptor specificities of SV40 and BKPyV correlate well with the biosynthetically available variants of sialic acids in their respective hosts.

AFRICAN GREEN MONKEY VIRUS AND HUMAN POLYOMAVIRUS 9

Another interesting pair of monkey and human polyomaviruses is formed by the African green monkey polyomavirus (AGMPyV) and its close relative, human polyomavirus 9 (HPyV9). AGMPyV was originally isolated from African green monkey lymph node cultures (13) and attracted interest because of its narrow tropism for a human B-lymphoblastoid tumor cell line. The virus is therefore also known as B-lymphotropic polyomavirus (LPyV). While

sialic acid had long ago been established as a crucial component of the AGMPyV receptor (14), the structural basis of this interaction was unveiled only very recently (15). The AGMPyV VP1 protein harbors a deeply recessed binding site that essentially buries the sialic acid (Fig. 2C). Although crystal soaking experiments were conducted only with Neu5Ac-bearing glycans, the architecture of the binding site suggested that AGMPyV VP1 can accommodate α2,3-linked sialyloligosaccharides terminating in either Neu5Ac or Neu5Gc (15). This was later confirmed by glycan microarray analysis (6).

HPyV9 was identified in 2011 and detected in human serum, plasma, urine, and skin (16, 17). Although there are no data yet on the pathogenicity of HPyV9 in the human population, the seroprevalence of the virus is 47% in adults (18). With 87% identical amino acid residues, the VP1 proteins of AGMPyV and HPyV9 are exceedingly closely related, and HPyV9 was therefore hypothesized to engage α2,3-linked sialic acid in a manner similar to that observed in AGMPyV (15). The receptor-binding specificity of HPyV9 VP1 was recently defined with glycan microarray screening and crystallographic analysis (6). The virus was found to bind linear, α2,3-linked sialylated glycans in a largely conserved binding region. Surprisingly, HPyV9 displays a preference for glycans terminating in Neu5Gc. Crystal structure analysis of HPyV9 VP1 in complex with a trisaccharide terminating in Neu5Gc identified key residues that underlie this preference. The CH_2-OH group of Neu5Gc projects into a narrow pocket that is lined with polar residues. The side chains of two polar residues, H274 and N282, form hydrogen bonds to the extra hydroxyl group and thus help to define the enhanced specificity of HPyV9 for Neu5Gc-bearing sialyloligosaccharides. Only one of these residues is conserved in AGMPyV (H271), while the other is replaced with a valine (V279) (Fig. 2C).
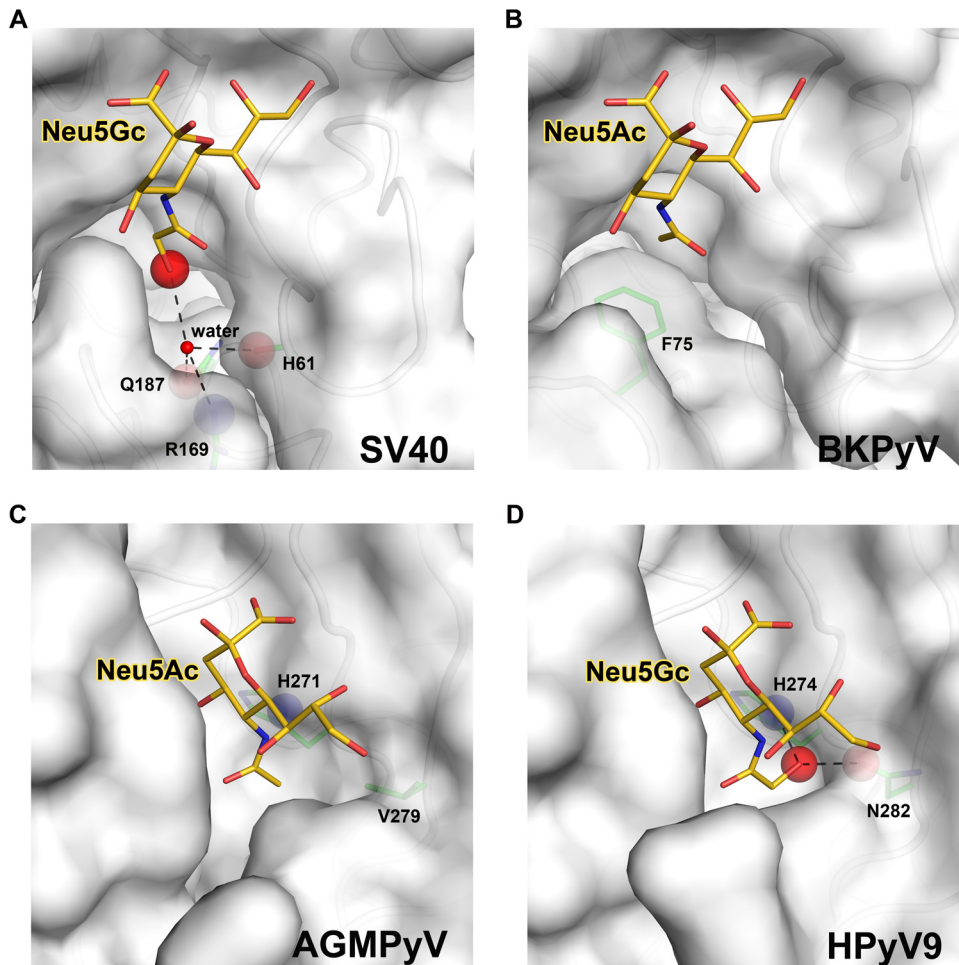
FIG 2 Sialic acid-binding sites in human and simian polyomaviruses. Views into the receptor-binding pockets of SV40 (A), BKPyV (B), AGMPyV (C), and HPyV9 (D) VP1. The sialic acid receptors (Neu5Ac or Neu5Gc) bound by each structure are shown in stick representation and colored gold (carbons), blue (nitrogens), and red (oxygens). The extra oxygen of Neu5Gc is highlighted with a red sphere. VP1 residues that define the specificity for either type of sialic acid are shown as sticks and colored green (carbons), blue (nitrogens), and red (oxygens). The protein main chain or side chain atoms interacting with the extra oxygen of Neu5Gc are highlighted as blue spheres for nitrogens or red spheres for oxygens. Hydrogen bonds are shown as black dashed lines. (A) Model of SV40 VP1 in complex with Neu5Gc (modified from PDB ID 3BWR). Neu5Gc can easily be accommodated in the conformation shown, and the extra oxygen of Neu5Gc could make favorable interactions with several side chains. (B) Structure of BKPyV VP1 in complex with Neu5Ac (modified from PDB ID 4MJ0). The presence of F75 prevents BKPyV from accommodating Neu5Gc. (C) Structure of AGMPyV VP1 in complex with Neu5Ac (modified from PDB ID 4MBZ). The binding site is less polar and does not favor an interaction with Neu5Gc. (D) Structure of HPyV9 VP1 in complex with Neu5Gc (modified from PDB ID 4POT). HPyV9 has enhanced specificity for Neu5Gc due to contacts of the extra hydroxyl group with two polar side chains.

Although they are unable to synthesize it, humans can acquire Neu5Gc from sources such as red meat and milk products (19). Gut endothelium and kidney vasculature in particular display Neu5Gc on their surfaces, and this is, for example, exploited by bacterial toxins that engage glycans terminating in Neu5Gc (7). Similarly, the expression of Neu5Gc on certain human tumors has also been reported (20). The ability to bind to Neu5Gc therefore likely confers an evolutionary advantage to HPyV9, perhaps by directing it to specific tissues.

CONCLUSIONS

Sialic acids are used as initial attachment receptors by a large number of viruses, probably owing to the prominent display of these glycans at the cell surface. It is therefore not surprising that the many biologically occurring sialic acid variants are specifically recognized by some viruses to gain an advantage, such as restricting its host range. This strategy is exemplified by the human BKPyV and JCPyV viruses, which in contrast to their simian relative SV40 cannot interact with Neu5Gc-bearing glycans. The example of HPyV9 reveals a different strategy. Here, a human virus has adapted its binding site to specifically recognize a Neu5Gc-bearing receptor that is not normally found in humans but can be acquired only by diet. In this case, the apparent advantage for HPyV9 may lie in increased access to cells that incorporate dietary glycan structures. The two examples demonstrate that the attachment strategies of many viruses are highly sophisticated and cannot easily be predicted, or even rationalized, from studies of highly homologous viruses. It is often difficult to obtain sialic acid variants such as Neu5Gc in large amounts and even harder to obtain Neu5Gc in the context of a specific glycan structure such as GM1-Gc. Thus, structure-function analyses have been performed with Neu5Ac-based compounds for the bulk of sialic acid-binding viruses. As

the case of HPyV9 demonstrates, limiting the analysis of receptor recognition to the readily available Neu5Ac compound may miss key aspects of the life cycle of a virus.

REFERENCES

1. **Ströh L, Stehle T.** Glycan engagement by viruses: receptor switches and specificity. Annu. Rev. Virol., in press.
2. **Palma AS, Feizi T, Childs RA, Chai W, Liu Y.** 2014. The neoglycolipid (NGL)-based oligosaccharide microarray system poised to decipher the meta-glycome. Curr. Opin. Chem. Biol. **18:**87–94. http://dx.doi.org/10.1016/j.cbpa.2014.01.007.
3. **Mayer M, Meyer B.** 2001. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J. Am. Chem. Soc. **123:**6108–6117. http://dx.doi.org/10.1021/ja0100120.
4. **Varki A.** 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. Nature **446:**1023–1029. http://dx.doi.org/10.1038/nature05816.
5. **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:**13456–13466. http://dx.doi.org/10.1128/JVI.06975-11.
6. **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.
7. **Byres E, Paton AW, Paton JC, Lofling JC, Smith DF, Wilce MC, Talbot UM, Chong DC, Yu H, Huang S, Chen X, Varki NM, Varki A, Rossjohn J, Beddoe T.** 2008. Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. Nature **456:**648–652. http://dx.doi.org/10.1038/nature07428.
8. **DeCaprio JA, Garcea RL.** 2013. A cornucopia of human polyomaviruses. Nat. Rev. Microbiol. **11:**264–276. http://dx.doi.org/10.1038/nrmicro2992.
9. **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.
10. **Campanero-Rhodes MA, Smith A, Chai W, Sonnino S, Mauri L, Childs RA, Zhang Y, Ewers H, Helenius A, Imberty A, Feizi T.** 2007. *N*-glycolyl GM1 ganglioside as a receptor for simian virus 40. J. Virol. **81:**12846–12858. http://dx.doi.org/10.1128/JVI.01311-07.
11. **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.
12. **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.
13. **zur Hausen H, Gissmann L.** 1979. Lymphotropic papovaviruses isolated from African green monkey and human cells. Med. Microbiol. Immunol. **167:**137–153. http://dx.doi.org/10.1007/BF02121180.
14. **Haun G, Keppler OT, Bock CT, Herrmann M, Zentgraf H, Pawlita M.** 1993. The cell surface receptor is a major determinant restricting the host range of the B-lymphotropic papovavirus. J. Virol. **67:**7482–7492.
15. **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.
16. **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.
17. **Sauvage V, Foulongne V, Cheval J, Ar Gouilh M, Pariente K, Dereure O, Manuguerra JC, Richardson J, Lecuit M, Burguiere A, Caro V, Eloit M.** 2011. Human polyomavirus related to African green monkey lymphotropic polyomavirus. Emerg. Infect. Dis. **17:**1364–1370. http://dx.doi.org/10.3201/eid1708.110278.
18. **Trusch F, Klein M, Finsterbusch T, Kuhn J, Hofmann J, Ehlers B.** 2012. Seroprevalence of human polyomavirus 9 and cross-reactivity to African green monkey-derived lymphotropic polyomavirus. J. Gen. Virol. **93:**698–705. http://dx.doi.org/10.1099/vir.0.039156-0.
19. **Tangvoranuntakul P, Gagneux P, Diaz S, Bardor M, Varki N, Varki A, Muchmore E.** 2003. Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. Proc. Natl. Acad. Sci. U. S. A. **100:**12045–12050. http://dx.doi.org/10.1073/pnas.2131556100.
20. **Padler-Karavani V, Hurtado-Ziola N, Pu M, Yu H, Huang S, Muthana S, Chokhawala HA, Cao H, Secrest P, Friedmann-Morvinski D, Singer O, Ghaderi D, Verma IM, Liu YT, Messer K, Chen X, Varki A, Schwab R.** 2011. Human xeno-autoantibodies against a non-human sialic acid serve as novel serum biomarkers and immunotherapeutics in cancer. Cancer Res. **71:**3352–3363. http://dx.doi.org/10.1158/0008-5472.CAN-10-4102.