Specificity and Affinity of Sialic Acid Binding by the Rhesus Rotavirus VP8* Core

Philip R. Dormitzer,1* Zhen-Yu J. Sun,2 Ola Blixt,3 James C. Paulson,3 Gerhard Wagner,2 and Stephen C. Harrison1,4

Laboratory of Molecular Medicine, Children’s Hospital,1 and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,2 Boston, Massachusetts 02115; Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 920373; and Howard Hughes Medical Institute and Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 021384

Received 7 March 2002/Accepted 2 July 2002

Nuclear magnetic resonance spectroscopy demonstrates that the rhesus rotavirus hemagglutinin specifically binds α-anomeric N-acetylneuraminic acid with a Kd of 1.2 mM. The hemagglutinin requires no additional carbohydrate moieties for binding, does not distinguish 3′ from 6′ sialyllactose, and has approximately tenfold lower affinity for N-glycolyneuraminic than for N-acetyleneuraminic acid. The broad specificity and low affinity of sialic acid binding by the rotavirus hemagglutinin are consistent with this interaction mediating initial cell attachment prior to the interactions that determine host range and cell type specificity.

Rotavirus causes approximately 600,000 childhood deaths through dehydrating diarrhea annually (43) and is also an important veterinary pathogen. Understanding early events in cell entry by rotavirus could aid in the development of therapeutic and preventive measures against rotavirus and could help explain the host range and cell type specificity of the virus. A number of rotavirus strains agglutinate human type O red blood cells (31, 39). Hemagglutination and infection of cultured cells by these strains is sensitive to the treatment of cells by neuraminidase (8, 9, 11, 25), indicating that these strains bind sialic acid (N-acetyleneuraminic acid or NeuNAc) on the cell surface. In contrast, a majority of rotavirus strains, including most of those that cause disease in humans, do not hemagglutinate, nor do they require cell surface sialic acid for entry (9, 31).

The specificity of viral receptors for specific linkages of sialic acid is a significant determinant of phenotype in other viruses. For influenza viruses, sialic acid specificity influences host range: avian and equine strains of influenza virus recognize sialic acid in an α2,3 linkage to galactose, while human strains prefer sialic acid in an α2,6 linkage (10, 36). Changes in sialic acid binding specificity have been linked to the spread of influenza virus strains from animal to human populations (28). For polyomaviruses, sialic acid specificity correlates with cell tropism and pathogenicity: while all polyomavirus strains recognize straight-chain oligosaccharides ending in an α2,3-linked sialic acid, strains that also recognize branched oligosaccharides containing an additional α2,6-linked sialic acid are less tumorigenic and spread less extensively in mice (5, 14, 17).

The sialic acid dependence of rotavirus strains is associated with one key difference in entry. While sialic acid-independent strains can infect cells through either the apical or basolateral membrane, sialic acid-dependent strains infect cells only through the apical membrane (7). Sialic acid binding by rotavirus has, however, been clearly linked to any other host range restriction, cell tropism, or pathogenicity. The specificity of rotavirus for sialic acid-containing oligosaccharides and sialylmimetics has been examined by using small molecules as competitors of hemagglutination, infection, or binding to cells (15, 22, 24, 27, 37, 44). The low affinity of rotavirus for monovalent sialosides complicates this approach. Nevertheless, the results of two studies have suggested that rotavirus inhibition is not strictly dependent on the glycosidic linkage of the sialic acid moiety: 2.7 mM 3′ sialyllactose inhibited the binding of rotavirus strain OSU to cultured cells by 50%, as did 1.9 mM 6′ sialyllactose (37), and a synthetic thiosialoside at 6.25 mM in either α2,3 or α2,4 linkage to galactose inhibited rotavirus strain NCDV infection of monolayers by 50% (22).

Studies of rotavirus binding have suggested that either glycoproteins or glycolipids may be the initial sialic acid-containing cellular receptors for some rotavirus strains (11, 16, 37, 41, 44). One intriguing result from such studies is the finding that some rotavirus strains (SA11, OSU, and NCDV) preferentially bind immobilized gangliosides containing the N-glycolyl variant of sialic acid, which is hydroxylated on C11 (Fig. 1) (11, 37). Since their divergence from chimpanzees, humans have sustained a mutation that prevents the synthesis of N-glycolyneuraminic acid from an N-acetyleneuraminic acid precursor (6). As a result, little if any N-glycolyneuraminic acid is present on the surface of human cells (32). A widespread preference of sialic acid-dependent rotavirus strains for N-glycolyneuraminic acid would offer an explanation for the sialic acid independence of most human rotavirus strains (15 out of 16 tested [9]).

We have measured the specificity and affinity of the rhesus rotavirus (RRV) hemagglutination domain for sialosides by nuclear magnetic resonance (NMR) spectroscopy. NMR can be used to determine dissociation constants in equilibrium mixtures of receptors and ligands in solution; it has been used previously to study sialoside binding by influenza virus hemag-
glutinin (38). In the NMR experiments reported here, we titrated the rotavirus hemagglutination domain with sialosides and detected binding by monitoring the changes in the protein backbone amide nitrogen and hydrogen chemical shifts. The characteristic pattern of affected chemical shifts ensured the specificity of the observed changes. Figure 2A shows the migration of the R101 backbone amide peak induced by the binding of 3′ sialyllactose.

Both the protein and the carbohydrates used in this binding study have been extensively characterized. The protein is an *Escherichia coli*-expressed construct (EcVP8<sub>846-231</sub>) that corresponds to a protease-resistant core of the VP8* region of the RRV outer capsid spike protein, VP4 (12). The construct contains the rotavirus hemagglutination domain, which consists of residues L65 to L224. The construction, purification, and NMR solution structure of this recombinant protein and the crystal structure of a similar recombinant protein complexed with a sialoside are described elsewhere (13). The recombinant VP8* core is monomeric (12), so multivalent binding does not complicate the affinity measurements.

The monovalent sialoside ligands used in this study are synthetic products with the structures shown in Fig. 1. N-Acetylneuraminic acid (sialic acid), 2-O-methyl α-N-acetylneuraminic acid (2-O-methyl-α-sialic acid), and lactose were obtained from Sigma. NeuNAc-(α2,3)-Gal-(β1,4)-Glc (3′ sialyllactose) and NeuNAc-(α2,6)-Gal-(β1,4)-Glc (6′ sialyllactose) were obtained from Glyko, Inc. NeuNAc-(α2,3)-Gal-(β1,4)-Glc NeuNAc-β-OCH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub> (N-acetyl trisaccharide) and NeuNGc-(α2,3)-Gal-(β1,4)-Glc NeuNAc-β-OCH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub> (N-glycolyl trisaccharide) were synthesized by using a combined chemical and enzymatic synthesis procedure (3, 4). All charged carbohydrates were prepared as sodium salts at pH 7.0.

Two-dimensional 15N-1H heteronuclear single quantum correlation (HSQC) spectra of mixtures of the VP8* core and carbohydrates were obtained at 25°C by using a 500-MHz Varian spectrometer or a 500-MHz Bruker spectrometer equipped with a cryoprobe. Because only the protein was isotopically labeled with 15N, as previously described (13), only protein resonances were measured. The NMR buffer consisted of 18 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> (pH 7.0), 9 mM NaCl, 0.018% sodium azide, and 10% D<sub>2</sub>O in water. Data were processed with PROSA software (19), spectra were analyzed with the XEASY program (1), binding curves were fitted with Origin 5.0 (Microcal Software, Inc.), and figures were prepared with XEASY, Origin 5.0, Illustrator 8.0 (Adobe Systems, Inc.), and ISIS/DRAW 2.4 (MDL Information Systems, Inc.).

To determine binding affinities, we followed the sialoside-induced changes in the backbone amide chemical shifts of 10 residues. R101, D142, Y155, G156, A166, V167, N178, G179, K187, and T192 were selected for the test set on the basis of their large binding-induced changes in chemical shift. K145, Y175, Y189, T191, and Y194 showed equally large sialoside-induced changes, but their peak positions could not be deter-
mined unambiguously at all sialoside concentrations due to chemical shift overlap and conformational averaging. The NMR and X-ray crystallographic structures of the VP8* core (13) show that most of the residues listed above either bind the sialoside directly or participate in interactions that stabilize sialoside-binding residues. A166 and V167 do not appear to stabilize the binding site, but their backbone amides are positioned close to several aromatic rings so that small binding-induced movements may significantly change their electronic environment. N178 and G179 also do not appear to participate directly in sialoside binding but are adjacent to an area of multiple conformations in the bound structure near the sialoside-binding site.

To determine the binding affinities, the binding-induced changes in chemical shifts were quantified by the following index (40): \[ \Delta f = \sum \left( \frac{(N_0 - N)}{0.5} \right)^2 + \left( \frac{(H_0 - H)}{0.1} \right)^2 \]

where \( N_0 \) and \( N \) are the amide \( ^{15} \text{N} \) chemical shifts in the absence and presence of ligand, respectively; \( H_0 \) and \( H \) are the backbone amide hydrogen chemical shifts in the absence and presence of ligand, respectively; chemical shifts are expressed in parts per million; and the summation is over the 10 residues listed above. The total concentration of protein, \( [P_{\text{tot}}] \), was determined by absorbance using an extinction coefficient at 280 nm of 39.6 \times 10^3 M^{-1} cm^{-1}, based on prediction from the primary amino acid sequence by Protein (DNASTar, Inc.). The total concentration of ligand, \( [L_{\text{tot}}] \), was determined by weight. Under conditions of rapid chemical exchange, the equilibrium chemical shifts of VP8* core atoms in the presence of ligand are the average of the shifts of the VP8* core in an unbound and a ligand-bound state, weighted by the proportion of the VP8* core in each state. The chemical shifts in the unbound state were measured directly. The chemical shifts in the fully bound state were extrapolated from the chemical shifts measured during the titration (Fig. 2B) by using a hyperbolic function, consistent with a simple adsorption isotherm. These extrapolated chemical shifts were used to estimate \( \Delta f_{\text{holo}} \), the chemical shift index of the VP8* core saturated with ligand. The portion of the VP8* core in the bound state (\( \nu \)) is given by \( \nu = \Delta f_{\text{obs}}/\Delta f_{\text{holo}} \), where \( \Delta f_{\text{obs}} \) is the observed chemical shift index. The concentration of free ligand, \( [L_{\text{free}}] \), is given by \( [L_{\text{free}}] = [L_{\text{tot}}] - \nu [P_{\text{tot}}] \). The \( K_d \) was determined by fitting the binding data to the following formula: \[ 1/\nu = 1 + K_d [L_{\text{free}}] \]

with \( [L_{\text{free}}] \) in millimolar.

The chief sources of error in the calculations of \( K_d \) are the errors of predicting an extinction coefficient from sequence data (to determine \( [P_{\text{tot}}] \)), of weighing ligands (to determine \( [L_{\text{tot}}] \)), and of extrapolating \( \Delta f_{\text{holo}} \) (to determine \( \nu \)). Due to these uncertainties, we estimate that the final error of the \( K_d \) calculations is approximately \( \pm 20\% \).

This analysis shows that the VP8* core binds 3’sialyllactose with a \( K_d \) of 1.2 mM (Fig. 2B). Chemical shifts in the presence of 0.5 or 9.9 mM sialic acid in an \( \alpha,\beta \) linkage to lactose (6’sialyllactose) match those in the presence of the same quantities of sialic acid in an \( \alpha,2 \) linkage to lactose (3’sialyllactose) (Fig. 2B). Free sialic acid in aqueous solution is in equilibrium between the \( \alpha \) and \( \beta \) anomers (Fig. 1), with 7% in the \( \alpha \)-anomeric configuration (20). Free sialic acid at a concentration of 10.0 mM produces the same \( \Delta f_{y} \) value as 0.56 mM 3’sialyllactose (Fig. 2B), consistent with the binding of approximately 6% of the free sialic acid. In 2-O-methyl-\( \alpha \)-sialic acid, the lactose moiety of 3’sialyllactose is replaced by a methyl group, which locks the sialoside in the \( \alpha \) anomer (Fig. 1). The changes in chemical shift induced by 1.0 or 10.0 mM 2-O-methyl-\( \alpha \)-neuraminic acid match those induced by 3’sialyllactose (Fig. 2B). Taken together, these data demonstrate that the RRV VP8* core is specific for the \( \alpha \) anomer of sialic acid, that no sugar residues other than sialic acid are required for binding, that the presence of linked galactose does not significantly enhance binding, and that the VP8* core binds sialic acid in \( \alpha,2,3 \) and \( \alpha,2,6 \) linkages to the lactose moiety of lactose equally well.

The VP8* core binds a trisaccharide containing N-glycolyl-neuraminic acid with a \( K_d \) of 11 mM, while it binds an equivalent trisaccharide containing N-acetylenuraminic acid with a \( K_d \) of 1.2 mM (Fig. 2B). This strong preference for an N-acetyl sialoside over an N-glycolyl sialoside indicates that the sialic acid independence of most human rotavirus strains cannot be explained by a general preference of sialic acid-dependent rotavirus strains for N-glycolyneuraminic acid. Rather, there appears to be strain variation in the preference for N-glycolyneuraminic or N-acetylenuraminic acid. The possibility that strain variation in sialic acid binding arises during passage in cell culture cannot be ruled out, as this phenomenon has been observed with passage of influenza virus (35).

The variation in preference for N-acetylenuraminic or N-glycolyneuraminic acid may be linked to a specific genetic variation. The strains documented to preferentially bind N-glycolyneuraminic acid-containing glycosphingolipids (OSU, SA11, and NCDV [11, 37]) have a glycine at position 187, while RRV and strain UK (which gives mixed results in glycosphingolipid binding [11]) have lysine in this position (12, 21, 33, 34). Residue 187 is adjacent to the C-11 methyl group of bound N-acetylenuraminic acid (13), and a lysine in this position might sterically hinder binding of N-glycolyneuraminic acid, which is hydroxylated at C-11 (Fig. 1). Two studies have linked a lysine-to-arginine mutation at residue 187 of RRV VP8* to the acquisition of sialic acid-independent entry despite preserved hemagglutination, suggesting that mutations at this site may affect the specificity, affinity, or conformational effects of sialoside binding (26, 29). Future studies will address the chief sources of error in the calculations of \( K_d \) are the errors of predicting an extinction coefficient from sequence data (to determine \( [P_{\text{tot}}] \)), of weighing ligands (to determine \( [L_{\text{tot}}] \)), and of extrapolating \( \Delta f_{\text{holo}} \) (to determine \( \nu \)). Due to these uncertainties, we estimate that the final error of the \( K_d \) calculations is approximately \( \pm 20\% \).

This analysis shows that the VP8* core binds 3’sialyllactose with a \( K_d \) of 1.2 mM (Fig. 2B). Chemical shifts in the presence of 0.5 or 9.9 mM sialic acid in an \( \alpha,\beta \) linkage to lactose (6’sialyllactose) match those in the presence of the same quantities of sialic acid in an \( \alpha,2 \) linkage to lactose (3’sialyllactose) (Fig. 2B). Free sialic acid in aqueous solution is in equilibrium between the \( \alpha \) and \( \beta \) anomers (Fig. 1), with 7% in the \( \alpha \)-anomeric configuration (20). Free sialic acid at a concentration of 10.0 mM produces the same \( \Delta f_{y} \) value as 0.56 mM 3’sialyllactose (Fig. 2B), consistent with the binding of approximately 6% of the free sialic acid. In 2-O-methyl-\( \alpha \)-sialic acid, the lactose moiety of 3’sialyllactose is replaced by a methyl group, which locks the sialoside in the \( \alpha \) anomer (Fig. 1). The changes in chemical shift induced by 1.0 or 10.0 mM 2-O-methyl-\( \alpha \)-neuraminic acid match those induced by 3’sialyllactose (Fig. 2B). Taken together, these data demonstrate that the RRV VP8* core is specific for the \( \alpha \) anomer of sialic acid, that no sugar residues other than sialic acid are required for binding, that the presence of linked galactose does not significantly enhance binding, and that the VP8* core binds sialic acid in \( \alpha,2,3 \) and \( \alpha,2,6 \) linkages to the lactose moiety of lactose equally well.

The VP8* core binds a trisaccharide containing N-glycolyl-neuraminic acid with a \( K_d \) of 11 mM, while it binds an equivalent trisaccharide containing N-acetylenuraminic acid with a \( K_d \) of 1.2 mM (Fig. 2B). This strong preference for an N-acetyl sialoside over an N-glycolyl sialoside indicates that the sialic acid independence of most human rotavirus strains cannot be explained by a general preference of sialic acid-dependent rotavirus strains for N-glycolyneuraminic acid. Rather, there appears to be strain variation in the preference for N-glycolyneuraminic or N-acetylenuraminic acid. The possibility that strain variation in sialic acid binding arises during passage in cell culture cannot be ruled out, as this phenomenon has been observed with passage of influenza virus (35).

The variation in preference for N-acetylenuraminic or N-glycolyneuraminic acid may be linked to a specific genetic variation. The strains documented to preferentially bind N-glycolyneuraminic acid-containing glycosphingolipids (OSU, SA11, and NCDV [11, 37]) have a glycine at position 187, while RRV and strain UK (which gives mixed results in glycosphingolipid binding [11]) have lysine in this position (12, 21, 33, 34). Residue 187 is adjacent to the C-11 methyl group of bound N-acetylenuraminic acid (13), and a lysine in this position might sterically hinder binding of N-glycolyneuraminic acid, which is hydroxylated at C-11 (Fig. 1). Two studies have linked a lysine-to-arginine mutation at residue 187 of RRV VP8* to the acquisition of sialic acid-independent entry despite preserved hemagglutination, suggesting that mutations at this site may affect the specificity, affinity, or conformational effects of sialoside binding (26, 29). Future studies will address...
the effects of mutations at this site on binding affinity and specificity.

The relatively low affinity of the VP8* core for sialosides parallels NMR measurements of the affinity of influenza virus hemagglutinin for sialosides: influenza virus hemagglutinin (H3, strain X-31) binds 3’ sialyllactose with a $K_D$ of 3.2 mM and 6’ sialyllactose with a $K_D$ of 1.9 mM (38). The multivalency resulting from the 120 molecules of VP4 arrayed on the rotavirus surface and the many sialic acid moieties presented on cell surface glycolipids and glycoproteins should allow for relatively tight binding despite the weak individual interactions. Consistent with this conclusion, sialic acid presented on sialophospholipid vesicles inhibits rotavirus infection with 1,000-fold greater potency than does soluble sialic acid (23). Tight binding due to multivalency has been quantitated for rosettes folded greater potency than does soluble sialic acid (23). Tight binding despite the weak individual interactions.


38. Sauter, N. K., M. D. Bednarski, B. A. Wurzburg, J. E. Hanson, G. M. Whitesides, J. J. Skehel, and D. C. Wiley. 1989. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar disso-