Binding of the Influenza A Virus to Cell-Surface Receptors: Structures of Five Hemagglutinin-
Sialyoligosaccharide Complexes Determined by X-Ray Crystallography

Michael B. Eisen,*† Subramaniam Sabesan,§ John J. Skehel,§ and Don C. Wiley*‡

*Department of Molecular and Cellular Biology, †Committee on Higher Degrees in Biophysics, ‡Howard Hughes Medical Institute, 7 Divinity Avenue, Harvard University, Cambridge, Massachusetts 02138; §National Institute for Medical Research, Mill Hill, London, United Kingdom NN17AA; and ‡Central Research and Development, DuPont Experimental Station, Wilmington, Delaware 19880

Received November 5, 1996; revised December 3, 1996; accepted March 3, 1997

The structures of five complexes of the X-31 influenza A (H3N2) virus hemagglutinin with sialyoligosaccharide receptor analogs have been determined from 2.6 to 2.9 Å resolution by X-ray crystallography. There is well-defined electron density for three to five saccharides in all five complexes and a striking conformational difference between two linear pentasaccharides with the same composition but different linkage (α(2→6) or α(2→3)) at the terminal sialic acid. The bound position of the terminal sialic acid (NanaAc) is the same in all five complexes and is identical to that reported previously from the study of mono- and trisaccharides. The two oligosaccharides with NeuAcα(2→6)Gal linkages and GlcNAc at the third position have a folded conformation with the GlcNAc double-bonded back to contact the sialic acid. The pentasaccharide with a terminal NeuAcα(2→3)Gal linkage and GlcNAc at the third position has an extended (not folded) conformation and exits from the opposite side of the binding site than the α(2→6)-linked molecule of the same composition. These differences between the conformation of the pentasaccharide with a 2,6 linkage and the trisaccharide 2,6-sialyllactose suggest that α(2→6)-sialyllactose is not, as previously believed, an appropriate analog of natural influenza A virus receptors. The oligosaccharides studied are NeuAcα(2→3)Galβ(1→4)Glc, NeuAcα(2→6)Galβ(1→4)Glc, NeuAcα(2→3)Galβ(1→3)GlcNAcβ(1→3)Galβ(1→4)Glc, NeuAcα(2→6)Galβ(1→4)GlcNAcβ(1→3)Galβ(1→4)Glc, and NeuAcα(2→6)Galβ(1→4)GlcNAcβ(1→3)Galβ(1→4)Glc- O(CH₂)₃COOCH₂. © 1997 Academic Press

INTRODUCTION

The initial step in infection of a cell by influenza viruses is the attachment of a virus particle to the target cell surface which is accomplished through the interaction of a glycoprotein found on the viral surface (hemagglutinin; HA), with cell-surface oligosaccharides containing sialic acids (Gotshall, 1959). The three-dimensional structure of the soluble ectoderm of the hemagglutinin (HBA) from the virus strain X-31 (H3N2) was determined by X-ray crystallography, and the sialic acid-binding site was proposed to consist of a pocket of conserved residues at the membrane distal end of the hemagglutinin trimer (Wilson et al., 1981). This was confirmed by the identification of a mutation in this site that altered receptor binding properties (Rogers et al., 1983) and by X-ray crystallographic studies of BHA complexed with compounds containing sialic acid, namely the N-acetylneuraminic acid (NeuAc) (Sauter et al., 1992a,b; Weis et al., 1988). Sialic acid in a number of natural and synthetic sialoglycans has been observed to interact in the same manner with conserved amino acids and the main chain of the HA as shown in Fig. 2A (Sauter et al., 1992ab; Watowich et al., 1994; Weis et al., 1988). The interaction between DI/A and NeuAc has also been characterized by biochemical studies (Kelm et al., 1992; Matrosovich et al., 1991, 1993) and biological measurements of the affinities of various sialoglycans for BHA (Hanson et al., 1992; Sauter et al., 1989, 1992b; Takemoto et al., 1998).

While all influenza A virus attachment requires the presence of terminal sialic acids in oligosaccharide chains (Rogers and Paulson, 1983), different viral strains have different affinities for different sialyglycosamides. In particular, certain viral strains show a preference for binding to erythrocytes enzymatically modified to bear only sialic acids attached with an α(2→3) linkage, while other strains show a similar preference for the α(2→6) linkage (Rogers and Paulson, 1983). A single amino acid substitution in the receptor binding site of the X-31 hemagglutinin (leucine HAI 226 to glutamine) results in a change in this linkage specificity from α(2→6) to α(2→3) (Rogers et al., 1983). While the opposite mutation in a related strain (A/clock/Ukraine/1/83) with a natural preference for the α(2→3) linkage results in a shift to an α(2→6) preference (Rogers et al., 1985). Mutation studies have, however, shown that position 226 is not an exclusive determinant of linkage specificity (Doria et al., 1987) and in H1 subtype strains position 226 does not appear to be significant in defining linkage specificity as both α(2→3) and α(2→6) preferences are observed for strains.

1Current address: Department of Genetics, Stanford University School of Medicine, Stanford, CA.
2To whom reprint requests should be addressed.
with the same amino acid at 226 (Rogers and D'Souza, 1999). Surveys of the linkage specificities of a large number of natural H3 and its subtype strains also indicate that the amino acid at position 226 correlates with linkage specificity (Miatrosoovich et al., 1993), although it appears to vary in concert with changes at position 228 (Connor et al., 1994).

A strong correlation between linkage specificity and a virus’ host of origin has suggested that different host species may exert receptor-based or inhibitor-based selective pressures (Baum and Paulson, 1990; Connor et al., 1994; Couceiro et al., 1993; Rogers and Paulson, 1983; Ryan-Poirier and Kawaoka, 1991).

Weiss et al. (1989) attempted to gain insight into the structural basis for linkage specificity by comparing the three-dimensional structure of the binding site of the BHA from X-31 to that of an X-31 variant selected to have α(2→3) linkage specificity (X-31/HS) (Rogers et al., 1982). However, X-ray crystallographic studies of X-31 and X-31/HS BHA complexes with the trisaccharides 2,3-sialyllactose and 2,6-sialyllactose failed to produce electron density for most of the asialo portions of the sugars (Weis et al., 1988, 1990). Recent low-resolution (5.5 Å) X-ray crystallographic studies of X-31 BHA complexes with higher purity sialyllactose produced electron density for all three sugars of these two trisaccharides (Eisen, 1996). We report here the results of higher resolution (2.6 to 2.8 Å) X-ray crystallographic studies of complexes of X-31 BHA with five sialyloligosaccharides that represent diversity in sialic acid linkage type, length of oligosaccharide, and composition of the asialo portion of the oligosaccharide chain. The five oligosaccharides are

1. NeuAcα(2→3)Galβ(1→4)Glc (2,3-sialyllactose; 23SL),
2. NeuAcα(2→6)Galβ(1→4)Glc (2,6-sialyllactose; 26SL),
3. NeuAcα(2→3)Galβ(1→3)GlcNAcβ(1→3)Galβ(1→4)Glc (lactoseries tetrasaccharide α; LSta),
4. NeuAcα(2→6)Galβ(1→3)GlcNAcβ(1→3)Galβ(1→4)Glc (lactoseries tetrasaccharide α; LSta), and
5. [NeuAcα(2→6)Galβ(1→4)GlcNAcβ(1→3)Galβ- O-(CH₂)₃-COOCH₃ (3,6-disialyllactosamine; 36DSL).

For clarity in the remainder of this text, the sugar monomers in each oligosaccharide are numbered sequentially beginning with 101 for the NeuAc group (this corresponds to the numbering of the coordinate files deposited with the protein data bank).

MATERIALS AND METHODS

2,3-Sialyllactose and 2,6-sialyllactose were purchased from BioCarb Chemicals. LSTA and LSTc were purchased from Oxford Glycosystems. 3,6-Disialyllactosamine was made according to the published methods (Sabesan et al., 1992). These compounds were used without further purification.

X-31 BHA was prepared as previously described (Brand and Skehel, 1972) and concentrated to 40 mg/ml in 150 mM sodium chloride. Crystals were grown by vapor diffusion against a well solution containing approximately 1.4 M sodium citrate, pH 7.5 (Willey and Skehel, 1977). These crystals were isomorphous to those used in the structure determination of X-31 BHA (Wilson et al., 1981) and were in the spacegroup P4₁ with unit cell dimensions approximately a = b = 160 Å, c = 175 Å, and α = β = γ = 90°. Crystals were harvested in 1.41 M sodium citrate, 0.02% sodium azide, pH 7.5, and serially transferred to solutions containing gradually increasing amounts of the oligosaccharide ligands (at a concentration of 16 to 25 mM; see Table 1) and the cryoprotectant xylitol (Watovich et al., 1984). Two ligand/cryoprotectant solutions were prepared by dissolving lyophilized ligand in a cryoprotectant solution made by dissolving xylitol (Fluka) to 1.5 M in harvest buffer (this results in a reduction in the sodium citrate concentration from the 1.41 M of the harvest buffer; the final citrate concentration has not been measured). Serial transfer solutions were prepared by mixing the final ligand/cryoprotectant solution with harvest buffer. In general crystals were transferred through 32×, 16×, 8×, 4×, 2× dilutions of the ligand/cryoprotectant solution for 30 min per solution before a 30-min soak in the undiluted ligand/cryoprotectant solution.

Crystals were drawn into loops made of 70-μm Nichrome wire (approximate diameter of 1 mm) (Teng, 1990) and flash-cooled to approximately −160°C in a stream of liquid nitrogen-cooled nitrogen gas (Rodgers, 1994). Data were collected at −160°C on a Siemens multiwire area detector or a 340-mm MAR image plate detector using X rays from a GX-13 rotating anode, or on Fuji image plates using X rays from the CHESS A-1 and F-1 beamlines (see details in Table 1). Data from the area detector were processed using XDS (Kabsch, 1988) and scaled with programs in the CCP4 suite (Collaborative Computational Project, 1994). Data from the MAR and Fuji image plates were processed using DENZO (Z. Otwinowski) and scaled with SCALEPACK (Z. Otwinowski).

For refinement, a starting model from the 2.2 Å structure of X-31 BHA complexed with a high affinity NeuAα antigen (α-2-O-β-naphthylmethylimidocarboxylhexyl-5-N-acetylneuraminic acid) (Watovich et al., 1994) was used for each complex. The non-sialic acid portion of this ligand was removed and replaced by models for the non-sialic acid portions of the oligosaccharide chains in the Fα electron-density maps computed using phases from this initial model and Fα from a 2.15 Å native X-31 BHA dataset (S. Watovich, unpublished data). One hundred fifty-nine water molecules were included from the NCSA model (none were in the binding site and none were within 7.0 Å of ligand atoms in the final model). Refinement was performed with X-PLOR (Brünger, 1992b). Five percent of the reflections were set aside as an Rfree set (Brünger, 1992a). A brief energy minimization
TABLE 1
Data Collection and Model Statistics

<table>
<thead>
<tr>
<th></th>
<th>2,3SL</th>
<th>2,6SL</th>
<th>3,6OSL</th>
<th>LSTa</th>
<th>LSTc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligand concentration</td>
<td>25 mM</td>
<td>25 mM</td>
<td>25 mM</td>
<td>16 mM</td>
<td>18 mM</td>
</tr>
<tr>
<td>X-ray source</td>
<td>GX-13</td>
<td>CHESS F-1</td>
<td>GX-13</td>
<td>GX-13</td>
<td>GX-13</td>
</tr>
<tr>
<td>Detector</td>
<td>Siemens AD</td>
<td>Fuji IP</td>
<td>Siemens AD</td>
<td>MAR 30 mm</td>
<td>MAR 30 mm</td>
</tr>
<tr>
<td>Resolution</td>
<td>2.6 Å</td>
<td>2.5 Å</td>
<td>2.6 Å</td>
<td>2.6 Å</td>
<td>2.6 Å</td>
</tr>
<tr>
<td>Completeness (to real)</td>
<td>68.3%</td>
<td>92.7%</td>
<td>70.3%</td>
<td>92.2%</td>
<td>97.3%</td>
</tr>
<tr>
<td>R merge</td>
<td>0.080</td>
<td>0.077</td>
<td>0.089</td>
<td>0.096</td>
<td>0.118</td>
</tr>
<tr>
<td>Refinement R-work (R-free after)</td>
<td>27.2 (31.8)</td>
<td>25.1 (28.4)</td>
<td>24.4 (28.2)</td>
<td>21.8 (26.4)</td>
<td>21.2 (25.7)</td>
</tr>
<tr>
<td>B individual</td>
<td>22.1 (26.5)</td>
<td>21.5 (24.8)</td>
<td>21.3 (26.4)</td>
<td>19.3 (24.0)</td>
<td>18.7 (23.2)</td>
</tr>
<tr>
<td>Rebuilding/modeling of ligand</td>
<td>23.7 (27.4)</td>
<td>22.3 (25.3)</td>
<td>22.6 (26.3)</td>
<td>19.4 (24.3)</td>
<td>18.8 (23.2)</td>
</tr>
<tr>
<td>Simulated annealing</td>
<td>22.4 (26.3)</td>
<td>21.9 (24.8)</td>
<td>21.3 (25.0)</td>
<td>19.4 (23.6)</td>
<td>18.8 (23.0)</td>
</tr>
<tr>
<td>Ligand B individual</td>
<td>19.2 (23.5)</td>
<td>18.8 (22.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dihedral angles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NeuAc-Gal</td>
<td>-62°</td>
<td>-133°</td>
<td>-57°</td>
<td>-68°</td>
<td>-57°</td>
</tr>
<tr>
<td></td>
<td>-12°</td>
<td>-176°</td>
<td>-162°</td>
<td>-18°</td>
<td>-154°</td>
</tr>
<tr>
<td></td>
<td>-69°</td>
<td>56°</td>
<td>62°</td>
<td>60°</td>
<td></td>
</tr>
<tr>
<td>Gal-Glc/GlcNAc</td>
<td>59°</td>
<td>27°</td>
<td>30°</td>
<td>34°</td>
<td>34°</td>
</tr>
<tr>
<td></td>
<td>-12°</td>
<td>30°</td>
<td>-22°</td>
<td>-40°</td>
<td>-25°</td>
</tr>
<tr>
<td></td>
<td>32°</td>
<td>19°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9°</td>
<td>-23°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gal/Glc</td>
<td>51°</td>
<td>38°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4°</td>
<td>40°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ligand B-factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NeuAc</td>
<td>53 Å²</td>
<td>38 Å²</td>
<td>49 Å²</td>
<td>33 Å²</td>
<td>28 Å²</td>
</tr>
<tr>
<td>Gal</td>
<td>63 Å²</td>
<td>60 Å²</td>
<td>71 Å²</td>
<td>63 Å²</td>
<td>55 Å²</td>
</tr>
<tr>
<td>Glc/GlcNAc</td>
<td>76 Å²</td>
<td>66 Å²</td>
<td>87 Å²</td>
<td>70 Å²</td>
<td>74 Å²</td>
</tr>
<tr>
<td>Gal</td>
<td>78 Å²</td>
<td>90 Å²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glc</td>
<td>85 Å²</td>
<td>97 Å²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein B-factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All residue</td>
<td>37 Å²</td>
<td>26 Å²</td>
<td>30 Å²</td>
<td>22 Å²</td>
<td>24 Å²</td>
</tr>
<tr>
<td>Binding-site residues</td>
<td>46 Å²</td>
<td>31 Å²</td>
<td>36 Å²</td>
<td>27 Å²</td>
<td>30 Å²</td>
</tr>
</tbody>
</table>

Note. For each of five complexes, data collection and model refinement statistics are given along with final refined dihedral angles and average B-factors for non-hydrogen atoms in ligand and protein models. Dihedral angles are determined using the following atoms: NeuAcα(2 → 3)Gal φ = C1-C2-O3'-C3', ψ = C2-O3'-C3'-H3'; NeuAcα(2 → 6)Gal Φ = C1-C2-O6'-Gal, Ψ = C2-O3'-C6'-O6', ω = O6'-C6'-O5'-O5', β(1 → 4) linkages Φ = H1-C1-O4'-C4', Ψ = C1-O4'-C4'-H4'; β(1 → 3) linkages Φ = H1-C1-O3'-C3', Ψ = C1-O3'-C3'-H3'.

was followed by simulated annealing refinement (temperature from 3000 to 300uK) using X-PLOR verlet dynamics and then refinement of individual B-factors. Only the NeuAc portion of the oligosaccharide was used in this cycle of refinement. Weak noncrystallographic symmetry restraints (X-PLOR ncs-weight = 13.0) were used, with separate constraints on the HA1 and HA2 chains to allow for adjustment to a new molecular threefold symmetry axis. Ten cycles of noncrystallographic averaging and solvent flattening were performed using the averaging program AVE (Jones, 1992) masks generated with the programs ENVAT (Madden, 1992) and MAPMAN (Kleywegt and Jones, 1996) and programs in the CCP4 suite (Collaborative Computational Project, 1994). Averaged and unaveraged Fo-Fc and 2Fo-Fc maps were computed using the native dataset discussed above, and phases were computed following averaging and solvent flattening. Ligand atoms were modeled using unaveraged Fo-Fc and 2Fo-Fc maps iteratively, with model building followed by energy minimization in X-PLOR (Brünger, 1992b). Protein atoms within 15 Å of the receptor binding site were rebuilt using a combination of averaged and unaveraged 2Fo-Fc maps. A second round of simulated annealing refinement (temperature from 3000 to 300u)K) was performed with weak ncs-restraints on the entire monomer. Occupancies of 1.0 were used for all modeled ligand atoms in the second round of refinement. For all five complexes, there was clear electron density for at least three sugar monomers (including NeuAc). For LSTa and LSTc, electron density was observed for all...
FIG. 1. Stereo views of electron density for each ligand in the X-31 binding site. 2,3-Sialyllactose, averaged $F_o - F_c$ map contoured at 3 $\sigma$ with a 2 Å cover radius. 2,6-Sialyllactose, averaged $F_o - F_c$ map contoured at 2.6 $\sigma$ with a 1.5 Å cover radius. 3,6-Disialyllactosamine, averaged $F_o - F_c$ map contoured at 2.5 $\sigma$ with a 2.0 Å cover radius. LS1a, unaveraged $F_o - F_c$ map contoured at 1.5 $\sigma$ with a 1.5 Å cover radius. And LS1c, unaveraged $F_o - F_c$ map contoured at 1.0 $\sigma$ with a 1.3 Å cover radius. All maps computed using phases calculated following 10 cycles of averaging and solvent flattening of final refined model. Figures generated with MOLSCRIPT (Krulwich, 1991).
five sugars. In all cases Gal-102 and Glc/GlcNAc-103 could be unambiguously modelled, based on (1) the necessity of placing linkage atoms adjacent to neighboring residues, (2) the flattened overall shape of the electron density (see Fig. 1), which constrains the position of the pyranose ring, and (3) features in the density corresponding to exocyclic hydroxymethyl groups, the N-acetyl group of GlcNAc, and the axial 4-hydroxyl group of galactose (see Fig. 1). For LSTa and LSTc approximate models for the position of Gal-104 and Glc-105 could also be built. For 36DSI, no electron density for a fourth sugar was visible, and only a model for the terminal trisaccharide could be built.

RESULTS

Electron density maps for the five complexes are shown in Fig. 1. In all cases the electron density for the terminal NeuAc is excellent and allows for unambiguous modeling of this sugar. The position of the NeuAc group and the interactions between the protein and this sugar are the same (within experimental error) in all five complexes and correspond to those previously reported (Sauer et al., 1992b; Watowich et al., 1994; Weis et al., 1988).

Figure 2 shows the final, refined models of each complex with potential hydrogen bonds between the ligand and the HA as dashed lines. The torsion angles of the glycosidic linkages that describe the overall conformation of the oligosaccharides and temperature factors for each monosaccharide are listed in Table 1.

LSTc: NeuAcα(2→6)Galβ(1→4)GlcNAcβ(1→3) Galβ(1→4)Glc

The pentasaccharide LSTc binds in a folded conformation (Φ = −67°, Ψ = −164°, ω = 82°), such that the third saccharide, GlcNAc, is directly over the terminal NeuAc (Figs. 1 and 2C). The GlcNAc 3-OH group is positioned close enough to hydrogen bond to the ring oxygens of either the NeuAc or the intervening Gal residue or to the NeuAc glycosidic oxygen and to make van der Waal contacts with NeuAc C-2 and C-3. The carbonyl oxygen.
carbonyl carbon, and methyl group of GlcNAc are also within van der Waal contact of the NeuAc 4-OH and acetamido carbonyl oxygen. [The GlcNAc is located over the “leading” edge of the NeuAc (Fig. 2C) near NeuAc 4-OH and over 4Å away from the 7-OH of the NeuAc glycerol substituent (see also Fig. 3B).] Weaker electron density for Gal-104 and Glc-105 is visible, indicating that the oligosaccharide chain exits the site to the right (Fig. 2C), away from the subunit boundary in the trimer and an N-linked oligosaccharide of the adjacent protein subunit (Figs. 3A and 3B). The four non-sialic acid residues appear to make no significant contacts (<4.0 Å) with the HA, although the distal-most Glc-105 cannot be accurately modeled (Fig. 1C).

36DSL: [NeuAcα(2→6)Galβ(1→4)GlcNAc]β(1→3/6)Gal-β-O-(CH₂)₃-COOCH₃

The 3,6-disialyllactosamine has the same terminal 2,6-sialyllactosamine as LSTc, and its first three saccharides bind to HA essentially the same conformation (Φ = -0°, Ψ = -126°, ω = 56°) as observed for LSTa. The fourth sugar (which bridges the two sialyllactosamine moieties) of this bidentate molecule is not seen in the electron-density maps. The high quality of the density for the sialyllactosamine indicates that only one of the two sialylated branches is more likely that either one can bind in nearly identical conformations.

LSTa: NeuAcα(2→3)Galβ(1→3)GlcNAcβ(1→3)Galβ(1→4)Glc

The pantasaccharide LSTa has the same composition as LSTc, but with NeuAc in an α(2→3) linkage rather than α(2→6), and the second linkage is β(1→3) rather than β(1→4). LSTa, however, binds in an extended conformation (Φ = -98°, Ψ = -18°) (Figs. 1 and 2D). There is a potential hydrogen bond between the Gal-102 exocyclic hydroxymethyl group and the main chain of Gly-229 and another between the 2-OH of Gal-104 and the Glc-189 on the top of the α-helix in the HA-binding site (Fig. 2D). None of the asialo portion of the oligosaccharide is tightly packed against the HA; only one interatomic distance is <3.5 Å. LSTa exits from the binding site near the interface between the two HA monomers (Fig. 3A) pointing toward the trimer's long axis (Fig. 3C) on the opposite side of the binding site from where the LSTc chain exits (Fig. 3A).

26SL: NeuAcα(2→6)Galβ(1→4)Glc

The triasaccharide 2,6-sialyllactose binds in an extended conformation (Φ = -133°, Ψ = -175°, ω = -89°) with the lactose portion reaching toward the “left end” of the binding site that is bordered by an N-linked carbohydrate from the adjacent monomer in the HA trimer in a manner similar to LSTa (Figs. 1 and 2E). Gal-102 is elevated from the site (closest backbone contact 4.5 Å, average closest contact 5.8 Å) while Glc-103 contacts a portion of the binding site near Ser-227 (closest contact 2.8 Å; average closest contact 3.3 Å) (Fig. 2E). The 1-OH and 2-OH groups of Glc-103 appear to form hydrogen bonds with Ser-227 (1-OH to Ser-227 Oγ; 2-OH to Ser-227 NH and Oγ) and the ring oxygen, O-6, to the indole NH of Trp-222.

23SL: NeuAcα(2→3)Galβ(1→4)Glc

The trisaccharide 2,3-sialyllactose binds in an extended conformation (Φ = -62°, Ψ = -12°) and in a location similar to LSTa (op. Figs. 2D and 2F). The same hydrogen bond between the exocyclic hydroxymethyl group of Gal-102 and the main chain of Gly-225 is observed in the 23SL complex as the LSTa complex (Figs. 2D and 2F). GlcNAc-103, which is attached to Gal-102 through a β(1→4) rather than a β(1→3) linkage as in LSTa, is twisted and slightly lower in the site in the 23SL complex. Neither Gal-102 nor Glc-103 is tightly packed against the protein. The closest contact between the pyranose backbone of Gal-102 and the protein is 3.9 Å and the average closest contact distance for atoms in the Gal-102 backbone is 4.7 Å. For Glc-103 the corresponding values are 4.8 and 5.7 Å, respectively.

Differences in non-crystallographically related binding sites

For all of the complexes examined, there were differences in the quality of the electron density in the three non-crystallographically related binding sites (because the molecular tetradecameric symmetry axis is not a crystallographic axis, in the crystal each binding site exists in a different environment). Some of these differences were most dramatic for LSTa, where absolutely no ligand density was present in one of the three binding sites. We found that in the site where no LSTa electron density was observed, a crystal contact involving residues in the N-terminus of the HA1 chain, the C-terminus of the HA2 chain, and a short loop from residues 144 to 145 of the HA2 chain precludes the binding of LSTa in its observed conformation. While no protein electron density was observed in the crystal that would directly prevent the binding of the other four ligands (LSTc, 36DSL, 26SL, 23SL) in their modeled orientations, there are some protein residues for which no electron density has been observed, particularly the N-terminus of HA1, which might affect binding of these ligands and account for the variation in quality of the ligand electron density in this site.

DISCUSSION

The observation that LSTc and 36DSL, both containing terminal α(2→6) linked sialic acid and GlcNAc as the third saccharide, bind to the HA in a folded conformation (Figs. 2B and 2C) is consistent with both conformational energy calculations and NMR studies on sialosides in
solution that contain an α(2→6) linkage (Berman, 1984; Breg et al., 1989; Ohru et al., 1991; Poppe et al., 1992; Sabesan et al., 1991a,b, Vliegenthart et al., 1983, 1981). Evidence for a folded conformation in solution includes: (1) the measurement of Galβ(1→4)GalNAc is sialylated with an α(2→6) linkage, but not with an α(2→8) linkage, indicating close spatial interactions between GalNAc 3-OH and NeuAc in the α(2→6) saccarides (Berman, 1984; Sabesan et al., 1991a; Vliegenthart et al., 1983, 1981), and (3) the modification of certain GalNAc chainlinking effects with the removal of partial NeuAc chemical groups (Sabesan et al., 1991a,b).

The values of the torsion angles Φ and Ψ are less constrained in solution according to both conformational energy calculations and NMR measurements (Breg et al., 1989; Poppe et al., 1992; Sabesan et al., 1991a). Both LSTc and 3DSL bound to HA show ΦΨ angles (Φ = −57°, Ψ = −154°) near those favored by Poppe et al. (1992), but although we observe the intramolecular Glc 3-OH to Gal ring-O hydrogen bond, the bound conformation does not make the Glc 3-O to NeuAc 7-OH bond predicted. The HA-bound conformation that we observed is considered by others to have a minor role in solution (Sabesan et al., 1991a). Of the four examples of α(2→6)-linked sialic acid X-ray structures reported elsewhere, the two glycoprotein fragments bound to wheat-germ agglutinin (WGA) (Wright, 1992; Wright and Jaeger, 1993) and one of two isoclinic–oligosaccharide complexes (Bourne et al., 1994) have ω angles similar to those of LSTc and 3DSL, but only one of the four, a WGA complex, has similar ΦΨ values.

The structures of LSTc and 3DSL complexes with HA suggest that the folded conformation of the NeuAcα(2→6)-Gal(β(1→4))GlcNAcβ(1→2) trisaccharide is stabilized by a hydrogen bond between NeuAc and GlcNAc and by van der Waals contacts in the interface between these two saccharides; the GlcNAc3-OH group is positioned close enough to hydrogen bond to the ring oxygens of either the NeuAc or the intervening Gal residue or to the NeuAc glycosidic oxygen. Approximately 40 Å2 of the oligosaccharide's solvent-accessible surface is buried by intramolecular contacts in the folded conformation relative to the extended conformation. The lack of contacts between the asialo portion of these oligosaccharides and the HA in both bound LSTc and 3DSL suggests either that the folded conformation is an intrinsic property of the NeuAcα(2→6)Gal(β(1→4))GlcNAc moiety in solution or that this conformation is chosen during binding from among conformations stable in solution because it will fit into the HA-binding site with the least energetic cost. If the observed conformation is intrinsically the most stable, then sialeyl oligosaccharides containing this α(2→6)-linked terminal trisaccharide may have this folded conformation when bound to other sialoside-binding proteins, including the receptors of the influenza C virus (Hoffing et al., 1996) and some reoviruses (Paul et al., 1989). The apparent inability of other sialoside-binding viruses, such as some paramyxoviruses (Cahan and Paulson, 1980), polioviruses (Markwell et al., 1991), and coronavirus (Schultze et al., 1996; Vlasak et al., 1988), to recognize the α(2→6) linkage may result from poor complementarity with this folded conformation.

The observed conformation of bound NeuAcα(2→6)lactose differs in two ways from that of NeuAcα(2→6)lactosamine. First, the C-6 arm of the galactosamine adopts a near "gg" orientation (ω = −89°), assumed to be disfavored due to repetitive nonbonded interaction with C-4-OH of the galactose. This observation also contrasts to that seen in the free trisaccharide (Sabesan et al., 1991a), where it has been shown that the C-6 rotamer is preferentially populated in "gt" orientation (ω = 71°). The bound NeuAcα(2→6)lactose also exhibits the hitherto unobserved anti orientation with respect to the C-6 of galactose. Although conformational calculations have indicated the presence of anti (Sabesan et al., 1991a) and syn (Breg et al., 1989) conformations, our observations demonstrate the presence of the anti conformation in a NeuAcα(2→6)lactoside, where it is presumably stabilized by the interactions observed in the complex.

The trisaccharide 2,8-sialyl lactoside, which lacks the N-acetyl group on the third saccharide, binds in an extended conformation with the third saccharide (glucose) making specific contacts with the protein involving the Glc 1- and 2-OH groups (Fig. 2E). Because in natural receptors the 1-position of the third sugar is glycosylated and the 2-position is N-acetylated, the observed 2,6-sialyllactose-binding mode cannot be that of natural receptors containing NeuAcα(2→6)lactosamine units. The binding of LSTc, a sialoside with α(2→6)-linked sialic acid, therefore, appears to demonstrate a binding mode more likely to represent that of the influenza virus receptor in vivo (Figs. 2C and 3B).

In the extended conformations of LSTc and 2,3-sialyllactose bound to HA, the α(2→3) linkage (Φ = −68°, Ψ = −18°) corresponds to the syn conformer identified in modeling studies (Breg et al., 1989; Poppe et al., 1989; Sabesan et al., 1991a). Two low-energy conformations, syn and anti, with ΦΨ angles of approximately −70°, 5° and −160°, −20°, are common to conformational calculations of α(2→3)-linked sialosides, and NMR studies of such oligosaccharides have been interpreted to show that both of these conformations are significantly populated in solution (Breg et al., 1989; Poppe et al., 1989; Sabesan et al., 1991a). Both syn (Sauer et al., 1992a,b; Stehle and Harrison, 1996; Stehle et al., 1994; Wright, 1990, 1992; Wright and Jaeger, 1993) and anti (Merritt et al., 1994) conformations have been found in other protein–sialoside complexes.

The fact that LSTc exits the site toward the molecular
threefold symmetry axis provides a model for the design of trimeric sialosides where three sialic acids, one to reach into each binding site, could be synthetically linked to a hub on the threefold axis (Priebe, 1993; Watowich et al., 1994).

The large differences observed here in the way that the longer oligosaccharides, LSTc and LSTA with α(2→6)- or α(2→3)-linked sialic acid, bind to the influenza virus HA (Fig. 3A) may provide a structural foundation for the linkage specificity observed among different viral strains. Viruses bearing the HA used in the studies reported here agglutinates erythrocytes specifically modified to contain only α(2→6) sialic acid better than those with α(2→3)-linked sialic acid (Rogers and Paulson, 1983). The structure of a mutant HA that contains a single amino acid substitution, HA1 L220Q, and prefers α(2→3)-linked sialic
FIG. 3. Conformations of LSTα and LSTc. (A) LSTα (red), LSTc (green), and HA monomers colored white and gold; (B and C) The molecular surface of BHA viewed along the threefold molecular symmetry axis. BHA monomers are colored white, light-blue, and gold, with (B) LSTc in green and (C) LSTα in red. Ligand models are from one binding site, inset-squares fitted onto the other binding sites. Generated with GRASP (Nicholls et al., 1991).

Acid (Rogers et al., 1983) has been determined by X-ray crystallography (Eisen, 1996; Wells et al., 1988, 1990). When LSTc and LSTα are modeled into the binding site of the L226Q HA by superimposing their sialic acids with those of bound ligands, the C6 linkage atom of LSTc and the C6, C4, and C3 atoms of the Gal of LSTα are about 4 Å from Gln-226 (Eisen, 1996). Although the consequences of such modeled contacts cannot be deduced from the structures alone, in each case water molecules associated with the Gln Oe and NH2 in the ligand-free site would be displaced. It may be that the more extensive contacts between the α(2→3)- than the α(2→6) sialosides and Gln-226 can better compensate for the loss of the contacts to water. Measurement of the differences in affinity of the wild-type and mutant HA molecules for the polyaaccharides, LSTc and LSTα,
is required to determine whether they are larger than the very small differences in affinity observed between α(2→6)- and α(2→3)-sialyllactose (Sauter et al., 1989).

In addition to HA1 L226Q, as noted in the introduction, amino acid substitutions at other positions in HA have been suggested to influence receptor-binding preferences (Connor et al., 1994). For example, replication of influenza viruses in cells from different species or in
different tissues can select HAs that differ at residues HA1 156 (E→K), 186 (S→I), 193 (N→K), and 194 (L→I) (Gubareva et al., 1994; Hardy et al., 1995; Robertson et al., 1995) and these changes may be related to differences in binding specificities. The locations of these residues in the receptor-binding site suggests that the substitutions might contribute to a preference for α(2→3)- or α(2→6)-linked sialylated oligosaccharides as virus receptors.
Studies of the H-As of viruses isolated from different species and in different cells have also been interpreted to show that sequence changes near the receptor-binding site may modify not only receptor-binding properties but also antigenicity (Schild et al., 1983; Skehel and Wiley, 1988). These interpretations were supported by the findings that, in antigenic variants selected by monoclonal antibodies, amino acid substitutions near the receptor-binding site can change both receptor-binding specificity and antigenicity (Daniels et al., 1987). It is possible also that receptor binding properties may influence the efficiency of neutralization of infectivity by antibodies that block HA–receptor interactions. Antibodies that bind close to the receptor-binding site may efficiently prevent its interactions with any receptor; those that bind at a distance from the site may more efficiently block interactions with receptor oligosaccharides that exit the site in one of another of the orientations observed here. Changes in receptor-binding specificity may in this way increase or decrease the effectiveness of different antibodies to neutralize virus infectivity and as a consequence contribute to the overall mechanism of natural antigenic drift (reviewed in Wiley and Skelchol, 1987).

ACKNOWLEDGMENTS

We are grateful to K. Leong, P. Rosenthal, S. Garman, S. Watovich, T. Jardetzky, M. Lawrence, and the staff at the CHESS synchrotron for assistance in data collection and D. Stevens and R. Gonsalves for technical assistance. This work was supported by NIH Grant AI-13654 and by the MRC. M.B.E. was supported by a postdoctoral fellowship from the National Science Foundation. D.C.W. is an investigator of the Howard Hughes Medical Institute.

REFERENCES


Nicholls, A., Sharp, K., and Henig, B. (1991). Protein folding and associ-


