Structural Studies of the Parainfluenza Virus 5 Hemagglutinin-Neuraminidase Tetramer in Complex with Its Receptor, Sialyllactose

Ping Yuan,1 Thomas B. Thompson,1 Beth A. Wurzburg,1 Reay G. Paterson,1 Robert A. Lamb,1,2 and Theodore S. Jardetzky1,*
1Department of Biochemistry, Molecular Biology, and Cell Biology
2Howard Hughes Medical Institute
Northwestern University
Evanston, Illinois 60208-3500

Summary
The paramyxovirus hemagglutinin-neuraminidase (HN) functions in virus attachment to cells, cleavage of sialic acid from oligosaccharides, and stimulating membrane fusion during virus entry into cells. The structural basis for these diverse functions remains to be fully understood.(0.29,0.45)

Introduction
The paramyxovirus family of enveloped, negative-stranded RNA viruses has many members and includes parainfluenza viruses 1–5, mumps virus, measles virus, Newcastle disease virus (NDV), Sendai virus, Nipah virus, Hendra virus, and respiratory syncytial virus. The viruses possess two major spikes, glycoprotein one involved in cell attachment (HN, H, or G), and the other glycoprotein (F) involved in mediating pH-independent fusion of the viral envelope with the plasma membrane of the host cell.

For many of the paramyxoviruses (e.g., parainfluenza virus 5 [SV5]), human parainfluenza viruses 1–4, NDV, Sendai virus, mumps virus), the attachment protein (HN) has hemagglutinating and neuraminidase (NA) activities. HN binds to sialic acid-containing cell surface molecules; furthermore, HN mediates enzymatic cleavage of sialic acid from the surface of virions and the surface of infected cells (reviewed in Lamb and Kolakofsky, 2001). It is thought that NA activity is required to remove sialic acid from the viral and cellular conjugates during viral assembly, to allow release of newly budded virus from the virus-infected cell (reviewed in Choppin and Scheid, 1986).

In addition to receptor binding and NA activity, HN has a fusion-promoting activity that plays an important role in virus-cell fusion. The coexpression of HN and F derived from the same paramyxovirus lowers the activation energy barrier for triggering F protein-mediated fusion (Russell et al., 2001). Coimmunoprecipitation assays suggest that F and HN exist in a complex, and a great deal of effort has been spent to map the regions of F and HN that interact. Mutations have been identified in several domains of HN that decrease or abolish fusogenic activity with no effect on receptor recognition (reviewed in Colman and Lawrence, 2003). However, a consensus view of the domains of interaction between HN and F has not been obtained to date.

HN is a type II integral membrane protein that spans the membrane once and contains an N-terminal cytoplasmic tail, a single N-terminal transmembrane (TM) domain, a membrane-proximal stalk domain, and a large C-terminal globular head domain (Hiebert et al., 1985). The globular head domain contains the receptor binding and enzymatic activity (Parks and Lamb, 1990; Scheid et al., 1972; Thompson and Portner, 1987). HN is glycosylated and noncovalently associated to form a tetramer, based on biochemical, crosslinking, and electron microscopy studies, which, depending on the virus, can be composed of two disulfide-linked dimers (McGinnes et al., 1993; Ng et al., 1989, 1990; Thompson et al., 1988). The covalent linkage occurs through a cysteine residue in the stalk domain. The specific residues involved in tetramer association have not been identified but have been suggested to reside in the TM domain, the cytoplasmic domain, the extracellular domain, or in all three domains (Parks and Lamb, 1990; Takimoto et al., 1992; Thompson et al., 1988).

The structure of the enzymatically active head domain of HN had been predicted to be similar to other neuraminidases or sialidases, such as influenza NA (Epa, 1997), with the globular head composed of identical subunits arranged with 4-fold symmetry. Each NA domain was expected to exhibit the six-blade-propeller fold typical of other neuraminidase/sialidase structures from viral, protozoan, or bacterial origin (Buschiazzo et al., 2000, 2002; Langedijk et al., 1997; Taylor, 1996). The predicted structure of the HN NA domain was confirmed in the X-ray structures of soluble NDV and HPIV3 HN globular NA regions (Crennell et al., 2000; Lawrence et al., 2004). While the NDV and HPIV3 HN-soluble NA domains are monomeric in solution, both form similar dimer interactions in their respective crystals. A potential tetrameric arrangement for the NDV HN NA domains was recently proposed, and a second sialic acid binding site at the previously observed dimer interface was identified (Zaitsev et al., 2004).

Here, we report the structure of the tetrameric ectodomain of SV5 HN and crystallographic studies with sialic acid, the inhibitor DANA (2,3-dehydro-2-deoxy-N-acetylneuraminic acid), and the substrate/receptor tri-saccharide sialyllactose. Sialyllactose is observed as...
an intact molecule in the active site, revealing an authentic receptor complex. Crystals soaked with sialic acid reveal electron density consistent with the inhibitor DANA, suggesting that SV5 HN can catalyze this reaction, as observed for other neuraminidases. SV5 HN forms dimers that are very similar to the previously described NDV and HPIV3 dimers, suggesting that this oligomeric arrangement is well conserved. SV5 HN also forms a potential tetramer in the crystals, through the association of two dimers. The observed tetramer arrangement is analogous to that described for NDV HN, but with significant shifts in dimer packing angles and residue positions. Implications from these structures for understanding the coupling of receptor binding and membrane fusion are discussed.

Results and Discussion

Structure Determination

Both the SV5 HN globular NA domain (HN_{NA}), residues 115–565, and the entire HN ectodomain (HN_{ecto}), residues 37–565, were expressed in insect cells by using the baculovirus system (Figure 1). Secreted HN protein was purified to homogeneity by using metal chelate and gel filtration chromatography. Purified HN_{NA} migrates as a monomer, as analyzed by both gel filtration chromatography and SDS-PAGE. In contrast, purified HN_{ecto} elutes from a gel filtration column with an apparent molecular weight of approximately 265 kDa, close to the predicted tetramer molecular weight (250–260 kDa). In reducing SDS-PAGE gels, HN_{ecto} migrates with the expected monomer molecular weight of 58 kDa, while, on nonreducing SDS-PAGE, HN_{ecto} migrates quantitatively at the predicted dimer molecular weight (Figure 1B). These observations indicate that HN_{ecto} forms disulfide-linked dimers through C111 and that the dimers associate into tetramers in solution, consistent with previous studies of the HN oligomeric arrangement (Ng et al., 1990). The formation of a stable HN tetramer in solution requires the N-terminal stalk region of the protein. Gel filtration analysis of HN_{ecto} at concentrations from 30 nM to 2 μM show no changes in the elution peak of the protein, suggesting that the tetrameric arrangement is stable over this concentration range. Both HN_{ecto} and HN_{NA} exhibit NA activity.

SV5 HN_{ecto} (henceforth referred to as SV5 HN) was crystallized (Table 1), and the structure was solved by molecular replacement (Figure 2). SDS-PAGE analysis of dissolved crystals (Figure 1B) shows that the intact HN_{ecto} protein is present in the crystals. Complexes with sialic acid (at pH 7.0 and 8.0), with the inhibitor DANA (at pH 7 and 8), and with sialyllactose (at pH 8) have also been determined. The resolution of the crystal structures ranges from 2.3 to 2.8 Å; final refinement statistics were collected and are shown in Table 1.

The globular NA domain of SV5 HN (residues 118–565) exhibits the typical sialidase/neuraminidase fold consisting of six antiparallel β strands organized as a superbarrel, with a centrally located active site (Figure 2A). In all of the structures, the stalk region (residues 37–117) is not visible in electron density maps, although some additional electron density is observed extending from residue 118 that could not be modeled. These observations suggest that the HN stalk is unstructured or adopts multiple conformations in the crystal. The HN structure contains seven disulfide linkages. In a comparison of 16 HN sequences, 4 of these are between absolutely conserved cysteines (C255–C285, C227–C240, C448–C458, and C528–C539). The remaining three are less well conserved. C175–C236 and C333–C454 are conserved in 10 of 16 sequences, but they are not observed in HPIV3 HN. C365–C375 is not observed in either HPIV3 or NDV sequences, but it is still found in 10 of 16 HN sequences.

The electron density for residues 186–190 (Figure 2) is variable in the data sets collected. Good electron density for these amino acids is observed in the complexes with sialyllactose soaked at pH 8.0, sialic acid soaked at both pH 7.0 and pH 8.0, and DANA soaked at pH 8.0. However, for the unliganded crystals at pH 8.0 and for DANA soaked at pH 7, the electron density for this region is poor. Residues 186–190 form a loop located on the top of the active site (Figure 2) and contain a highly conserved aspartic acid (D187, Table 2). In SV5 HN, this loop appears to be flexible, but this flexibility is not simply correlated with ligand binding or pH. It is not clear whether loop flexibility is important for HN function, but its conformation differs in the SV5, NDV (Crennell et al., 2000), and HPIV3 (Lawrence et al., 2004) HN structures. In NDV HN, mutation of D198 (D187 in SV5) affects receptor binding, NA activity, and fusion promotion (Deng et al., 1999). D198 is ordered in the NDV unliganded and sialic acid bound structures, but it is disordered in the inhibitor complex (Crennell et al., 2000).

In the SV5 HN structures, electron density is observed at predicted N-linked glycosylation sites (residues 139, 267, and 504) and for a calcium ion at a conserved binding site. The Ca^{2+} is coordinated by the carbonyl oxygen of D250, the hydroxyl and carbonyl oxygens of S253, the carbonyl oxygen of A285, and the carbonyl oxygen of A255.

Comparison of the SV5, NDV, and HPIV3 HN Monomer Structures

Superposition of the SV5 HN monomer onto NDV HN and HPIV3 HN monomers over 447 and 431 equivalent Cα positions (Figure 2C) yields rmsds in Cα positions of 1.2 Å² for NDV and 1.9 Å² for HPIV3. A structure-based sequence alignment of SV5, NDV, and HPIV3 (Figure 1C) indicates that SV5 HN has higher similarity to NDV HN (30%) and lower similarity to HPIV3 HN (21%). Overall, there are 79 conserved residues across all 3 sequences, corresponding to ~15% of the extracellular region. The major structural differences appear to be near insertion and deletion sites (Figure 1C) that are located at surface regions in the NA domain.

Many of the largest structural differences in the SV5, NDV, and HPIV3 HN proteins cluster to one side of the NA domain, forming a highly variable face (Figures 2C and 2D). In particular, differences in residues in the N-terminal 133 region, the α3-α4 helical region, the 424 loop, and the 478 loop define a markedly different broad surface in each of the HN proteins, with a number of potential N-linked glycosylation sites. Both the high variability and location of potential N-linked glycosylation sites suggest that this surface is not critical for
conserved functions of HN, such as interactions with F, and may instead represent adaptive changes associated with evasion of the immune response. However, it is possible that changes in this highly variable face or some of the other surface regions could influence the specificity of HN/F interactions.

Complexes of HN with Ligands
Ligand complexes were obtained by soaking crystals with sialic acid, DANA, and sialyllactose at the different pH values and time points listed in Table 1. Good electron density for ligands within the active sites was observed for all soaks (Figure 3). Comparison of the
Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Complexes</th>
<th>Sialic Acid (pH 7.0)</th>
<th>Sialic Acid (pH 7.0)</th>
<th>DANA (pH 7.0)</th>
<th>DANA (pH 8.0)</th>
<th>Sialyllactose (pH 8.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unliganded (pH 7.0)</td>
<td>2.6</td>
<td>2.5</td>
<td>2.8</td>
<td>2.3</td>
<td>2.7</td>
</tr>
<tr>
<td>Complexes</td>
<td>Unliganded (pH 7.0)</td>
<td>2.5</td>
<td>2.6</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td>a = b (Å)</td>
<td>137.6</td>
<td>137.9</td>
<td>137.7</td>
<td>137.8</td>
</tr>
<tr>
<td></td>
<td>c (Å)</td>
<td>184.9</td>
<td>184.8</td>
<td>184.9</td>
<td>184.9</td>
</tr>
<tr>
<td>Completeness/last shell (%)</td>
<td>96.5/96.6</td>
<td>92.0/98.3</td>
<td>100.0/100.0</td>
<td>98.1/98.7</td>
<td>83.4/95.4</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.8</td>
<td>6.0</td>
<td>7.2</td>
<td>5.4</td>
<td>7.1</td>
</tr>
<tr>
<td>Rmerge (last shell) %</td>
<td>9.8 (42.7)</td>
<td>7.8 (41.1)</td>
<td>11.9 (39.4)</td>
<td>7.9 (43.3)</td>
<td>8.3 (34.6)</td>
</tr>
<tr>
<td>Data Collection Statistics</td>
<td>Number of reflections</td>
<td>30,405/926</td>
<td>34,874/1,013</td>
<td>25,368/725</td>
<td>44,732/1,308</td>
</tr>
<tr>
<td></td>
<td>Number of protein atoms</td>
<td>3,449</td>
<td>3,491</td>
<td>3,491</td>
<td>3,461</td>
</tr>
<tr>
<td></td>
<td>Number of ligand atoms</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Number of water molecules</td>
<td>192</td>
<td>203</td>
<td>192</td>
<td>217</td>
</tr>
<tr>
<td></td>
<td>Average B factor</td>
<td>30.4</td>
<td>34.9</td>
<td>27.6</td>
<td>33.6</td>
</tr>
<tr>
<td></td>
<td>Protein</td>
<td>29.0</td>
<td>34.0</td>
<td>26.4</td>
<td>32.6</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>34.3</td>
<td>35.5</td>
<td>31.4</td>
<td>36.1</td>
</tr>
<tr>
<td></td>
<td>Rmsd bonds (Å)</td>
<td>0.007</td>
<td>0.006</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>Rmsd angles (°)</td>
<td>1.47</td>
<td>1.43</td>
<td>1.45</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>Rwork / Rfree, (%)</td>
<td>21.6/23.6</td>
<td>20.3/22.6</td>
<td>19.9/22.4</td>
<td>20.6/22.6</td>
</tr>
</tbody>
</table>

Ramachandran Statistics

| Favorable (%) | 85.9 | 85.7 | 84.9 | 85.7 | 84.6 | 84.9 |
| Additional (%) | 13.5 | 13.8 | 14.6 | 13.8 | 14.6 | 14.6 |
| Generous (%) | 0.5 | 0.3 | 0.5 | 0.5 | 0.8 | 0.5 |
| Disallowed (%) | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 |

active sites of the ligand bound and free SV5 HN structures reveals no major changes of the residue positions upon ligand binding (Figure 4), in contrast to observations made for both the NDV and HPIV3 HN structures. In this way, the SV5 site appears to be more similar to the rigid active sites observed for influenza A and B neuraminidases.

The Unbound Active Site

Seven highly conserved active site residues in neuraminidases and sialidases (R163, D187, E390, R405, R495, Y523, and E544) are shown in Table 2; the numbers corresponding to the SV5, NDV, and HPIV3 HN sequences are also shown. The three conserved arginines, R163, R405, and R495, are positioned to interact with the sialic acid carboxylate, even in the absence of ligand. Two water molecules are observed at positions corresponding to the substrate carboxylate oxygens, and these water molecules form a hydrogen bonding network with the conserved arginine triad (Figures 4A and 4B). Residues Y523 and E544 are positioned at the base of the active site. In complexes with ligand, these two residues are positioned underneath the glycosidic bond to be cleaved and participate directly in catalysis. E544 is located behind R163, forming a bidentate salt bridge interaction between the carboxylate oxygens and two of the guanidinium group nitrogens. There is no evidence for a bridging water molecule between R163 and E544, as observed in the HPIV3 and NDV HN structures. Instead, the tight R163:E544 interaction corresponds more closely to influenza A and B NA structures and may partially explain the apparent rigidity of the SV5 HN active site as compared to NDV and HPIV3 HN. Residues 186–190, including the conserved D187, are disordered.

Complex Soaked with DANA

Electron density maps from DANA-soaked crystals under both pH conditions show the bound inhibitor (Figure 3A). The active site is open to solvent with DANA positioned at the bottom of a large cavity. Twelve residues, comprising a total surface area of 232 Å², are buried by DANA binding. The buried residues include 5 of the 7 conserved active site residues, R163, E390, R405, R495, and Y523 (Figures 4C and 4D). E544, located behind R163, is already substantially buried from solvent, while D187 is disordered in this complex. In addition to the conserved residues, I164, S226, F241, and 2 residues are positioned underneath the glycosidic bond to be cleaved and participate directly in catalysis. E544 is located behind R163, forming a bidentate salt bridge interaction between the carboxylate oxygens and two of the guanidinium group nitrogens. There is no evidence for a bridging water molecule between R163 and E544, as observed in the HPIV3 and NDV HN structures. Instead, the tight R163:E544 interaction corresponds more closely to influenza A and B NA structures and may partially explain the apparent rigidity of the SV5 HN active site as compared to NDV and HPIV3 HN. Residues 186–190, including the conserved D187, are disordered.

The Unbound Active Site

Seven highly conserved active site residues in neuraminidases and sialidases (R163, D187, E390, R405, R495, Y523, and E544) are shown in Table 2; the numbers corresponding to the SV5, NDV, and HPIV3 HN sequences are also shown. The three conserved arginines, R163, R405, and R495, are positioned to interact with the sialic acid carboxylate, even in the absence of ligand. Two water molecules are observed at positions corresponding to the substrate carboxylate oxygens, and these water molecules form a hydrogen bonding network with the conserved arginine triad (Figures 4A and 4B). Residues Y523 and E544 are positioned at the base of the active site. In complexes with ligand, these two residues are positioned underneath the glycosidic bond to be cleaved and participate directly in catalysis. E544 is located behind R163, forming a bidentate salt bridge interaction between the carboxylate oxygens and two of the guanidinium group nitrogens. There is no evidence for a bridging water molecule between R163 and E544, as observed in the HPIV3 and NDV HN structures. Instead, the tight R163:E544 interaction corresponds more closely to influenza A and B NA structures and may partially explain the apparent rigidity of the SV5 HN active site as compared to NDV and HPIV3 HN. Residues 186–190, including the conserved D187, are disordered.
R405 N12 at distances of 2.6 Å and 3.1 Å, respectively.

Beneath the ligand, Y523 and E390 hydrogen bond to each other, with Y523 making additional potential hydrogen bonds to the ligand ring oxygen (O6) and R405. At the other end of the ligand, N5 forms hydrogen bonds to a water molecule located underneath DANA that is in a tight network with S226, Y306, and E390. This water molecule is also observed in the free HN structure. The glycerol moiety of DANA fits into a side pocket of the active site formed by E247, Y251, Y306, F353, E390, and R405. E247 makes two hydrogen bonds to O7 and O9 of DANA. O8 of the glycerol moiety points down into the active site, forming a hydrogen bond with a buried water molecule involved in a network of interactions that includes the E390 main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407. The E544 carboxylate main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407. The E544 carboxylate main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407.

The glycerol moiety of DANA fits into a side pocket of the active site formed by E247, Y251, Y306, F353, E390, and R405. E247 makes two hydrogen bonds to O7 and O9 of DANA. O8 of the glycerol moiety points down into the active site, forming a hydrogen bond with a buried water molecule involved in a network of interactions that includes the E390 main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407. The E544 carboxylate main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407. The E544 carboxylate main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407. The E544 carboxylate main chain nitrogen, another buried water, the carbonyl of G388, and the side chain of N407.

Complex Soaked with Sialic Acid
While the reaction mechanism of viral NAs is not completely understood, hydrolysis is thought to involve the induction and stabilization of an oxocarbonium ion intermediate (Janakiraman et al., 1994) and involve proton donors, which may be either protein residues (Ghate and Air, 1998) or water molecules (Taylor and von Itzstein, 1994). For the sialic acid complexes, we observed electron density consistent with the inhibitor DANA at both pH 7 and pH 8 (Figure 3B). Electron density for the O2 atom is not evident, and the sialic acid moiety appears to adopt a planar conformation, consistent with either the direct synthesis of DANA from sialic acid (Burmeister et al., 1993) or the formation of an oxocarbonium intermediate in the active site (Janakiraman et al., 1994). The analysis of the substrate density suggests that SV5 HN, like NDV HN and influenza B virus NA (Burmeister et al., 1993; Crennell et al., 2000), can synthesize DANA from sialic acid. For influenza B virus NA, the conversion of sialic acid to DANA has been shown to be a slow process, resulting from the high concentrations of both NA and sialic acid present in the crystallographic experiments, and it is not likely to affect NA activity in vivo. Overall, the structures of sialic acid-soaked HN crystals do not reveal any major differences compared to the DANA-soaked complexes.

Complex Soaked with Sialyllactose
Crystals were soaked with either a mixture of α2,3 and α2,6-sialyllactose, or only α2,3-sialyllactose (Neu5Acα2 → 3 Galβ1 → 4Glc). The intact sialyllactose ligand, representing a model receptor:HN complex, is observed. HN\textsubscript{ecto} enzymatic activity is optimal at low pH (4.0–6.0), but it drops off at pH 6.5–7.0 and is down by over 5-fold at pH 8 in solution. The pH of these soaking experiments (Scheid and Choppin, 1974). HN\textsubscript{ecto} NA activity is also further inhibited (~2x) by 250 mM NaCl present in the crystallization buffer. Under all soaking conditions, we observed electron density con-
Figure 3. Electron Density Observed for Ligand-Soaked Crystals

(A) DANA electron density at 2.3 Å resolution.
(B) Sialic acid-soaked crystal (modeled as DANA) at 2.5 Å resolution.
(C) α2,3-sialyllactose at 2.5 Å resolution. Stereo views of the electron density for the ligands from composite omit 2Fo −Fc maps contoured at 0.9 σ are shown. The figures were generated with Pymol (DeLano, 2002).

To test whether residual NA activity in the crystal could convert the α(2,3')-sialyllactose to a mixture of product and substrate during the time course of the soaking experiments, we collected data sets for crystals soaked from 2 hr up to 3 days (Table 3). For all of the crystals, the observed electron density was similar, suggesting that the enzymatic turnover rate was slow enough to prevent conversion of the substrate pool over this time.

The observed sialyllactose electron density is consistent with a deformed conformation for the saccharide (Figure 3C), similar to the structure of a Michaelis complex observed for a Trypanosoma cruzi trans-sialidase and other glycosidases (Amaya et al., 2004). In this conformation, the α(2,3') glycosidic bond to be cleaved is oriented in a pseudo-axial position, while the carboxylate moiety projects equatorially toward the conserved arginine triad (Figures 4E and 4F). The HN complex likely represents a distorted state prior to the authentic transition state (Ghate and Air, 1998) or oxocarbonium ion intermediate (Janakiraman et al., 1994).
Figure 4. Comparison of Active Sites in the Absence and Presence of Ligands
(A–F) Active sites are shown in two views, rotated 90° from each other. Surrounding protein side chains and water molecules are shown, with waters as spheres and hydrogen bonds as blue dotted lines. (A and B) Unliganded active site. (C and D) DANA complex. (E and F) Sialyllactose complex.

The lactose moiety is positioned around the α(2,3) linkage to one side of the large active site cavity close to W352. Fifteen HN residues, with a total buried surface of 329 Å², are buried by the receptor. Only 107 Å² of the active site are buried by the larger lactose moiety, consistent with its limited interactions in the cavity. Compared to the complex with DANA, only 3 additional residues—Q186, W352, and L459—are buried by sialyllactose, although the lactose does bury additional surface area of a subset of the other 12 residues. For example, while DANA buries 15 Å² of R495, in the sialyllactose complex, 32 Å² of this residue is buried. The lactose moiety is positioned above the end of the side chain of R495. The electron density in composite omit maps is best for the sialic acid region and is progressively weaker for the galactose and glucose moieties, consistent with the fewer contacts made by the sugar as it extends out of the active site (Figure 3C). This difference is also observable as a gradient of increasing temperature factors along the polysaccha-
Table 3. Sialyllactose Soaks

<table>
<thead>
<tr>
<th>Mixture (Concentration, mM)</th>
<th>Time (hr)</th>
<th>Resolution (Å)</th>
<th>$R_{merge}$ (%)</th>
<th>Density for 2,3'-SL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixture (10)</td>
<td>16</td>
<td>2.5</td>
<td>7.8</td>
<td>+</td>
</tr>
<tr>
<td>2,3'-sialyllactose (15)</td>
<td>2</td>
<td>3.0</td>
<td>13.5</td>
<td>[ ]</td>
</tr>
<tr>
<td>2,3'-sialyllactose (15)</td>
<td>16</td>
<td>3.0</td>
<td>10.2</td>
<td>+</td>
</tr>
<tr>
<td>2,3'-sialyllactose (15)</td>
<td>72</td>
<td>2.95</td>
<td>11.0</td>
<td>+</td>
</tr>
</tbody>
</table>

ride chain, with values near 30 Å$^2$ for sialic acid and near 75 Å$^2$ for the terminal glucose moiety.

Overall, the sialic acid moiety interacts with residues of the active site similarly to the DANA complex, and only minor adjustments of active site residues are observed. The conserved arginine triad interacts with the sialic acid carboxylate. In sialyllactose, the C2 carbon of sialic acid adopts a tetrahedral geometry and therefore projects slightly higher from the base of the active site above Y523 and E390. Y523 lies underneath the sialic acid and is stabilized by hydrogen bonds with E390 and R405, in addition to a potential hydrogen bond with the sialic acid ring oxygen (O6). The hydrogen bond networks around the glycerol moiety and N5 of the N-acetyl group are also observed in this complex. E544 and R163 retain their bidentate interaction, as observed in all of the HN structures.

Comparison of HN Dimers
SV5 HN forms a disulfide bonded dimer based on non-reducing SDS-PAGE. Although the crystals contain a monomer in the asymmetric unit, crystallographic symmetry generates HN dimers that are similar to those previously described for NDV and HPIV3 HN proteins. In the electron density maps, the N-terminal residues of the SV5 HN monomers extend toward the dimer interface, potentially meeting at A116. C111 (Figure 1) in each monomer could then form an interchain disulfide bond between these monomers, consistent with the observed oligomeric state in solution. The calculated buried surface area for each monomer in the SV5 HN dimer is 1810 Å$^2$. Superposition of SV5 HN with NDV HN and HPIV3HN (Figure 5) shows that these proteins all form similar dimers with rmsds in C$\alpha$ positions of 1.4 Å$^2$ and 1.9 Å$^2$, respectively. The dimer places the two HN active sites at nearly 90° angles relative to each other, with a 2-fold axis oriented approximately 45° from the top of the NA domain.

The highly variable surface of the HN monomers is located on the outer edges of the long axis of the dimer (Figure 5), while 47 residues (Figure 1) of more structurally conserved regions are involved in the dimer interaction. Four prolines interact at the center of the dimer interface; of these, P149 and P153 from both chains and aromatic residues (F194, F209, and Y217) form a hydrophobic cluster on one side of the prolines.

There is no evidence from the SV5 HN complex structures that the dimer interface changes upon ligand binding, as suggested by structural studies of the NDV HN (Crennell et al., 2000). The SV5 HN crystals do not crack upon ligand soaking, at either pH studied, suggesting that binding does not significantly alter HN dimerization and packing interactions.

SV5 Lacks a Second Binding Site for Sialyllactose
Recently, a potential second binding site for sialic acid in NDV HN, located at the dimer interface, was described (Zaitsev et al., 2004). In our experiments, 15 mM sialyllactose (Table 3), similar to the concentrations used to identify the NDV second binding site with a sialyllactose analog, was soaked into crystals. However, no electron density for this second site is observed in the SV5 HN complexes, suggesting that the second binding site is nonfunctional in SV5. The proposed second binding site is located near a solvent channel in the SV5 HN crystals and is not blocked by crystal packing.

In the NDV HN structures, the second sialic acid binding site is formed by four polypeptide segments from the two HN monomers (Figure 6A). In monomer 1, residues 514–519 and 155–156 make up two of the segments, and, in monomer 2, residues 553 and 168–
The SV5 HN tetramer is the only crystallographic arrangement for a tetramer consistent with the HN dimer (Figures 7A and 7B), and it places the N termini of the four monomers on one face, consistent with anchoring to a membrane surface. In addition, the variable face of the HN NA domain is positioned to the outside of the tetramer, consistent with a common oligomer arrangement for diverse paramyxovirus HN proteins. However, unlike typical protein tetramers, the SV5 HN tetramer is arranged with two 2-fold symmetry axes (Figures 7A–7C) that are not oriented at 90° to each other, allowing neighboring dimers and tetramers to associate into infinitely long oligomers. Because of this, we describe the HN tetramer as a dimer-of-dimers lacking closed symmetry.

While the observed SV5 HN tetramer can in principle form infinite helical ribbons because of its open symmetry, the interaction between NA domain dimers is weak. HNecto migration on gel filtration columns does not change at concentrations of up to ~2 μM, indicating that any interactions between tetramers are substantially weaker than this. In a lipid membrane, the assembly of multiple neighboring tetramers into a longer oligomer, as observed in the crystal, is not compatible with the anchoring of HN to the cell membrane, and this also likely contributes to limiting the physiological oligomerization of the protein. Based on EM images, Sendai virus HN, like influenza virus NA, exhibits a box-shaped arrangement consistent with four discrete subunits (Thompson et al., 1988; Varghese et al., 1983).

169 make up the final two segments. The majority of hydrogen bonds between sialic acid and HN are formed with monomer 1, with only a single hydrogen bond formed by G169 in monomer 2 (Zaitsev et al., 2004).

Superposition of the SV5 dimer onto the NDV dimer reveals a number of structural changes that would significantly disrupt the second sialic acid binding site. The sialic acid moiety at the NDV HN second site interacts with nonconserved residues (Figure 6B). In SV5 HN, the NDV G169 residue in monomer 2 is replaced by P158, which cannot hydrogen bond to the ligand. In monomer 1, SV5 residues 510–515 and 144–145, corresponding to the two interacting segments in NDV HN, are displaced such that additional hydrogen bond interactions with the sialic acid would be disrupted. The observed structural differences are consistent with a lack of electron density in the sialyllactose-soaked SV5 HN complexes and suggest that this second site is not functional in SV5.

**SV5 HN Tetramers**

Based on gel filtration chromatography, SV5 HN likely forms a tetramer in solution. In the HN crystal lattice, the SV5 HN NA domain forms a continuous ribbon of dimers and tetramers aligned along the crystallographic 6-fold axis. The potential tetramer is the only crystallographic arrangement for a tetramer consistent with the HN dimer (Figures 7A and 7B), and it places the N termini of the four monomers on one face, consistent with anchoring to a membrane surface. In addition, the variable face of the HN NA domain is positioned to the outside of the tetramer, consistent with a common oligomer arrangement for diverse paramyxovirus HN proteins. However, unlike typical protein tetramers, the SV5 HN tetramer is arranged with two 2-fold symmetry axes (Figures 7A–7C) that are not oriented at 90° to each other, allowing neighboring dimers and tetramers to associate into infinitely long oligomers. Because of this, we describe the HN tetramer as a dimer-of-dimers lacking closed symmetry.

While the observed SV5 HN tetramer can in principle form infinite helical ribbons because of its open symmetry, the interaction between NA domain dimers is weak. HNecto migration on gel filtration columns does not change at concentrations of up to ~2 μM, indicating that any interactions between tetramers are substantially weaker than this. In a lipid membrane, the assembly of multiple neighboring tetramers into a longer oligomer, as observed in the crystal, is not compatible with the anchoring of HN to the cell membrane, and this also likely contributes to limiting the physiological oligomerization of the protein. Based on EM images, Sendai virus HN, like influenza virus NA, exhibits a box-shaped arrangement consistent with four discrete subunits (Thompson et al., 1988; Varghese et al., 1983).

In contrast to the HN dimer interaction, the dimer-of-dimers interface is much smaller, involving only ten residues and burying only 657 Å². The buried residues are primarily located on loop regions that project toward the dimer-of-dimers interface, including residues near the N terminus of the NA domain (121, 122, 124, 125), the 204 loop (204–205), the 533 loop (532–534), and K171. These regions are not the most variable regions of the NA domain, but they are also not particularly well conserved in the different HN proteins (Figure 1C), suggesting that the arrangement is not a very strong or highly conserved interaction. However, the SV5 HN tetramer is similar to a proposed tetramer arrangement observed for the NDV HN monomer in multiple crystal forms (Zaitsev et al., 2004). The SV5 and NDV HN tetramers do show significant differences in the dimer-of-dimers packing, with one dimer shifted and rotated compared to the other (Figure 7C). This shift could be due to differences in the HN proteins, due to the differing crystal lattices or conditions, or could indicate that tetramerization of the NA domain is generally weak and adaptable. There are no analogous tetramers formed in the reported structures for HPIV3 HN.

**Conclusions**

We have determined and analyzed the crystallographic structure of the SV5 HN ectodomain alone and in complexes with the inhibitor DANA and an authentic receptor, sialyllactose. Sialyllactose was found to bind to the active site in an intact form at pH 8.0, and the sialic acid moiety exists in a distorted conformation induced by interactions with HN. While structural studies of the HN NA domains from NDV (Crennell et al., 2000; Zaitsev et al., 2004) do not provide the same information about the HN ectodomain, the overall structural arrangement and the binding site for sialyllactose are well conserved.
et al., 2004) and HPIV3 (Lawrence et al., 2004) have been reported, these have not included complexes with a model sialic acid receptor such as sialyllactose. In addition, the SV5 HN ectodomain protein exists as a tetramer in solution, in contrast to the previous HN structures. The presence of the HN stalk region is critical for the formation of stable HN tetramers. However, in the SV5 HN crystal structures, the N-terminal stalk region is not visible in electron density maps, suggesting that the connection between the stalk and NA domains is flexible.

Based on the existing structures and biochemical evidence, it remains unclear whether the hemagglutinin and NA activities of HN involve one or two separate sialic acid binding sites (Portner et al., 1987; Thompson and Portner, 1987; Zaitsev et al., 2004). The disparate theories of one site with dual functions, or of two distinct sites that are intimately related, are both consistent with the observation that sialic acid-derived NA inhibitors also interfere with receptor binding (Iorio et al., 2001; Murrell et al., 2003; Scheid and Choppin, 1974). A single site can provide both hemagglutinin and NA activities by binding sialic acid tightly and hydrolyzing the molecule slowly (Scheid and Choppin, 1974). The SV5 HN structure with intact sialyllactose bound in the active site, but not at a second putative receptor binding site observed for NDV HN, provides further support for a single site model for HN action.

Structural studies of NDV HN (Crennell et al., 2000) suggested that the NA domain could form two distinct dimeric assemblies that were hypothesized to be ligand dependent. One of the dimers, observed after cocrystallization with ligand, formed an extensive buried interface, while the second dimer, crystallized in the absence of ligand and at low pH, formed a much smaller interface. Conformational changes were observed in the active site of the HN protein upon ligand binding that were correlated with changes in the dimer interface, suggesting a possible mechanism for coupling ligand recognition to changes in the oligomeric assembly of the HN protein. A model for the mechanism of the HN action in membrane fusion was proposed in which HN engagement of sialic acid through the active site leads to dimerization (and potentially tetramerization) of the NA domain, thus generating a second sialic acid binding site that would then function in adhesion. The formation of the HN dimer and tetramer would lead to release of F and its subsequent activation in membrane fusion.

Our structural studies of SV5 HN do not support the conclusion that there is a second sialic acid binding site common to HNs of different paramyxoviruses or that there are ligand-dependent oligomerization or conformational changes in the protein structure. Moreover, the dimer of HN that is observed in the SV5 (and HPIV3) HN structures occurs in the absence of ligand binding, and there is no crystallographic evidence that monomeric ligand binding influences the oligomeric arrangement for these HN proteins. The stalk region appears to be critical for stabilizing the tetramer in the absence of ligand, although ligand binding could in principle affect the HN oligomer. It remains unclear how changes in the HN NA domain may influence the HN stalk region, which has been implicated by functional studies to contain the primary sites for specificity of the HN/F interaction. Mutagenesis studies have also demonstrated that the stalk region of NDV HN can affect the NA activity and fusion promotion activity (Wang and Iorio, 1999).

While the mechanism by which HN binding to receptors can trigger F activation remains to be understood, the structural and functional studies of NDV, HPIV3, and SV5 HN proteins do demonstrate that HNs from different viruses form similar dimer and potentially tetrameric arrangements. The stabilization of the HN tetramer and NA domain interactions by the N-terminal stalk region suggests that rearrangement or partial dissoc-
ation of the HN NA domains may not require a significant amount of energy. Based on the structural analysis, it would appear that NA domain dimer interactions are of higher affinity, since the interface is larger and the arrangement is well conserved across three paramyxovirus HN proteins. The dimer of dimers appears to be a weaker interaction that is less well conserved and perhaps much easier to perturb with a lower energy penalty. The rearrangement of the HN dimer or tetramer, in particular through changes in the HN NA domain interactions, could be a critical feature of the F protein activation process.

We propose an alternative model for HN involvement in membrane fusion that is consistent with the available data and that involves ligand-dependent changes in the HN oligomer that are driven by surface-surface interactions (Figure 7D). In this model, the HN dimer/tetramer forms in the absence of ligands and can interact with the F protein, potentially through lateral interactions on two sides of the tetramer. Engagement of cell surface receptors could trigger the partial disassembly of the HN tetramer, assuming that the energy of binding of the individual HN active sites to distinct sialic acid receptors is sufficient to perturb the weak NA domain interactions. Opening of the tetrameric head, driven by the energy of receptor engagement, could lead to changes in both the HN stalk region and the interaction with F, thus activating F for membrane fusion.

Experimental Procedures

Construction of Expression Vectors

The HNecto and HNNA fragments were amplified separately by PCR from SV5 HN cDNA (Hiebert et al., 1985) and were cloned into the baculovirus transfer vector pBAGus-3 (Invitrogen) at the Smal/Xhol sites. Clones were identified and verified by DNA sequencing. The cloning strategy leaves three extra residues (SPG) after enterokinase cleavage at the N terminus of HNecto (residue 37) and HNNA (residue 115). Both recombinant proteins terminate at residue 565.

Insect Cell Expression

Baculovirus transfer vectors containing HNecto or HNNA inserts were cotransfected with Baculogold DNA (Pharmingen) into SF9 insect cells at 2 x 10^6 cells/ml grown in a T-25 flask with SF900 II serum-free medium (GIBCO-BRL) by using standard calcium phosphate conditions (Pharmingen). Virus stocks were produced by using Entero-kinase cleavage of the N terminus of HNecto (residue 37) and HNNA (residue 115). Both recombinant proteins terminate at residue 565.

Protein Preparation

Baculovirus-infected cells expressing HN were pelleted by centrifugation (13,000 x g, 30 min at 4°C), and one-tenth volume of 10 mM calcium, and the N-terminal fusion tags were cleaved by using enterokinase. Enterokinase was removed by using a capture column and dialyzed against enterokinase reaction buffer without calcium, and the N-terminal fusion tags were separated from the protein preparation (Novagen), and the cleaved fusion tags were separated from the HN proteins by repurification over Ni-NTA resin. Cleaved HNecto was further purified by using size-exclusion chromatography (Superdex-200) in 25 mM Tris (pH 8.0), 200 mM NaCl. Protein concentrations were determined by absorbance at 280 nm by using calculated extinction coefficients of 1.47 and 1.57 cm⁻¹(mg/ml)⁻¹ for HNecto and HNNA, respectively. The HNecto protein was concentrated to 10 mg/ml in 20 mM Tris (pH 7.4), 50 mM NaCl for crystallization.

Protein Characterization

Recombinant HN was separated by SDS-PAGE and was detected by staining with Coomassie brilliant blue or by immunoblotting. The HN protein was detected on blots by using His-tag-specific monoclonal antibody (Qiagen), followed by alkaline phosphatase-conjugated goat anti-mouse IgG. Detection was performed by using Enhanced Chemiluminescence (ECL) and was visualized by using a STORM phosphorimager (Amersham). Protein molecular weights were assessed by SDS-PAGE (reducing and nonreducing conditions) and by size exclusion chromatography. NA activities were measured by detection of released sialic acid from sialylactose by using the thiochorbic acid assay (Warren, 1959).

Crystallography and Data Collection

HNecto protein was crystallized at 18°C by the hanging drop vapor diffusion method. Drops containing 0.8 µl 10 mg/ml protein in 20 mM Tris (pH 7.4), 50 mM NaCl were mixed with the same volume of precipitant (10 mM Tris pH 8.0, 10% polyethylene glycol 2000 MME) and equilibrated over a well containing 600 µl precipitant solution. Crystals appeared in 5 days as needles and grew into long, thin hexagonal rods with maximum dimensions of 0.4 x 0.02 x 0.02 mm in 3 weeks. Crystals were harvested into a stabilizing buffer (10 mM Tris pH 8.0, 14% polyethylene glycol 2000 MME, 250 mM NaCl).

Harvested crystals could be transferred to pH 7 and soaked with ligands without any deterioration in diffraction. Crystals were flash-cooled in cryoprotectant solutions containing 25% polyethylene glycol 2000 MME, 250 mM NaCl, and 10% glycerol at either pH 7 or pH 8. When forming complexes, ligands were added into the stabilizing and cryoprotectant solutions at the following concentrations: 8 mM or 30 mM sialic acid, 8 mM DANA, and 10 mM or 15 mM sialyllactose (either a mixture of 2-3'- and 2-6'- or purified 2-3'-sialyllactose). Crystals were soaked in sialic acid or DANA overnight or in sialyllactose at the pH values and time points indicated in Table 3.

Diffraction data were collected at the COM-CAT beamline at the Advanced Photon Source by using a Mar CCD detector. Data were processed and scaled with the HKL suite of programs (Otwinowski and Minor, 1997). The crystals belong to space group P6_22 with unit cell dimensions (a = b = 137.5 Å, c = 184.8 Å) that remain similar for all data sets.

Structure Determination and Refinement

Molecular replacement was performed by using the SV5 HN unliganded data set and NDV HN (PDB code 1E8T) as a search model, providing a clear solution for a single monomer in the asymmetric unit. The structure was rebuilt and refined to crystallographic Rwork and Rref values near 30% at 2.6 Å by using the programs O (Jones et al., 1991) and CNS (Brünger et al., 1998). Further higher-resolution refinement (2.3 Å) was performed by using a DANA-soaked data set, using the same test reflections to 2.6 Å along with additional reflections between 2.6 and 2.3 Å, providing the initial refined structure for the remaining data sets. In order to test for any potential ligand or active site electron density bias from the DANA-refined model, composite omit electron density maps were also calculated in parallel with the partially refined native model for comparison. Geometric parameters of the refined model were analyzed with PROCHECK (Laskowski et al., 1993) and WHATCHECK (Hooft et al., 1996).

Acknowledgments

This research was supported in part by National Institutes of Health research grants GM-61050 (T.S.-J.) and AI-23173 (R.A.L.). R.A.L. is an investigator of the Howard Hughes Medical Institute.
Influenza B virus neuraminidase can synthesize its own inhibitor. Structure 1, 19–26.


Accession Numbers

The atomic coordinates and structure factors of the structures have been deposited in the Protein Data Bank under codes 1Z4V, 1Z4W, 1Z4X, 1Z4Y, 1Z4Z, and 1Z50.