JOURNAL OF VIROLOGY, June 1993, p. 2972–2980
0022-538X/93/062972-09$02.00/0

Vol. 67, No. 6

Sequence and Structure Alignment of Paramyxovirus Hemagglutinin-Neuraminidase with Influenza Virus Neuraminidase

P. M. COLMAN,[1]* P. A. HOYNE,[2] AND M. C. LAWRENCE[1]

Biomolecular Research Institute[1] and CSIRO Division of Biomolecular Engineering,[2] 343 Royal Parade, Parkville, Victoria, 3052 Australia

Received 16 December 1992/Accepted 24 February 1993

A model is proposed for the three-dimensional structure of the paramyxovirus hemagglutinin-neuraminidase (HN) protein. The model is broadly similar to the structure of the influenza virus neuraminidase and is based on the identification of invariant amino acids among HN sequences which have counterparts in the enzyme-active center of influenza virus neuraminidase. The influenza virus enzyme-active site is constructed from strain-invariant functional and framework residues, but in this model of HN, it is primarily the functional residues, i.e., those that make direct contact with the substrate sialic acid, which have identical counterparts in neuraminidase. The framework residues of the active site are different in HN and in neuraminidase and appear to be less strictly conserved within HN sequences than within neuraminidase sequences.

Neuraminidases (sialidases, EC 3.2.1.18) are found in a number of pathogenic microorganisms, including viruses, bacteria, and protozoa. Their biochemical role is to catalyze the cleavage of sialic acid from glycoconjugates (18). The biological significance of this event may vary depending on the microorganism, but generally it appears to be related to virulence, including the penetration of mucosal secretions (7, 8, 9), uptake of bacterial toxins (16), and release of virus from infected cells (5, 19, 36, 37).

The best characterized of these enzymes is that from influenza viruses, for which the primary structure from many different strains has been determined and several three-dimensional structures are also known (reviewed in reference 11). Among subtypes of influenza type A virus, neuraminidase sequences differ by about 50%, and influenza type B virus neuraminidase sequences differ from those of A strains by about 70% (11). The three-dimensional structures of two N2 strains (48, 49), one N9 strain (3, 46), and one B strain (6) have been reported, together with the structures of a number of so-called escape mutants, i.e., variants which differ from the parental strain by a single amino acid sequence change which renders them selectable by monoclonal antibodies (46, 51). All of these three-dimensional structures display the same folding motif, a so-called β-sheet propeller (49) comprising six four-stranded antiparallel β-strands connected internally by reverse turns and joined to each other by a connection between the outside strand of each sheet to the inside strand of the following sheet. The arrangement and twist of the sheets is reminiscent of the blades of a propeller (Fig. 1). The six β-sheets are referred to herein as β_i (i = 1,6), and the four strands within each sheet are labelled S_j (j = 1,4), reading in the sense of the polypeptide, i.e., from the center of the structure to the periphery. Loop structures are referred to as $L_{n,n+1}$, indicating connections between strands n and $n + 1$ within the sheet, with n = 0 indicating a connection from the preceding β-sheet. The neuraminidase structure may therefore be read in this terminology as: membrane anchor and stalk (not part of the head structure shown here), N-terminal arm, β_6S_4, β_1L_{01}, β_1S_1, β_1L_{12}, β_1S_2, β_1L_{23}, β_1S_3, β_1L_{34}, β_1S_4, β_2L_{01}, β_2S_2, β_2L_{12}, ···, β_6S_3, C-terminal arm at subunit interface (48, 49).

The folded structure of the polypeptide brings into close spatial proximity a number of amino acids which are invariant in all strains of influenza virus characterized to date (12). These amino acids line the walls of a pocket into which sialic acid and substrate analogs are observed to bind (6, 12, 50). Eight of these strain-invariant amino acids contact the substrate directly and are referred to here as functional (Fig. 2). Ten others appear to be important primarily for stabilizing the active-site structure and are referred to here as framework. Partitioning the site in this way is not a precise exercise (see the legend to Fig. 2). Some amino acids may be critical for function though not themselves in contact with substrate. Although no such amino acids have been characterized yet, E-277 could be one candidate through its proximity to Y-406 and a possible functional role that it may have in polarizing the phenolic hydroxyl group of that residue. In this respect, it may be more important than E-276, whose role in binding the eight and nine hydroxyls of the substrate could be played by another amino acid. A striking feature of the active site is the triarginyl cluster (R-118, R-292, and R-371) which encircles the carboxylate moiety of the bound sugar and which may contribute to the observed "boat" geometry of sialic acid bound to the enzyme (50). Binding of the sugar in this conformation to the neuraminidase is likely to be important for catalysis.

A mechanism for the enzyme action has been proposed on the basis of kinetic isotope methods, nuclear magnetic resonance, and molecular dynamics simulation (10) of the enzyme-substrate complex (50). The mechanism involves a sialosyl cation transition state complex, formed by acid attack on the glycosidic oxygen by a water molecule activated by D-151.

Paramyxoviruses also carry a neuraminidase activity on the hemagglutinin-neuraminidase (HN) glycoprotein, and primary structures of a number of these molecules are known, including representatives from Newcastle disease virus (NDV), mumps virus, Sendai virus, and parainfluenza viruses (reviewed in reference 34). The pH profile of the HN

* Corresponding author.

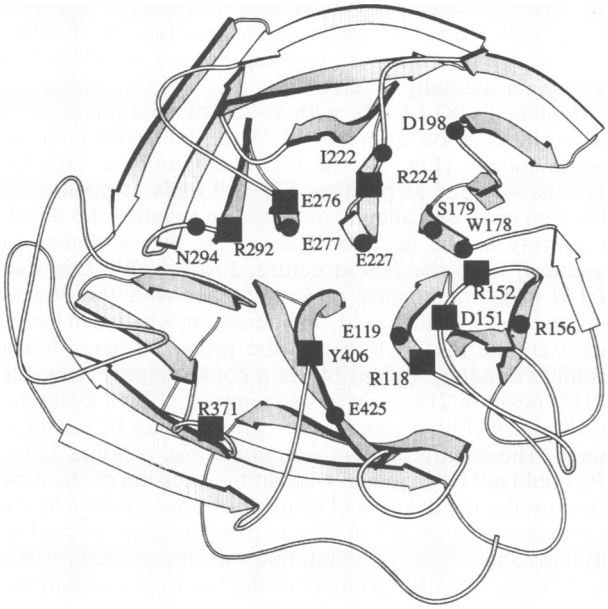

FIG. 1. Schematic of the chain tracing of the influenza virus neuraminidase polypeptide. Strain-invariant amino acids in the active site referred to in the legend to Fig. 2 are shown as solid squares (functional residues) or solid circles (framework residues). The sequence numbers of the first and last amino acids in each of the six sheets as drawn here are 118 and 175, 178 and 216, 226 and 268, 276 and 317, 352 and 398, and 406 and 103. The figure is based on a drawing produced by the MolScript program (29).

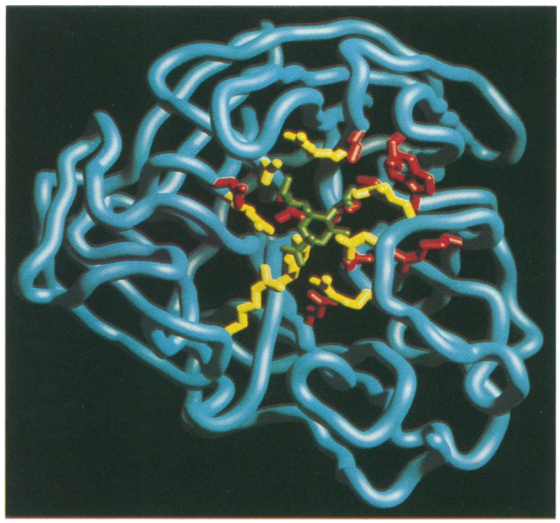

FIG. 2. Active-site structure of influenza virus type A neuraminidase (50). All amino acids shown are invariant across all known strains of types A and B influenza virus except for Asp-198, which is Asn in the N7 and N9 sequences (Fig. 3). Yellow residues (R-118, D-151, R-152, R-224, E-276, R-292, R-371, and Y-406) are defined as functional in the sense that they make contact with the reaction product sialic acid, shown in green. All of these interactions are polar in character, except that of R-224 in which the aliphatic component makes apolar contact with the lipophilic surface of the glycerol side chain. Red residues (E-119, R-156, W-178, S-179, D-198, I-222, E-227, E-277, N-294, and E-425) are defined as framework in the sense that they make no direct contact with the substrate but through interactions with the functional residues hold them in place for binding and catalysis. The distinction between these two types of residues is blurred in places, for example near the *N*-acetyl group binding residues, in which W-178 and I-222 are close to within the van der Waals contact with the bound sugar. The orientations of Fig. 1 and 2 are identical. The figure was generated with the HYDRASTER program, written by S. Watowich and L. Gross, based on work by D. Bacon and W. Anderson (RASTER3D) and R. Hubbard (HYDRA).

enzyme activity is bell shaped and peaks at around pH 4.5 (44), but these data do not allow an accurate comparison of the enzyme mechanism with that of influenza virus neuraminidase (10). No requirement for exogenous calcium ion for enzyme activity has been reported. The inhibitor 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acid is as effective against HN as it is against influenza virus neuraminidase (23), suggesting that both enzymes use a similar transition state in catalyzing cleavage, namely one in which the carboxylate of the sugar is equatorial to the pyranose ring (10, 50). Jorgensen et al. (26) have reported a homology between NDV HN and parts of the third and fourth β-sheets of influenza virus neuraminidase. The similarity of sequence is low, but secondary structure predictions for the regions in question are consistent with the claim. A shortcoming with their conclusion is that it identifies only a small fragment of the influenza virus neuraminidase active site in the HN sequence, yet if the overall structures are to be truly similar it would be expected that all of the binding site residues would be found within the HN sequence.

More recently, a number of bacterial and protozoal neuraminidase sequences have been determined and an aspartic acid box sequence motif common to influenza virus neuraminidases has been identified (27, 32, 38, 40). The motif is a short sequence of the consensus S-X-D-X-G-X-T-W, and it occurs four times in each of the bacterial neuraminidase sequences. However, the connection to the influenza virus (in β_6S_4, β_1L_{01}) enzyme is weak. The motif is found once in an N1 subtype sequence and twice in an N9 sequence (once as for N1 and again at β_5S_1, β_5L_{12}) (40), but in none of these examples is the consensus sequence strictly observed nor are significant elements of the motif conserved across other influenza virus neuraminidase subtypes. Furthermore, no active-site residues of the influenza virus enzyme lie within any of the identified motifs.

Comparative studies of protein and gene sequences are often not sufficiently sensitive to determine similarities in the three-dimensional structures of two protein molecules. In cases for which a functional similarity exists, such as among the neuraminidases, a more restricted search for elements of an active site might be undertaken. Even in such cases, convergent evolution may be operating, resulting in similar three-dimensional clustering of active-site elements supported on unrelated polypeptide backbones. One such example is provided by the four classes of enzymes which contain catalytic triads (35), viz., the eukaryotic serine proteases, the cysteine proteases, the subtilisins, and the α/β hydrolases.

Here we describe an alignment of influenza virus neuraminidase sequences with sequences from the HN glycoprotein of paramyxoviruses. The alignment extends the work of Jorgensen et al. (26) by including most of the amino acids known to be important for substrate binding in influenza virus neuraminidase and further suggests how bacterial neuraminidase sequences could be overlaid on the influenza virus neuraminidase structure so as to preserve functional elements of the active-site structure.

MATERIALS AND METHODS

Protein sequences were taken from the GenBank nucleotide data base (release 71) or the data base of the Protein Research Foundation of Japan (release 09/91) and aligned using the CLUSTAL program (22), with a gap penalty of 2 for pairwise sequence comparisons and a gap penalty of 10 for both fixed and varying gaps in multiple sequence comparisons. These alignments were not subsequently manually edited (with the exception of residues 609 to the C termini of the sequences shown in Fig. 4). The following 10 paramyxovirus HN protein sequences were extracted from GenBank, and the accession numbers are shown in parentheses: Simian virus 5 (K02870) (21), Sendai virus Z strain (X02808) (33), human parainfluenza virus type 1 (hPIV-1) (M31228) (17), hPIV-2 (X57559) (28), hPIV-3 strain Wash/1511/73 (M18759) (47), hPIV-4A (M34033) (4), NDV strain D26/76 (M24705) (43), NDV strain LAS/46 (M24709) (43), NDV strain CHI/85 (M24716) (43), and mumps virus (M19933) (55). Five influenza virus type A sequences and one type B neuraminidase sequence were extracted from GenBank; the accession numbers are shown in parentheses: N1, A/Puerto Rico/8/34 (J02146) (15); N2, A/Tokyo/3/67 (K01393) (30); N7, A/Cor/16/74 (M14916) (14); N8, A/Ken/1/81 (M14917) (14); N9, A/Tern/Australia/G70C/75 (M11445) (2); and B, B/Victoria/3/85 (M30639) (1). The sequence of an N5 neuraminidase was obtained from the data base of the Protein Research Foundation of Japan: N5, A/Shearwater/Australia/72 (Z1506542A) (20).

The alignments were examined for the presence of conserved amino acids which would preserve the sialic acid binding site observed in influenza virus neuraminidase. Initially, both the functional and framework components (see the legend to Fig. 2) of the neuraminidase active site were sought, but when it was apparent that that level of homology did not exist between the influenza virus neuraminidase and HN, the search focused primarily on functional residues.

RESULTS

The alignment of influenza virus neuraminidase subtype sequences is shown in Fig. 3. Amino acid sequence numbering is according to the sequence of the N2 subtype and follows that of Colman (11). The alignment of representative HN sequences is shown in Fig. 4. In seven places, concentrations of three invariant residues within a span of four residues are observed. Four of these seven regions can be mapped in order to active-site structures on influenza virus neuraminidase as shown in the boxes in Fig. 5, and the remaining three have plausible structural interpretations as described below. Although the individual alignments show only weak resemblances between HN and neuraminidase, the sequences are very conserved within HN and neuraminidase sequences. Furthermore, the two arginines at neuraminidase positions 118 and 371 and the glutamic acid at position 276, all known to be important elements of the neuraminidase structure for binding sialic acid, map to identical residues in HN. From the starting point of the boxed residues in Fig. 5, the following conclusions can be drawn.

D-230 (Fig. 4) on HN can be aligned with D-151 on neuraminidase, the putative catalytic aspartic acid (50). This requires the deletion of nine residues from HN with respect to neuraminidase, possibly within β_1L_{23}, but it allows for the formation of a disulfide bond in HN between residues 204 and 228, a bond which is plausible between the neuraminidase counterparts at residues 116 and 149. It should be stressed that the disulfide bonding pattern in HN has not been experimentally determined.

Aligning R-267 of HN with R-224 of neuraminidase requires the deletion from HN of 36 residues with respect to neuraminidase (Fig. 5). In the neuraminidase structure, sheet two begins at position 176 and ends at position 217 (48). The proposed alignment therefore requires the deletion in entirety of the neuraminidase sheet two sequence (not structure) from the HN structure. From within this sheet, W-178 in neuraminidase makes contact with the *N*-acetyl moiety of bound sialic acid, an interaction which can have no direct counterpart in the structure proposed here. Further disulfide bonding is plausible as a consequence. A cysteine at HN position 218 is always accompanied by a cysteine at position 279 (Fig. 4), and these residues may be covalently joined. The counterparts in neuraminidase, residues 130 and 236, could not be so joined without the excision of sheet two. Structurally, the deletion of sequence corresponding to sheet two means that R-267 of HN is now on sheet two of that structure, whereas its counterpart in neuraminidase is on sheet three. The displacement of the Cα atom resulting from such a change is on the order of 6 Å (0.6 nm).

The third alignment suggested in Fig. 5 requires an insertion of 114 residues in HN with respect to neuraminidase. Some conserved sequences are evident, and two disulfide bonds internal to this domain are also implied by the pattern of conserved cysteine residues. In this model, these 100 or so amino acids have no sequence counterpart in the influenza virus neuraminidase structure. We propose that they are the structural counterpart of neuraminidase sheet three. This structure may be associated with the hemagglutinating activity of the HN protein. It is interesting to note that this alignment would place the beginning of HN sheet three near residue 350 which is in one of the seven conserved sequence windows referred to above. It is common among the neuraminidase sequences shown in Fig. 3 to find similarly conserved windows near the beginning or within the first strand of each of the six β-sheets. By analogy, it could be argued that the conserved HN sequence near residue 350 is a marker for the beginning of sheet three.

Toward the C terminus from E-435, a conserved arginine at position 450 may be the third of the three carboxylate-binding residues observed at R-292 in neuraminidase. Four residues toward the C terminus, a (nearly) conserved tryptophan may be homologous to W-295 in neuraminidase.

Downstream from this tryptophan in HN is found the conserved sequence R-487-P-G. The arginine may correspond to the conserved arginine at position 327 in neuraminidase. In neuraminidase, that arginine is frequently followed by proline. This arginyl residue is buried in the neuraminidase structure and makes hydrogen bonds with four peptide carbonyl groups (48). A similar structural role is proposed here for R-487 in HN.

Residues 506 to 510 compose the last of the conserved sequence windows which do not relate to functional active-site residues. Previous and subsequent alignments suggest that residue 510 of HN corresponds approximately to residue 350 of neuraminidase. That residue in neuraminidase marks the beginning of the fifth β-sheet, and the alignment proposed is therefore consistent with the argument above concerning conserved sequences at the start of sheet structures. Two possible structural roles can be envisaged for HN D-510. As the counterpart of residue 350 in neuraminidase, it is within 6 Å of residue E-277 in neuraminidase and may

```
                  10        20        30        40        50        60        70        80
N2      MNPNQKIITIGSVSLTIATVCFLMQI-AILVTTVTLH---FKQHECDSPASNQVMPCEPIIIERNITEIVYLNNTTIEKEI----CPK--------VVEY
N1      MNPNQKIITIGSICLVVGLISLILQI-G----NIISI---WISHSIQTGSQNHTGICN------QNII-TYKNST-------WVKDTTSV---------I
N5      MNPNQKIITIGSASLGLVIFNILLHV-A----SITL-------GIISVTKDNKVHICN------TTEV--YNETVRVETVV--IPVNNTIYLNHEPE--F
N7      MNPNQKLFASSGIAIVLGIINLLIGI-SNMSLNISLY---SKG-ESH---KNNNLTCTNINQNDTTMVNTYINNATII-D-------KSTKIENPG---Y
N8      MNPNQKIIAIGSASLGILILNVILHV-V----SIIV-------TVLVLNNNGTGLNCN------GTIIREYNETVRVERITQWYNTNTIEYIERPSNEYY
N9      MNPNQKILCTSATALVIGTIAVLIGI-TNLGLNIGLH---LKP-SCN---CSHSQP--EATNASQTIINNYYNDTNIT-QI----SNTNIQVEERAIRDF
B       MLPS----TIQTLTLFLTSGGVLLSLYVSASLSYLLYSDILLKFSPKITAPTMTLDCTNASNVQAVNRSATKEMTFLLPEPEWT--------------Y
                * *

                   90       100       110       120       130       140       150       160       170       180
N2      RNWSKPQCQITGFAPFSKDNSIRLSA---GGDIWVTREPYVSCDPVKCYQFALGQGTTLDNKHSNDTVHDRIPHRTLLMNELGVPFHLGTRQV-CIAWSS
N1      LTGNSSLCPIRGWAIYSKDNSIRIGS---KGDVFVIREPFISCSHLECRTFFLTQGALLNDRHSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSA
N5      LNNTEPLCDVSGFAIVSKDNGIRIGS---RGHIFVIREPFVSCGPSECRTFFLTQGALLNDKHSNNTVKDRSPYRALMSVPLGSSPNAYQAKFESVGWSA
N7      LLLNKSLCNVEGWVVIAKDNAIRFGE---SEQIIVTREPYVSCDPLSCKMYALHQGTTIRNKHSNSTTHDRTAFRGLISTPLGSPPTVSNSEFICVGWSS
N8      MNNTEPLCEAQGFAPFSKDNGIRIGS---RGHVFVIREPFVSCSPLECRTFFLTQGSLLNDKHSNGTVKDRSPYRTLMSVKVGQSPNVYQARFESVAWSA
N9      NNLTKGLCTINSWHIYGKDNAVRIGE---DSDVLVTREPYVSCDPDECRFYALSQGTTIRGKHSNGTIHDRSQYRALISWPLSSPPTVYNSRVECIGWSS
B       PRLS---CQGSTFQKALLISPHRFGEARGNSAPLIIREPFIACGPKECKHFALTHYAAQPGGYYNGTREDRNKLRHLISVKLGKIPTVENSIFHMAAWSG
              *            *           ***   *   *   *          * *  **   * *                  **

                  190       200       210       220       230       240       250       260       270       280
N2      SSCHDGKAWLHVCITGDDKNATASFIYDGRLVDSIGSWSQNILRTQESECVCINGTCTVVMTDGSASGRADTRILFIEEGKIVHISPLAGSAQHVEECSC
N1      SACHDGMGWLTIGISGPDNGAVAVLKYNGIITETIKSWRKKILRTQESECACVNGSCFTIMTDGPSDGLASYKIFKIEKGKVTKSIELNAPNSHYEECSC
N5      TACHDGKKWMAIGVSGADDDAYAVIHYGGVPTDVIRSWRKQILRTQESSCVCIKGECYWVMTDGPANNQASYKIFKSQKGMVVDEKEISFQGGHIEECSC
N7      TSCHDGVNRMTICVQGDNENATATVYYNKRLTTTIKTWAKNILRTQESECVCHNSTCVVVMTDGPANNQAFTKVIYFHKGMIIKEESLKGSAKHIEECSC
N8      TACHDGKKWMTVGVTGPDNQAVAVVNYGGVPVDIINSWGRDILRTQESSCTCIKGDCYWVMTDGPANRQAKYRIFKAKDGRIIGQTDISFNGGHIEECSC
N9      TSCHDGKTRMSICISGPNNNASAVIWYNRRPVTEINTWARNILRTQESECVCHNGVCPVVFTDGSATGPAETRIYYFKEGKILKWEPLAGTAKHIEECSC
B       SACHDGREWTYIGVDGPDSNALIKIKYGEAYTDTYHSYANNILRTQESACNCIGGDCYLMITDGSASGISKCRFLKIREGRIIKEIFPTGRVEHTEECTC
          ****         *   *    *             ******* * *   *   ***             *            * *** *

                  290       300       310       320       330       340       350       360       370
N2      -YPRYPGVRCICRDNWKGSNRPVVDINMEDYSIDSSYVCSGLVGDTPRNDDRSSNSNCRNP-NNERGTQ-GVKGWAFDNGN-----DLWMGRTISKDLRS
N1      -YPDTGKVMCVCRDNWHGSNRPWVSFD-QNLDYQIGYICSGVFGDNPRPKDG--TGSCGPVYVDGAN---GVKGFSYRYGN-----GVWIGRTKSHSSRH
N5      -YPNMGKVECVCRDNWNGMNRPILIFD-EKLEYEVGYLCAGIPTDTPRVQDSSFTGSCTNAVGRSGTNNYGVKGFGFRQGN-----SVWAGRTISVSSRS
N7      -YGHNQRVTCVCRDNWQGANRPIIEIDMNKLEHTSRYICTGVLTDTSRPKDKTI-GECFNPITGSPGAP-GIKGFGFLNED-----NTWLGRTISPRLRS
N8      -YPNEGKVECVCRDNWTGTNRPILVIS-PDLSYTVGYLCAGIPTDTPRGEDSQFTGSCTSPLGNKG---YGVKGFGFRQGN-----DVWAGRTISRTSRS
N9      -YGERAEITCTCRDNWQGSNRPVIRIDPVAMTHTSQYICSPVLTDNPRPNDPTV-GKCNDPYPGN-NNN-GVKGFSYLDGV-----NTWLGRTISIASRS
B       GFASNKTIECACRDNNYTAKRPFVKLNVETDTAEIRLMCTETYLDTPRPDDGSITGPCESNGDKGRG---GIKG-GFVHQRMASKIGRWYSRTMSKTERM
              * ****     **               *    * * *     *          * **             *  ** *   *

                  380       390       400       410       420       430       440       450       460
N2      GYETFKVIGG--WSTPNSKSQINRQVIVDSDNRSGYSGIFS---V-EGKSCINTCFYVELIRGRKQETR-VWWTSNSIVVFCGTSGTYGTGSWPDGANIN
N1      GFEMIWDPNG--WTETD-SKFSVRQDVVAMTDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKEKT-I-WTSASSISFCGVNSDTVDWSWPDGAELP
N5      GFEVLLIEDG--WIRPS-KTISKKVEVLNNKNWSGYSGAFTIPTAMTSKNCIVPCFWLEMIRGKPEERTSI-WTSSSSTVFCGVSSEVPGWSWDDGAILP
N7      GFEMLKIPNA--GTDPESKIK-ERQEIVSNDNWSGYSGSFIDYWN-DNSECYNPCFYVELIRGRPEEAKYVEWTSNSLIALCGSPISVGSGSFPDGAQIK
N8      GFEIIKIRNG--WTQNS-KDQIRKQVIIDNLNWSGYSGSFTLPVELTKKGCLVPCFWVEMIRGKPED-TTI-WTSSSSIVMCGVDHKIASWSWHDGAILP
N9      GYEMLKVPNA--LTDDKSKPT-QGQTIVLNTDWSGYSGSFMDYWA-E-GECYRACFYVELIRGRPKEDK-VWWTSNSIVSMCSSTEFLGQWDWPDGAKIE
B       GMELYVKYDGDPWTDSDALDPSGVMVSIKEPGW--YSFGF----EIKDKKCDVPCIGIEMVHDGGKKT----WHSAATAIYCLMGS--GQLLWDTVTGVD
        * *                             ** *         *   *   *           * *      *

                  470
N2      FM--PI
N1      FTIDK-
N5      FDIDKM
N7      YF---S
N8      FDIDKI
N9      YF---L
B       M---AL
```

FIG. 3. Alignment of influenza virus neuraminidase protein sequences from the viral subtypes referred to in the Materials and Methods. Sequence numbering corresponds to that of the N2 sequence (11). Asterisks indicate amino acid residues which are identical in all sequences. N2, A/Tokyo/3/67; N1, A/Puerto Rico/8/34; N5, A/Shearwater/Australia/72; N7, A/Cor/16/74; N8, A/Ken/1/81; N9, A/Tern/Australia/G70C/75; B, B/Victoria/3/85.

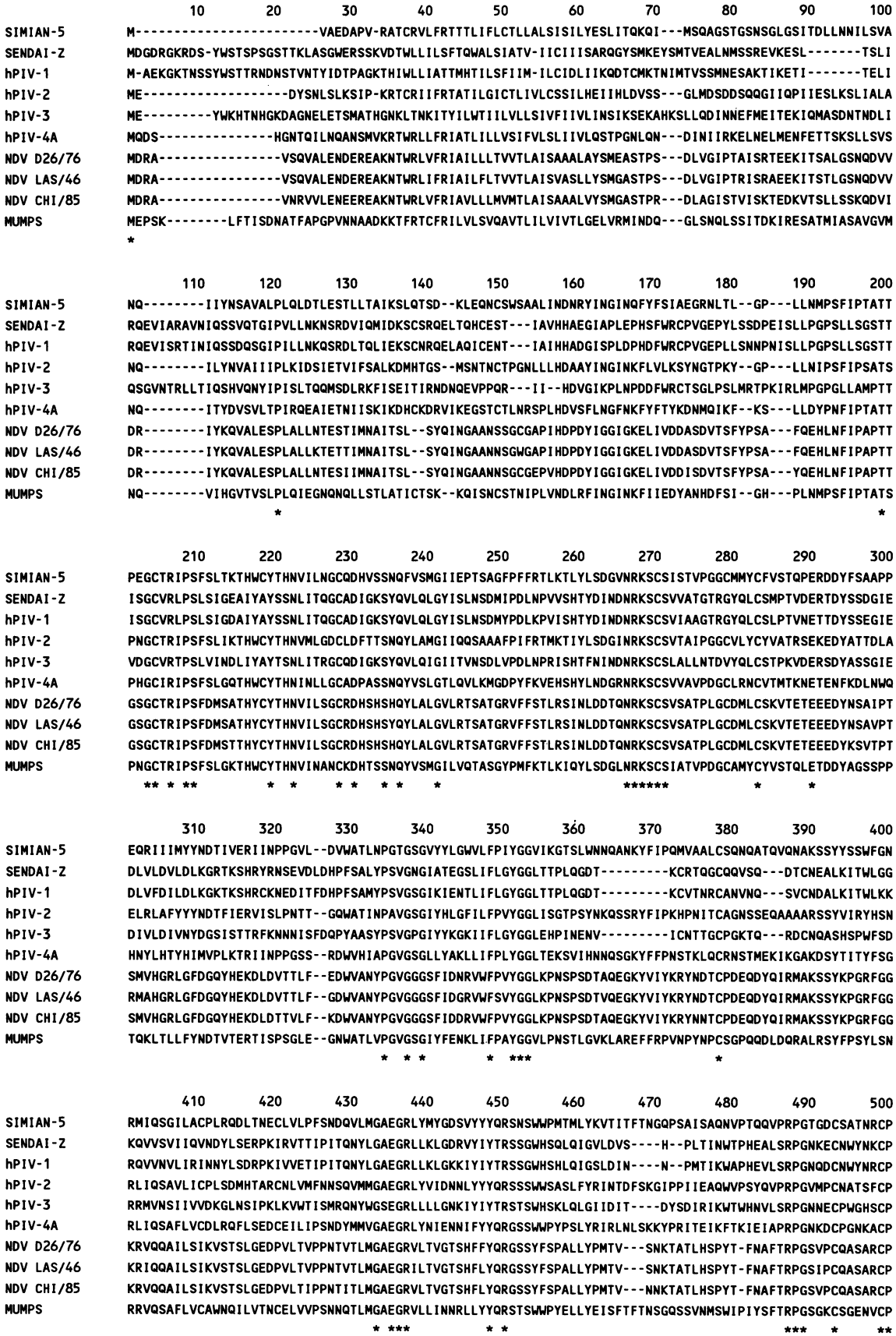
10 20 30 40 50 60 70 80 90 100
SIMIAN-5 M----------VAEDAPV-RATCRVLFRTTTLIFLCTLLALSISILYESLITQKQI---MSQAGSTGSNSGLGSITDLLNNILSVA
SENDAI-Z MDGDRGKRDS-YWSTSPSGSTTKLASGWERSSKVDTWLLILSFTQWALSIATV-IICIIISARQGYSMKEYSMTVEALNMSSREVKESL-------TSLI
hPIV-1 M-AEKGKTNSSYWSTTRNDNSTVNTYIDTPAGKTHIWLLIATTMHTILSFIIM-ILCIDLIIKQDTCMKTNIMTVSSMNESAKTIKETI-------TELI
hPIV-2 ME----------DYSNLSLKSIP-KRTCRIIFRTATILGICTLIVLCSSILHEIIHLDVSS---GLMDSDDSQQGIIQPIIESLKSLIALA
hPIV-3 ME---------YWKHTNHGKDAGNELETSMATHGNKLTNKITYILWTIILVLLSIVFIIVLINSIKSEKAHKSLLQDINNEFMEITEKIQMASDNTNDLI
hPIV-4A MQDS----------HGNTQILNQANSMVKRTWRLLFRIATLILLVSIFVLSLIIVLQSTPGNLQN---DINIIRKELNELMENFETTSKSLLSVS
NDV D26/76 MDRA----------VSQVALENDEREAKNTWRLVFRIAILLLTVVTLAISAAALAYSMEASTPS---DLVGIPTAISRTEEKITSALGSNQDVV
NDV LAS/46 MDRA----------VSQVALENDEREAKNTWRLIFRIAILFLTVVTLAISVASLLYSMGASTPS---DLVGIPTRISRAEEKITSTLGSNQDVV
NDV CHI/85 MDRA----------VNRVVLENEEREAKNTWRLVFRIAVLLLMVMTLAISAAALVYSMGASTPR---DLAGISTVISKTEDKVTSLLSSKQDVI
MUMPS MEPSK--------LFTISDNATFAPGPVNNAADKKTFRTCFRILVLSVQAVTLILVIVTLGELVRMINDQ---GLSNQLSSITDKIRESATMIASAVGVM
*
110 120 130 140 150 160 170 180 190 200
SIMIAN-5 NQ--------IIYNSAVALPLQLDTLESTLLTAIKSLQTSD--KLEQNCSWSAALINDNRYINGINQFYFSIAEGRNLTL--GP---LLNMPSFIPTATT
SENDAI-Z RQEVIARAVNIQSSVQTGIPVLLNKNSRDVIQMIDKSCSRQELTQHCEST---IAVHHAEGIAPLEPHSFWRCPVGEPYLSSDPEISLLPGPSLLSGSTT
hPIV-1 RQEVISRTINIQSSDQSGIPILLNKQSRDLTQLIEKSCNRQELAQICENT---IAIHHADGISPLDPHDFWRCPVGEPLLSNNPNISLLPGPSLLSGSTT
hPIV-2 NQ--------ILYNVAIIIPLKIDSIETVIFSALKDMHTGS--MSNTNCTPGNLLLHDAAYINGINKFLVLKSYNGTPKY--GP---LLNIPSFIPSATS
hPIV-3 QSGVNTRLLTIQSHVQNYIPISLTQQMSDLRKFISEITIRNDNQEVPPQR---II--HDVGIKPLNPDDFWRCTSGLPSLMRTPKIRLMPGPGLLAMPTT
hPIV-4A NQ--------ITYDVSVLTPIRQEAIETNIISKIKDHCKDRVIKEGSTCTLNRSPLHDVSFLNGFNKFYFTYKDNMQIKF--KS---LLDYPNFIPTATT
NDV D26/76 DR--------IYKQVALESPLALLNTESTIMNAITSL--SYQINGAANSSGCGAPIHDPDYIGGIGKELIVDDASDVTSFYPSA---FQEHLNFIPAPTT
NDV LAS/46 DR--------IYKQVALESPLALLKTETTIMNAITSL--SYQINGAANNSGWGAPIHDPDYIGGIGKELIVDDASDVTSFYPSA---FQEHLNFIPAPTT
NDV CHI/85 DR--------IYKQVALESPLALLNTESIIMNAITSL--SYQINGAANNSGCGEPVHDPDYIGGIGKELIVDDISDVTSFYPSA---YQEHLNFIPAPTT
MUMPS NQ--------VIHGVTVSLPLQIEGNQNQLLSTLATICTSK--KQISNCSTNIPLVNDLRFINGINKFIIEDYANHDFSI--GH---PLNMPSFIPTATS
*
210 220 230 240 250 260 270 280 290 300
SIMIAN-5 PEGCTRIPSFSLTKTHWCYTHNVILNGCQDHVSSNQFVSMGIIEPTSAGFPFFRTLKTLYLSDGVNRKSCSISTVPGGCMMYCFVSTQPERDDYFSAAPP
SENDAI-Z ISGCVRLPSLSIGEAIYAYSSNLITQGCADIGKSYQVLQLGYISLNSDMIPDLNPVVSHTYDINDNRKSCSVVATGTRGYQLCSMPTVDERTDYSSDGIE
hPIV-1 ISGCVRLPSLSIGDAIYAYSSNLITQGCADIGKSYQVLQLGYISLNSDMYPDLKPVISHTYDINDNRKSCSVIAAGTRGYQLCSLPTVNETTDYSSEGIE
hPIV-2 PNGCTRIPSFSLIKTHWCYTHNVMLGDCLDFTTSNQYLAMGIIQQSAAAFPIFRTMKTIYLSDGINRKSCSVTAIPGGCVLYCYVATRSEKEDYATTDLA
hPIV-3 VDGCVRTPSLVINDLIYAYTSNLITRGCQDIGKSYQVLQIGIITVNSDLVPDLNPRISHTFNINDNRKSCSLALLNTDVYQLCSTPKVDERSDYASSGIE
hPIV-4A PHGCIRIPSFSLGQTHWCYTHNINLLGCADPASSNQYVSLGTLQVLKMGDPYFKVEHSHYLNDGRNRKSCSVVAVPDGCLRNCVTMTKNETENFKDLNWQ
NDV D26/76 GSGCTRIPSFDMSATHYCYTHNVILSGCRDHSHSHQYLALGVLRTSATGRVFFSTLRSINLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYNSAIPT
NDV LAS/46 GSGCTRIPSFDMSATHYCYTHNVILSGCRDHSHSYQYLALGVLRTSATGRVFFSTLRSINLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYNSAVPT
NDV CHI/85 GSGCTRIPSFDMSTTHYCYTHNVILSGCRDHSHSHQYLALGVLRTSATGRVFFSTLRSINLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYKSVTPT
MUMPS PNGCTRIPSFSLGKTHWCYTHNVINANCKDHTSSNQYVSMGILVQTASGYPMFKTLKIQYLSDGLNRKSCSIATVPDGCAMYCYVSTQLETDDYAGSSPP
** * ** * * * * * * * ****** * *
310 320 330 340 350 360 370 380 390 400
SIMIAN-5 EQRIIIMYYNDTIVERIINPPGVL--DVWATLNPGTGSGVYYLGWVLFPIYGGVIKGTSLWNNQANKYFIPQMVAALCSQNQATQVQNAKSSYYSSWFGN
SENDAI-Z DLVLDVLDLKGRTKSHRYRNSEVDLDHPFSALYPSVGNGIATEGGSLIFLGYGGLTTPLQGDT---------KCRTQGCQQVSQ---DTCNEALKITWLGG
hPIV-1 DLVFDILDLKGKTKSHRCKNEDITFDHPFSAMYPSVGSGIKIENTLIFLGYGGLTTPLQGDT---------KCVTNRCANVNQ---SVCNDALKITWLKK
hPIV-2 ELRLAFYYYNDTFIERVISLPNTT--GQWATINPAVGSGIYHLGFILFPVYGGLISGTPSYNKQSSRYFIPKHPNITCAGNSSEQAAAARSSYVIRYHSN
hPIV-3 DIVLDIVNYDGSISTTRFKNNNISFDQPYAASYPSVGPGIYYKGKIIFLGYGGLEHPINENV---------ICNTTGCPGKTQ---RDCNQASHSPWFSD
hPIV-4A HNYLHTYHIMVPLKTRIINPPGSS--RDWVHIAPGVGSGLLYAKLLIFPLYGGLTEKSVIHNNQSGKYFFPNSTKLQCRNSTMEKIKGAKDSYTITYFSG
NDV D26/76 SMVHGRLGFDGQYHEKDLDVTTLF--EDWVANYPGVGGGSFIDNRVWFPVYGGLKPNSPSDTAQEGKYVIYKRYNDTCPDEQDYQIRMAKSSYKPGRFGG
NDV LAS/46 RMAHGRLGFDGQYHEKDLDVTTLF--GDWVANYPGVGGGSFIDGRVWFSVYGGLKPNSPSDTVQEGKYVIYKRYNDTCPDEQDYQIRMAKSSYKPGRFGG
NDV CHI/85 SMVHGRLGFDGQYHEKDLDTTVLF--KDWVANYPGVGGGSFIDDRVWFPVYGGLKPNSPSDTAQEGKYVIYKRYNNTCPDEQDYQIRMAKSSYKPGRFGG
MUMPS TQKLTLLFYNDTVTERTISPSGLE--GNWATLVPGVGSGIYFENKLIFPAYGGVLPNSTLGVKLAREFFRPVNPYNPCSGPQQDLDQRALRSYFPSYLSN
* * * * * * * *
410 420 430 440 450 460 470 480 490 500
SIMIAN-5 RMIQSGILACPLRQDLTNECLVLPFSNDQVLMGAEGRLYMYGDSVYYYQRSNSWWPMTMLYKVTITFTNGQPSAISAQNVPTQQVPRPGTGDCSATNRCP
SENDAI-Z KQVVSVIIQVNDYLSERPKIRVTTIPITQNYLGAEGRLLKLGDRVYIYTRSSGWHSQLQIGVLDVS----H--PLTINWTPHEALSRPGNKECNWYNKCP
hPIV-1 RQVVNVLIRINNYLSDRPKIVVETIPITQNYLGAEGRLLKLGKKIYIYTRSSGWHSHLQIGSLDIN----N--PMTIKWAPHEVLSRPGNQDCNWYNRCP
hPIV-2 RLIQSAVLICPLSDMHTARCNLVMFNNSQVMMGAEGRLYVIDNNLYYYQRSSSWWSASLFYRINTDFSKGIPPIIEAQWVPSYQVPRPGVMPCNATSFCP
hPIV-3 RRMVNSIIVVDKGLNSIPKLKVWTISMRQNYWGSEGRLLLLGNKIYIYTRSTSWHSKLQLGIIDIT----DYSDIRIKWTWHNVLSRPGNNECPWGHSCP
hPIV-4A RLIQSAFLVCDLRQFLSEDCEILIPSNDYMMVGAEGRLYNIENNIFYYQRGSSWWPYPSLYRIRLNLSKKYPRITEIKFTKIEIAPRPGNKDCPGNKACP
NDV D26/76 KRVQQAILSIKVSTSLGEDPVLTVPPNTVTLMGAEGRVLTVGTSHFFYQRGSSYFSPALLYPMTV---SNKTATLHSPYT-FNAFTRPGSVPCQASARCP
NDV LAS/46 KRIQQAILSIKVSTSLGEDPVLTVPPNTVTLMGAEGRILTVGTSHFLYQRGSSYFSPALLYPMTV---SNKTATLHSPYT-FNAFTRPGSIPCQASARCP
NDV CHI/85 KRVQQAILSIKVSTSLGEDPVLTIPPNTITLMGAEGRVLTVGTSHFLYQRGSSYFSPALLYPMTV---NNKTATLHSPYT-FNAFTRPGSVPCQASARCP
MUMPS RRVQSAFLVCAWNQILVTNCELVVPSNNQTLMGAEGRVLLINNRLLYYQRSTSWWPYELLYEISFTFTNSGQSSVNMSWIPIYSFTRPGSGKCSGENVCP
* *** * * *** * *

FIG. 4

```
                  510       520       530       540       550       560       570       580       590       600
SIMIAN-5    GFCLTGVYADAWLLT-NPSSTS--TFGSEATFTGSYLNTATQRINPTMYIANNTQIISSQQFGSSGQEAAYGHTTCFRDTGSVMVYCIYIIELSSSLLGQ
SENDAI-Z    KECISGVYTDAYPLSPDAANV-----------ATVTLYANTSRVNPTIMYSNTTNIINMLRIKDVQLEAAYTTTSCITHFG--KGYCFHIIEINQKSLNT
hPIV-1      RECISGVYTDAYPLSPDAVNV-----------ATTTLYANTSRVNPTIMYSNTSEIINMLRLKNVQLEAAYTTTSCITHFG--KGYCFHIVEINQTSLNT
hPIV-2      ANCITGVYADVWPLN-DPEPTSQNALNPNYRFAGAFLRNESNRTNPTFYTASASALLNTTGFNNTNHKAAYTSSTCFKNTGTQKIYCLIIIEMGSSLLGE
hPIV-3      DGCITGVYTDAYPLNPTGSIV-----------SSVILDSQKSRVNPVITYSTATERVNELAIRNRTLSAGYTTTSCITHYN--KGYCFHIVEINHKSSNT
hPIV-4A     KECITGVYQDILPLS-YPNTAFPHL--KQAYYTGFYLNNSLERRNPTFYTADNLDYHQQERLGKFNLTAGYSTTTCFKQTTTARLYCLYIIEVGDSVIGD
NDV D26/76  NSCVTGVYTDPYPLVFYRNHTLRGVF-------GTMLDDEQARLNPVSAVFDSISRSRITRVSSSSTKAAYTTSTCFKVVKTNKTYCLSIAEISNTLFGE
NDV LAS/46  NPCVTGVYTDPYPLIFYRNHTLRGVF-------GTMLDGVQARLNPTSAVFDSTSRSRITRVSSSSTKAAYTTSTCFKVVKTNKTYCLSIAEISNTLFGE
NDV CHI/85  NSCITGVYTDPYPLIFHRNHTLRGVF-------GTMLDDEQARLNPVSAVFDNISRSRVTRVSSSSTKAAYTTSTCFKVVKTNKAYCLSIAEISNTLFGE
MUMPS       IACVSGVYLDPWPLT---PYSHQSGINRNFYFTGALLNSSTTRVNPTLYVSALNNLKVLAPYGTQGLSASYTTTTCFQDTGDASVYCVYIMELASNIVGE
                 * *** *  *                    *    * **                    * *   *         **  * *

                  610       620       630       640       650       660
SIMIAN-5    FQIVPFIRQVTLS
SENDAI-Z    LQPMLFKTSIPKLCKAES
hPIV-1      LQPMLFKTSIPKICKITS
hPIV-2      FQIIPFLRELIP
hPIV-3      FQPMLFKTEIPKSCS
hPIV-4A     FQITLFLAA
NDV D26/76  FRIVPLLVEILKDDGVREARSGRLSQLQEGWKDDIVSPIFCDAKNQTEYRRELESYAASWP
NDV LAS/46  FRIVPLLVEILKDDGVREARSG
NDV CHI/85  FRIVPLLVEILKDDRV
MUMPS       FQILPVLTRLTIT
```

FIG. 4. Alignment of 10 paramyxovirus HN protein sequences taken from GenBank. SIMIAN-5, Simian virus 5; SENDAI-Z, Sendai virus Z strain; hPIV-1; hPIV-2; hPIV-3 strain WASH/1511/73; hPIV-4A; NDV strain D26/76; NDV strain LAS/46; NDV strain CHI/85; MUMPS, mumps virus. Sequences were aligned from residue positions 1 to 608 inclusive. Asterisks indicate amino acid residues which are identical in all sequences.

therefore fulfil the active-site structural role (50) of binding to active-site residues R-292 and Y-406 (HN counterparts are R-450 and Y-571). Alternatively, it could be involved in stabilizing a calcium ion site located near that observed in neuraminidase. Note that there is no evidence for such a site in HN and that in neuraminidase the calcium ion ligands are D-324 and the backbone carbonyl groups of residues 293, 297, 345, and 347.

Beyond the fourth alignment in Fig. 5 involving HN R-543 as the homologous residue to R-371 in neuraminidase, conserved HN residues Y-571 and E-592 appear as the counterparts of Y-406 and E-425 in neuraminidase. The pair of cysteine residues between them may exist in a disulfide bond because a covalent link between their counterparts in neuraminidase at positions 411 and 420 is plausible.

Finally, a disulfide bond may join residues in the homologous positions to neuraminidase residues 88 and 447 (HN numbers 173 and 614). This connection would not be possible in neuraminidase without a deletion of residues from the loop β_6L_{12}, and indeed the requisite cysteines for this disulfide bond are only found in HN sequences which lack two amino acids precisely in this place, i.e., between C-576 and C-587 in HN (Fig. 4).

It is also possible to align the sequences of three bacterial

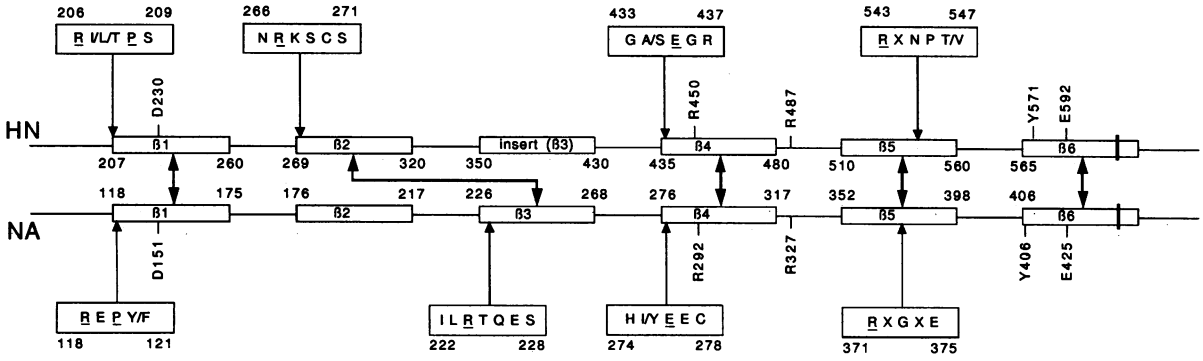

FIG. 5. Schematic representation of the proposed structural similarity of paramyxovirus HN and influenza virus neuraminidase (NA). The four highly conserved sequence motifs in HN (see text) that can be identified with active-site regions in NA are shown in the uppermost boxes. The locations of the corresponding NA regions are shown in the lower half of the diagram, with the key active-site residues underlined. The β-sheet secondary structure of NA is indicated in the boxes labelled β1 through β6; the large vertical arrows link these to the corresponding proposed secondary structural elements of HN. Also shown are the HN residues D-230, R-450, R-487, Y-571, and E-592 and the corresponding NA residues D-151, R-292, R-327, Y-406, and E-425. The vertical bar in sheet β6 indicates the transition from the C to the N terminus of the corresponding polypeptide.

TABLE 1. Putative active-site residues of bacterial neuraminidases

Neuraminidase from	Active-site residue							
Influenza virus	R-118	D-151	R-224	E-276	R-292	R-371	Y-406	E-425[a]
C. sordelli	R-55	D-80	R-115	E-253	R-270	R-330	Y-365	E-380
C. perfringens	R-37	D-62	R-97	E-235	R-252	R-312	Y-347	E-362
S. typhimurium	R-37	D-62	R-97	E-236	R-251	R-309	Y-342	E-364

[a] Manual correction of the automatic alignment described in the text for the three bacterial neuraminidase sequences was necessary to show conservation of sequence at the site homologous to E-425.

neuraminidases (*Clostridium perfringens* [41], *Clostridium sordelli* [42], and *Salmonella typhimurium* [24]) to capture the conserved elements of the influenza virus enzyme structure described above for HN. These alignments, not shown here, preserve the aspartic acid box motif first described by Roggentin et al. (40). Grafting these sequences onto a model of influenza virus neuraminidase is a more difficult task than the exercise described above for HN sequences, partly because there are fewer sequences in the alignment and conserved residues cannot be read as such with the same level of confidence. Tentative assignment of active-site functional residues as shown in Table 1 requires that, as for HN, the counterpart of influenza virus neuraminidase R-224 is on the second sheet of the propeller and not the third. Also, as for HN, an insertion between the counterparts of R-224 and E-276 shows no homology with influenza virus neuraminidase. A 10-kDa peptide of *S. typhimurium* neuraminidase has been labelled with an active-site photoprobe (53). The peptide includes R-309 and Y-342 which are elements of the proposed active site (Table 1). Functional residues in the active site (Fig. 2) appear conserved between viral and bacterial enzymes, but structural residues of the active site are not. The three-dimensional structure of a bacterial neuraminidase should be known shortly (45).

DISCUSSION

A summary of the essential elements of the HN model is shown in Fig. 5. The model provides for most of the functional elements of the active site of influenza virus neuraminidase within the sequence of the HN protein. Analogs of amino acids contacting the carboxylate of the bound sugar (R-118, R-292, and R-371) and the glycerol side chain (R-224 and E-276) are present in HN, as is the putative catalytic aspartic acid (D-151) and the tyrosine residue underlying the sugar-binding site (Y-406). The binding site for the *N*-acetyl moiety is apparently different in HN, since no direct counterparts of R-152, W-178, and I-222 of neuraminidase are seen in HN. On the basis of alignments focused on active-site functional residues, three additional conserved sequence windows have plausible structural interpretations and a number of potential disulfide bonds consistent with structure and conserved sequence patterns are identified.

The assignment (Fig. 4) of R-267 as the HN counterpart of R-224 suggests that it is located on the second of the six sheets in the propeller structure and not the third as found in influenza virus neuraminidase. The insertion in HN which follows, in order to regain the register of sequences around E-276 in neuraminidase, may correspond in structure, if not in sequence, to sheet three of the influenza virus enzyme. The lack of homology between this inserted region in HN and influenza virus neuraminidase suggests caution in regarding the insert as the third sheet of the HN structure, although this is clearly one possibility.

There is an alternative interpretation for the conservation of R-267 in the HN sequences and that is that it may play the role of R-152 in influenza virus neuraminidase, i.e., hydrogen bonding to the carbonyl group of the *N*-acetyl moiety (50). This role would have to be fulfilled from a structural loop directly adjacent to the β_1L_{23} loop holding R-152, viz., β_2L_{01}. The proximity of these loops makes such an interpretation plausible. Our preference for the first interpretation above is that R-224 in neuraminidase and R-267 in HN are both embedded in conserved hepta- and hexapeptide sequences, respectively, and the rare occurrence of such long stretches of conserved sequence in Fig. 3 and 4 suggests a structural correlation. The main findings of this work are unaffected by this ambiguity.

The model suggests that framework residues of the active site referred to in Fig. 2 can be replaced by amino acids other than those seen in influenza virus neuraminidase and that even some of the functional residues, especially those in the neighborhood of the *N*-acetyl group, are dispensable. Site-directed mutagenesis of the neuraminidase active site suggests that the enzyme activity is very sensitive to single amino acid sequence changes in both the functional and framework residues. In the latter category, substitutions of D-198 to Asn and E-277 to Asp both result in loss of enzyme activity (31).

Several studies have sought to directly identify amino acids in the neuraminidase active site of the HN protein. A number of laboratory variants of HN which have altered neuraminidase activities have been characterized, and gene sequences have been described for several of these.

Waxham and Aronowski (54) have sequenced a variant of mumps virus selected by growth in the presence of a neuraminidase inhibitor (56). This variant has no detectable neuraminidase activity, and the virus causes extensive syncytium formation in cell culture. The translated amino acid sequence of the HN protein of the variant showed two changes with respect to the parental strain, I-181 (position 207 in Fig. 4) replaced by threonine and Q-261 (position 288 in Fig. 4) to lysine. The first of these changes involves one of the framework residues of the active site, and the second is in the vicinity of β_2L_{23} (Fig. 5). D-198, part of this loop in neuraminidase, is a structural residue of the active site (50). The β_2L_{23} loop is further implicated in catalysis in HN by the observation (4) that the sequence of hPIV-4A (Fig. 4) differs from that of all other HN sequences by the substitution of D-293-Y-294 with the sequence N-F. It has been proposed (4) that this change correlates with reduced enzyme activity in the HN protein of hPIV-4A.

In another study (25), monoclonal antibodies were used to select variants of NDV. Some of these, in particular variants selected with a monoclonal antibody whose binding was competitive with that of a low-molecular-weight neuraminidase inhibitor, showed altered neuraminidase activity, and sequencing showed changes in these variants at residues 193 and 200 (positions 226 and 233 in Fig. 4). The homologous residue numbers in influenza virus neuraminidase are 147 and 154, i.e., either side of the functional residue D-151.

Iorio et al. (25) also studied a revertant of a temperature-sensitive mutant which has impaired neuraminidase activity. That variant has I-175 replaced by methionine (residue 207 in Fig. 4). Residue 119, the influenza virus neuraminidase counterpart, is one of the structural residues of the active site of neuraminidase. A second-step revertant which partially restores neuraminidase activity has F-192 (residue 225

in Fig. 4 and leucine in the NDV sequences shown there) replaced by leucine. A possible interaction between these two amino acids is plausible within the influenza virus neuraminidase by virtue of the proximity of the homologous residues 119 and 146.

The above three studies demonstrate that changes in the framework amino acids of the active site of the model proposed here influence enzyme activity.

The proposed association of the four highly conserved regions of HN with corresponding regions in the active site of influenza virus neuraminidase is supported indirectly by the fact that only one of these highly conserved regions (viz., the R X P motif, HN residues 206 to 208) appears to occur in the haemagglutinin (H) proteins of morbilliviruses. The morbilliviruses (39, 13), while also members of the *Paramyxoviridae* family, have a similar morphological and genetic organization to the paramyxoviruses, yet their attachment protein H specifically lacks neuraminidase activity.

The models for both HN and bacterial neuraminidases suggest that in neither case are there direct counterparts of the influenza virus neuraminidase active-site framework residues E-119 and E-227. These particular amino acids are believed to be important for tight binding of 4-amino and 4-guanidino substituted analogs of sialic acid (52). The replacement of E-119 by isoleucine or leucine suggests a more hydrophobic character to the binding site near the C-4 position of sialic acid in HN or bacterial neuraminidase. Some 2,3-unsaturated analogs of sialic acid (Neu5Ac2eu) with hydrophobic properties at C-4 have been tested for inhibiting activity against HN and bacterial neuraminidase but none are yet more potent than the parent molecule Neu5Ac2en (23).

In summary, earlier work of Jorgensen et al. (26), claiming a similarity in structure between influenza virus neuraminidase and the HN protein of paramyxovirus, has been extended to demonstrate how such a model can incorporate the essential amino acids of the neuraminidase active site within the HN structure. Many of the framework elements of the neuraminidase active site, i.e., amino acids which do not themselves contact substrate directly, are not conserved in this model of the HN structure. A similar model can be built for some bacterial neuraminidases.

REFERENCES

1. **Air, G. M., W. G. Laver, M. Luo, S. J. Stray, G. Legrone, and R. G. Webster.** 1990. Antigenic, sequence, and crystal variation in influenza B neuraminidase. Virology **177**:578–587.
2. **Air, G. M., L. R. Ritchie, W. G. Laver, and P. M. Colman.** 1985. Gene and protein sequence of an influenza neuraminidase with hemagglutinin activity. Virology **145**:117–122.
3. **Baker, A. T., J. N. Varghese, W. G. Laver, G. M. Air, and P. M. Colman.** 1987. The three-dimensional structure of neuraminidase of subtype N9 from an avian influenza virus. Proteins Struct. Funct. Genet. **2**:111–117.
4. **Bando, H., K. Kondo, M. Kawano, H. Komada, M. Tsurudome, M. Nishio, and Y. Into.** 1990. Molecular cloning and sequence analysis of human parainfluenza type 4A virus HN gene: its irregularities on structure and activities. Virology **175**:307–312.
5. **Basak, S., M. Tomana, and R. W. Compans.** 1985. Sialic acid is incorporated into influenza haemagglutinin glycoproteins in the absence of viral neuraminidase. Virus Res. **2**:61–68.
6. **Burmeister, W. P., R. W. H. Ruigrok, and S. Cusack.** 1992. The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid. EMBO J. **11**:49–56.
7. **Burnet, F. M.** 1948. Mucins and mucoids in relation to influenza virus action. IV. Inhibition by purified mucoid of infection and haemagglutinin with the virus strain WSE. Aust. J. Exp. Biol. Med. Sci. **26**:381–387.
8. **Burnet, F. M., J. F. McCrea, and S. G. Anderson.** 1947. Mucin as a substrate of enzyme action by viruses of the mumps influenza group. Nature (London) **160**:404–405.
9. **Burnet, F. M., and J. D. Stone.** 1947. The receptor destroying enzyme of *V. cholerae*. Aust. J. Exp. Biol. Med. Sci. **25**:227–233.
10. **Chong, K. J., M. S. Pegg, N. R. Taylor, and M. von Itzstein.** 1992. Evidence for a sialosyl cation transition-state complex in the reaction of sialidase from influenza virus. Eur. J. Biochem. **207**:335–343.
11. **Colman, P. M.** 1989. Influenza virus neuraminidase: enzyme and antigen, p. 175–218. *In* R. M. Krug (ed.), The influenza viruses. Plenum Publishing Co., New York.
12. **Colman, P. M., J. N. Varghese, and W. G. Laver.** 1983. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) **303**:41–44.
13. **Curran, M. D., D. K. Clarke, and B. K. Rima.** 1991. The nucleotide sequence of the gene encoding the attachment protein H of canine distemper virus. J. Gen. Virol. **72**:443–447.
14. **Dale, B., R. Brown, J. Miller, R. T. White, G. M. Air, and B. Cordell.** 1986. Nucleotide and deduced amino acid sequence of the influenza neuraminidase genes of two equine serotypes. Virology **155**:460–468.
15. **Fields, S., G. Winter, and G. G. Brownlee.** 1981. Structure of the neuraminidase gene in human influenza virus A/PR/8/34. Nature (London) **290**:213–217.
16. **Galen, J. E., J. M. Ketley, A. Fasano, S. H. Richardson, S. S. Wasserman, and J. B. Kaper.** 1992. Role of *Vibrio cholerae* neuraminidase in the function of cholera toxin. Infect. Immun. **60**:406–415.
17. **Gorman, W. L., D. S. Gill, R. A. Scroggs, and A. Portner.** 1990. The hemagglutinin-neuraminidase glycoproteins of human parainfluenza virus type 1 and Sendai virus have high structure-function similarity with limited antigenic cross-reactivity. Virology **175**:211–221.
18. **Gottschalk, A.** 1958. Neuraminidase: its substrate and mode of action. Adv. Enzymol. **20**:135–145.
19. **Griffin, J. A., S. Basak, and R. W. Compans.** 1983. Effects of hexose starvation and the role of sialic acid in influenza virus release. Virology **125**:324–334.
20. **Harley, V. R., C. W. Ward, and P. J. Hudson.** 1989. Molecular cloning and analysis of the N5 neuraminidase subtype from an avian influenza virus. Virology **169**:239–243.
21. **Hiebert, S. W., R. G. Paterson, and R. A. Lamb.** 1985. Hemagglutinin-neuraminidase protein of the paramyxovirus simian virus 5: nucleotide sequence of the mRNA predicts an N-terminal membrane anchor. J. Virol. **54**:1–6.
22. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**:237–244.
23. **Holzer, C. T., M. von Itzstein, B. Jin, M. S. Pegg, W. P. Stewart, and W.-Y. Wu.** Inhibition of sialidases from viral, bacterial and mammalian sources by analogues of 2-deoxy-2,3-didehydro-N-acetylneuraminic acid modified at the C-4 position. Glycoconjugate J., in press.
24. **Hoyer, L. L., A. C. Hamilton, S. M. Steenbergen, and E. R. Vimr.** 1992. Cloning, sequencing and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nan*H, provides evidence for interspecies gene transfer. Mol. Microbiol. **6**:873–884.
25. **Iorio, R. M., R. J. Syddall, R. L. Glickman, A. M. Riel, J. P. Sheehan, and M. A. Bratt.** 1989. Identification of amino acid residues important to the neuraminidase activity of the HN glycoprotein of Newcastle disease virus. Virology **173**:196–204.
26. **Jorgensen, E. D., P. L. Collins, and P. Lomedico.** 1987. Cloning and nucleotide sequence of Newcastle disease virus haemagglutinin-neuraminidase mRNA: identification of a putative sialic acid binding site. Virology **156**:12–24.
27. **Kahn, S., T. G. Colbert, J. C. Wallace, N. A. Hoagland, and H. Eisen.** 1991. The major 85-kDa surface antigen of the mammalian-stage forms of *Trypanosoma cruzi* is a family of sialidases. Proc. Natl. Acad. Sci. USA **88**:4481–4485.
28. **Kawano, M., K. Okamoto, H. Band, K. Kondo, M. Tsurudome, M. Nishio, and Y. Into.** 1991. Characterization of the human

parainfluenza type 2 virus gene encoding the L protein and the intergenic sequences. Nucleic Acids Res. **19:**2739–2746.
29. **Kraulis, P. J.** 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. **24:**946–950.
30. **Lentz, M. R., G. M. Air, W. G. Laver, and R. G. Webster.** 1984. Sequence of the neuraminidase gene of influenza virus and previously uncharacterized monoclonal variants. Virology **135:** 257–265.
31. **Lentz, M. R., G. M. Air, and R. G. Webster.** 1987. Site-directed mutation of the active site of influenza neuraminidase and implications for the catalytic mechanism. Biochemistry **26:** 5351–5358.
32. **Mejia, J. S., E. Ortega-Barria, D. Matzilevich, and R. P. Prioli.** 1991. The *Trypanosoma cruzi* neuraminidase contains sequences similar to bacterial neuraminidases, YWTD repeats of the low density lipoprotein receptor, and type III modules of fibronectin. J. Exp. Med. **174:**179–191.
33. **Miura, N., Y. Nakatani, M. Ishiura, T. Uchida, and Y. Okada.** 1985. Molecular cloning of a full-length cDNA encoding the hemagglutinin-neuraminidase glycoprotein of Sendai virus. FEBS Lett. **188:**112–116.
34. **Morrison, T., and A. Portner.** 1991. Structure, function and processing of the glycoproteins of *Paramyxoviridae*, p. 347–382. *In* D. W. Kingsbury (ed.), The paramyxoviruses. Plenum Publishing Co., New York.
35. **Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman.** 1992. The α/β hydrolase fold. Protein Eng. **5:**197–211.
36. **Palese, P., and R. W. Compans.** 1976. Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoro-acetyl-neuraminic acid (FANA): mechanism of action. J. Gen. Virol. **33:**159–163.
37. **Palese, P., K. Tobita, M. Ueda, and R. W. Compans.** 1974. Characterisation of temperature sensitive influenza virus mutants defective in neuraminidase. Virology **61:**397–410.
38. **Pereira, M. E. A., J. S. Mejia, E. Ortega-Barria, D. Matzilevich, and R. P. Prioli.** 1991. The *Trypanosoma cruzi* neuraminidase contains sequences similar to bacterial neuraminidases, YWTD repeats of the low density lipoprotein receptor, and type III modules of fibronectin. J. Exp. Med. **174:**179–191.
39. **Pringle, C. R.** 1991. The genetics of paramyxoviruses, p. 1–39. *In* D. W. Kingsbury (ed.), The paramyxoviruses. Plenum Publishing Co., New York.
40. **Roggentin, P., B. Rothe, J. B. Kapaer, J. Galen, L. Lawrisuk, E. R. Vimr, and R. Schauer.** 1989. Conserved sequences in bacterial and viral sialidases. Glycoconjugate J. **6:**349–353.
41. **Roggentin, P., B. Rothe, F. Lottspeich, and R. Schauer.** 1988. Cloning and sequencing of a *Clostridium perfringens* sialidase gene. FEBS Lett. **238:**31–34.
42. **Rothe, B., P. Roggentin, R. Frank, H. Blocker, and R. Schauer.** 1989. Cloning, sequencing and expression of a sialidase gene from *Clostridium sordelli* G12. J. Gen. Microbiol. **135:**3087–3096.
43. **Sakaguchi, T., T. Toyodo, B. Gotoh, N. M. Inocencio, K. Kuma, T. Miyata, and Y. Nagai.** 1989. Newcastle disease virus evolution: I. Multiple lineages defined sequence variability of the hemagglutinin-neuraminidase gene. Virology **169:**260–272.
44. **Scheid, A., L. A. Caliguiri, R. W. Compans, and P. W. Choppin.** 1972. Isolation of paramyxovirus glycoproteins. Association of both haemagglutinating and neuraminidase activities with the larger SV5 glycoprotein. Virology **50:**640–652.
45. **Taylor, G., E. Vimr, E. Garman, and G. Laver.** 1992. Purification, crystallisation and preliminary crystallographic study of neuraminidase from *Vibrio cholerae* and *Salmonella typhimurium*. J. Mol. Biol. **226:**1287–1290.
46. **Tulip, W. R., J. N. Varghese, A. T. Baker, A. van Donkelaar, W. G. Laver, R. G. Webster, and P. M. Colman.** 1991. Refined atomic structures of N9 subtype influenza virus neuraminidase and escape mutants, J. Mol. Biol. **221:**487–497.
47. **Van Wyke Coelingh, K. L., C. C. Winter, and B. R. Murphy.** 1988. Nucleotide and deduced amino acid sequence of hemagglutinin-neuraminidase genes of human type 3 parainfluenza viruses. Virology **162:**137–143.
48. **Varghese, J. N., and P. M. Colman.** 1991. Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. J. Mol. Biol. **221:**473–486.
49. **Varghese, J. N., W. G. Laver, and P. M. Colman.** 1983. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature (London) **303:**35–40.
50. **Varghese, J. N., J. McKimm-Breschkin, J. B. Caldwell, A. A. Kortt, and P. M. Colman.** 1992. The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. Proteins Struct. Funct. Genet. **14:**327–332.
51. **Varghese, J. N., R. G. Webster, W. G. Laver, and P. M. Colman.** 1988. Structure of an escape mutant of glycoprotein N2 neuraminidase of influenza virus A/Tokyo/3/67 at 3 Å resolution. J. Mol. Biol. **200:**201–203.
52. **von Itzstein, M., W.-Y. Wu., G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, J. N. Varghese, P. M. Colman, D. M. Ryan, J. M. Woods, R. C. Bethell, V. J. Hotham, J. M. Cameron, and C. R. Penn.** Rational design of potent sialidase-based inhibitors of influenza virus replication. Submitted for publication.
53. **Warner, T. G., R. Harris, R. McDowell, and E. R. Vimr.** 1992. Photolabelling of *Salmonella typhimurium* LT2 sialidase. Identification of a peptide with a predicted structural similarity to the active sites of influenza-virus sialidases. Biochem. J. **285:**957–964.
54. **Waxham, M. N., and J. Aronowski.** 1988. Identification of amino acids involved in the sialidase activity of the mumps virus hemagglutinin-neuraminidase protein. Virology **167:**226–232.
55. **Waxham, M. N., J. Aronowski, A. C. Server, J. S. Wolinsky, J. A. Smith, and H. M. Goodman.** 1988. Sequence determination of the mumps virus HN gene. Virology **164:**318–325.
56. **Waxham, M. N., and J. S. Wolinsky.** 1986. A fusing mumps virus variant selected from a nonfusing parent with the neuraminidase inhibitor 2-deoxy-2,3-dehydro-N-acetylneuraminic acid. Virology **151:**2286–2295.