PROTEINS: Structure, Function, and Genetics 14:327–332 (1992)

The Structure of the Complex Between Influenza Virus Neuraminidase and Sialic Acid, the Viral Receptor

Joseph N. Varghese, Jennifer L. McKimm-Breschkin, James B. Caldwell, Alexander A. Kortt, and Peter M. Colman

CSIRO, Division of Biomolecular Engineering, Parkville, Victoria 3052, Australia

ABSTRACT **Crystallographic studies of neuraminidase–sialic acid complexes indicate that sialic acid is distorted on binding the enzyme. Three arginine residues on the enzyme interact with the carboxylate group of the sugar which is observed to be equatorial to the saccharide ring as a consequence of its distorted geometry. The glycosidic oxygen is positioned within hydrogen-bonding distance of Asp-151, implicating this residue in catalysis.**

Key words: crystallographic, neuraminidase, influenza virus, sialic acid

INTRODUCTION

The influenza virus receptor is a glycoconjugate on which sialic acid is the terminal sugar.[1] The viral neuraminidase destroys the receptor by specifically cleaving the glycosidic linkage between sialic acid and the adjoining saccharide unit[2] thereby allowing the spread of infection to new host cells. The three-dimensional structure of the neuraminidase is known for two different subtypes of influenza, N2[3,4] and N9,[5] and, together with a preliminary analysis of binding of sialic acid to the enzyme,[6] these studies have implicated a strain-invariant cavity on the surface of the molecule as the likely enyzme active site. Here that result is confirmed and extended by high resolution studies of the complex of neuraminidase with sialic acid and an unsaturated derivative of sialic acid which is thought to mimic the transmition state structure. The crystallographic data suggest that sialic acid is distorted at C2 when bound to neuraminidase.

The neuraminidase is a tetramer of 50-kDa glycosylated polypeptides, with a β-sheet propellor fold, which display C_4 symmetry.[3] Sialic acid is the product of the reaction catalyzed by the enzyme and its binding site has previously been characterized as a cavity on the upper surface of the subunit, directly above the six, parallel, central strands of the propellor structure.[6] The site is unusual for the concentration of charged amino acids which project into it, in contrast to the sialic acid binding site of the viral hemagglutinin which has one histidine and one glutamic acid residue.[7]

Crystallographic studies of three enzyme–inhibitor complexes have now been undertaken (Table I). The earlier low resolution analysis of the sialic acid complex has been extended to include at least partial data to 1.8 Å resolution (Table I) in parallel with analyses of complexes of 2-deoxy-2,3-dehydro- *N*-acetylneuraminic acid and its trifluoracetyl derivative.[8] These unsaturated compounds are believed to approximate transition-state structures[9,10] by virtue of their planarity at C2 (Fig. 1), and are the most potent inhibitors of neuraminidase activity so far described.

MATERIALS AND METHODS

Details of the enzyme preparation and crystallization have been reported previously[3,11] and the atomic structure of the native enzyme has been determined to 2.2 Å resolution.[4] This model of the native structure includes amino acid residues 83 to 469, four oligosaccharide structures N-linked at asparagine residues 86, 146, 200, and 234, a putative Ca^{2+} ion site, and 85 water molecules. The crystallographic residual was 0.21 for data between 6 and 2.2 Å and the rms deviations from ideal geometry were 0.02 Å for bond lengths and 3.9° for bond angles, with a cis peptide at residue Lys-431.

X-ray diffraction data were collected for three inhibitor–enzyme complexes prepared by soaking crystals of native influenza virus A/Tokyo/3/67 neuraminidase[3,11] in inhibitor solutions. Two inhibitor–neuraminidase complexes NANA1 and NANA2 (see Table I) were prepared by soaking native crystals in a 0.5 mM solution of *N*-acetyl neuraminic acid (Sigma Chemicals) for 21 hr, and the inhibitor–neuraminidase complex DANA was prepared by soaking native crystals in a 5 mM solution of 2-deoxy-2,3-dehydro-*N*-acetyl neuraminic acid (Boehringer) for 2 days. Another dataset (FANA) was collected from crystals soaked in 2-deoxy-2,3-dehydro-*N*-trifluoracetyl neuraminic acid, a gift from Glaxo Group Research. Data were collected on film by screenless oscillation photography[12] with a rotating anode X-ray source for the

Received December 16, 1991; revision accepted March 6, 1992.

Address reprint requests to Dr. Joseph N. Varghese, CSIRO-Australia, Division of Biomolecular Engineering, Parkville Laboratory, 343 Royal Parade, Parkville, Victoria, 3052, Australia.

TABLE I. Crystallographic Studies of Enzyme–Inhibitor Complexes*

Dataset	d_{min}	Source	Observations	Images	Reflections	$R(I)$	Cmp. (%)
NANA1	2.9	RA	32565	12	11900	0.088	56
DANA	2.9	RA	40286	18	13750	0.088	65
NANA2	1.8	PF	175644	16	50917	0.058	46

*NANA and DANA refer to crystals of influenza virus A/Tokyo/3/67 neuraminidase[3,11] soaked, respectively, in *N*-acetyl neuraminic acid and 2-deoxy-2,3-dehydro-*N*-acetyl neuraminic acid. RA and PF refers to data collected on film by oscillation photography[12] with a rotating anode X-ray source (RA), and on image plates with Weissenberg geometry[13,14] at the Photon Factory (PF). $R(I) = \Sigma |I_i - \langle I \rangle| / \Sigma I_i$, the sum being over all observations of all reflections. Cmp is completeness of data for the sphere to d_{min}.

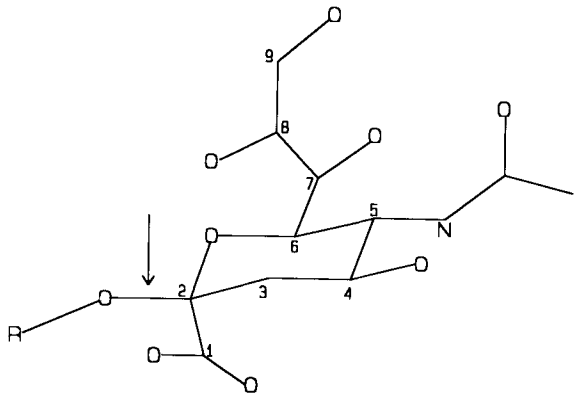

Fig. 1. α-Sialic acid (R = H), or *N*-acetyl neuraminic acid. The bond cleaved by bacterial neuraminidases is indicated by an arrow.[30] 2-Deoxy-2,3-dehydro-*N*-acetylneuraminic acid is an unsaturated derivative of sialic acid in which C2 is trigonal and in the plane of C6–O–C3–C4.

NANA1 and DANA data sets, and the NANA2 data were collected on image plates with Weissenberg geometry[13,14] at the Photon Factory synchrotron source. Details of the data collection statistics for these complexes are shown on Table I.

Difference Fouriers of NANA1-neuraminidase, DANA-neuraminidase, and NANA1-DANA, all at 2.9 Å resolution, were calculated to determine the location and orientation of the inhibitors in the enzyme. All difference maps were computed using phases from the uncomplexed neuraminidase structure.[4] Subsequently the NANA2 and DANA complex structures were independently refined by molecular dynamics[15] to 1.8 and 2.9 Å resolution respectively (Table I). The FANA structure was not refined as only limited data were collected on this complex. All charges on protein side chains were removed, but the substrate carboxylate charge was retained during refinement. For NANA2, the refined uncomplexed[4] neuraminidase structure containing 3,022 protein atoms, 164 carbohydrate atoms, and 85 water molecules was used as the starting model and a full cycle of the simulated annealing protocol was carried out on this model. A $2F_o - F_c$ difference Fourier was then calculated and the sialic acid moiety was built[16] into the largest positive feature in the map in a similar location and orientation in the active site as observed in the NANA1-neuraminidase difference map. Another cycle of energy refinement was performed with the dihedral restraints removed on angles C5–C4–C3–C2 and C5–C6–O–C2 in the pyranose ring of the substrate (see Fig. 1). A similar procedure was carried out for the DANA structure, albeit with fixed geometry for the unsaturated pyranose ring. The DANA structure is at much lower resolution and has converged to a crystallographic R value of 0.15 with rms deviations from ideality of 0.015 Å in bondlengths, 3.63° in bond angles, and 26.9° in dihedral angles. The comparative values for NANA2, including partial data to 1.8 Å resolution, are 0.22, 0.017 Å, 3.7°, and 27°. The NANA2 data set is 95% complete to 3.1 Å, and in shells from 3.1–2.5, 2.5–2.2, 2.2–2.0, and 2.0–1.8 Å the completeness is 63, 38, 25, and 19%, respectively. Only reflections for which F is greater than $2\sigma(F)$ were included.

RESULTS

The difference Fourier analyses lead to the conclusion that sialic acid binds to neuraminidase in a distorted configuration with the carboxylate in the equatorial position. The orientation of the inhibitor moiety is further confirmed by (1) the NANA1–DANA difference map which shows a single significant feature of density located at the position of the glycosidic oxygen in the axial position of the distorted structure of sialic acid and pointing away from the active site (Fig. 2) and (2) the FANA–DANA difference map which shows density only around the methyl group of the acetamido side chain as expected for the trifuoro substitution (data not shown). The difference density of DANA-neuraminidase indicates that the orientation of the unsaturated inhibitor is similar to that of sialic acid, with the carboxylate group in an equatorial position at a location close to that of the carboxylate in the bound sialic acid.

The final atomic model for the sialic acid complex is shown in Figure 3, and schematically in Figure 4, confirming the interpretation of the difference Fourier studies. Sialic acid binds to the enzyme through the same face as is used in its interaction with hemagglutin.[7] The NH group of the acetamido side chain is directed toward the floor of the active

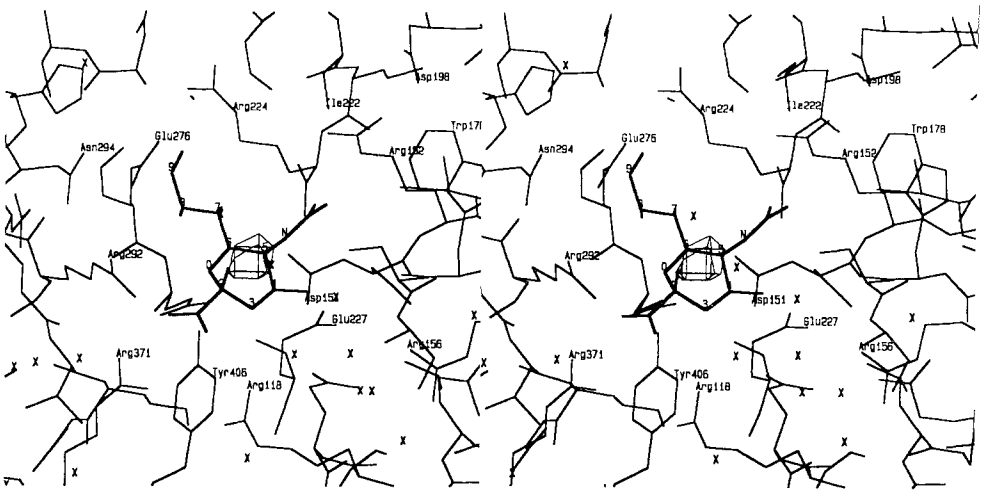

Fig. 2. A stereo drawing of the difference electron density using phases from the uncomplexed neuraminidase structure[4] and amplitudes NANA1-DANA. The refined structure of the sialic acid complex is overlaid. The peak shown coinciding with the glycosidic oxygen has a peak height of $+5.9\sigma$. Noise peaks are less than 4σ, the contour level.

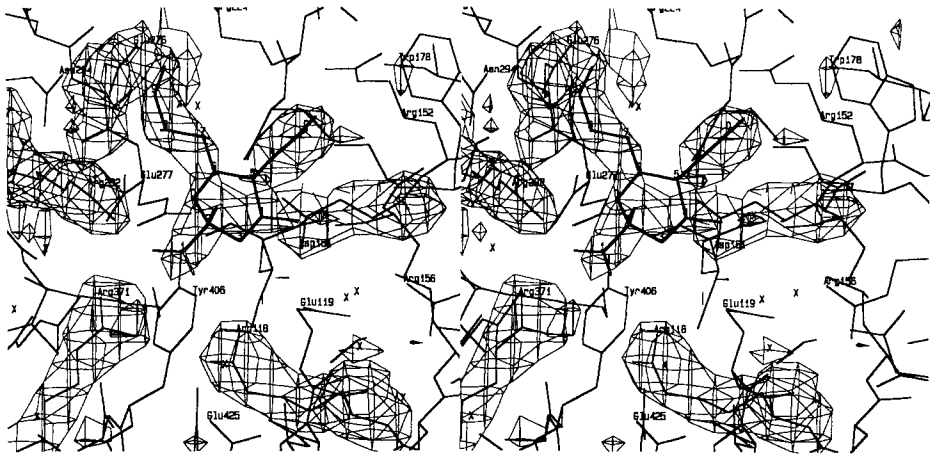

Fig. 3. Stereo diagram of the $F_o - F_c$ Fourier map with phases from the refined structure of NANA2 (Table I) except that the atoms of side chains of residues 118, 151, 276, 292, and 371 and of the sialic acid have been omitted from the structure factor calculation. The atomic model is overlaid. The conformation of Arg-152 is unusual. An alternative structure which can be accommodated nearly as well by the density places NE and NH_2 in bidentate hydrogen bonding with the carboxylate of Asp-198. In N9 strains of influenza, residue 198 is Asn, and refinement of N9[5] indicates that the conformation of Arg-152 shown here is consistent with that structure.

site cavity where it makes a hydrogen bond to a bound water molecule. The carboxylate group, axial toward the floor in the undistorted structure of the sugar, is held equatorial, possibly by reason of interactions it makes with the three guanidinium groups of arginine residues at sequence positions 118, 292, and 371. Refinement of the bound sialic acid (Table I) has resulted in a boat configuration for the pyranose ring. The oxygen of the acetamido group is hydrogen bonded to NE of Arg-152 and the methyl group lies in hydrophobic pocket near Ile-222 and Trp-178, but not over the six-membered ring of the tryptophan residue as observed in the hemagglutinin complex.[7] Hydroxyl groups of the glycerol side chain are hydrogen bonded to Glu-276, and the 4-hydroxyl is directed toward Glu-119. Features of the saccharide–protein interaction which are generally observed in other systems[17] include the participation of many planar, polar amino acids and the underlying network of hydrogen bonds in the protein which hold those groups in position to receive the sugar.

All of these interactions are similar to those observed in the refined DANA complex. Shown in Figure 5 is a superposition of the refined inhibitor structures in the active site. The orientation of the glycerol and acetamido side chains is similar and the saturated and unsaturated structures are distin-

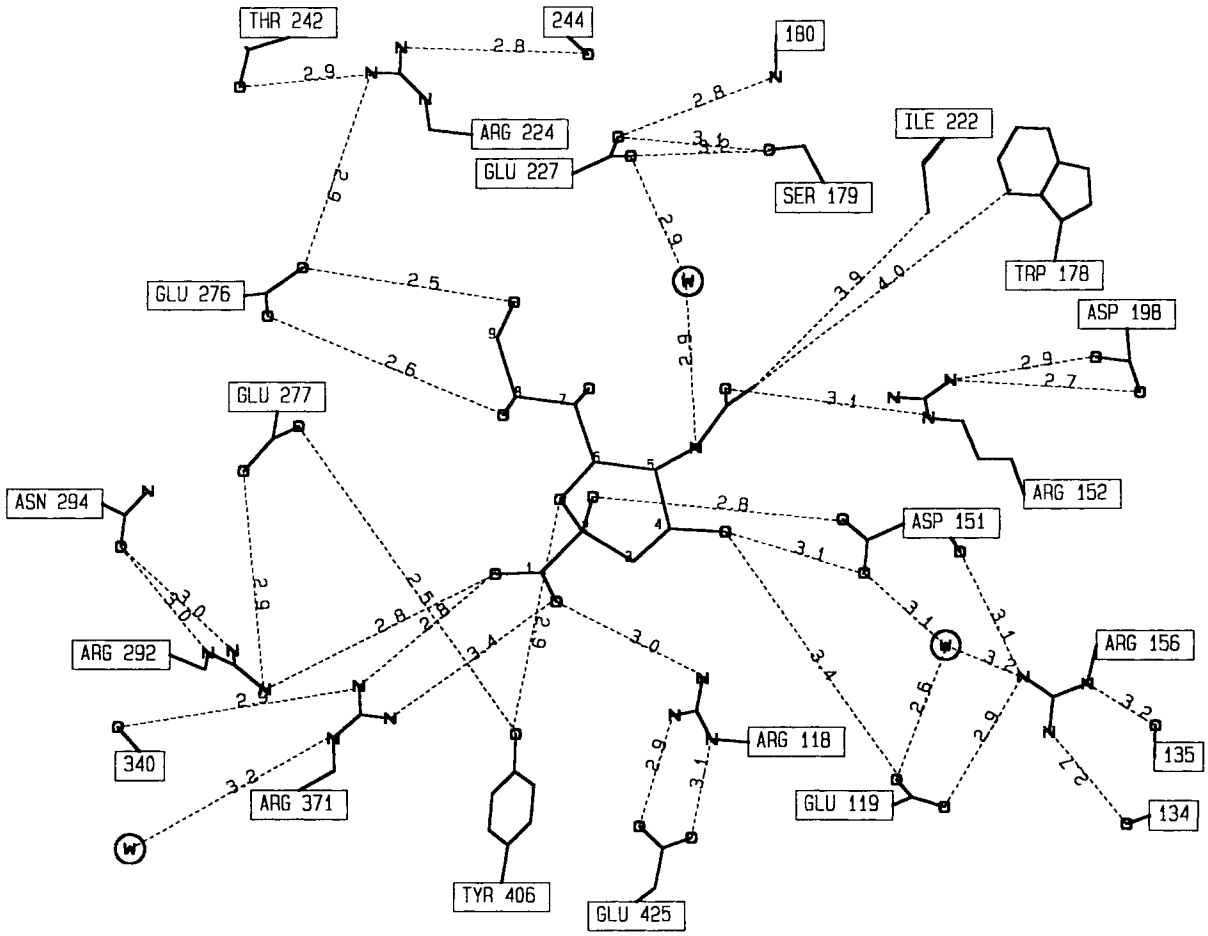

Fig. 4. Schematic representation of interactions within the catalytic side of neuraminidase and between the enzyme and the product, sialic acid. The estimated errors[31] in the coordinate are 0.27 Å.

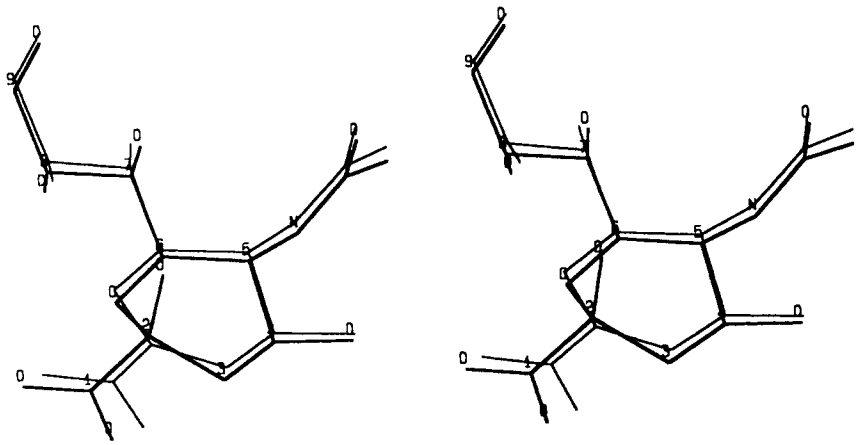

Fig. 5. A stereo drawing of the superposition of refined *N*-acetyl neuraminic acid (bolder lines) and 2-deoxy-2,3-dehydro-*N*-acetyl neuraminic acid structures in the active site. The numerals refer to the carbon atom positions. Excluding the glycosidic oxygen from the comparison the mean rms difference between the two inhibitors is 0.31 Å. The dihedral angles of the sialic acid pyranose ring determined from C5 to C2 via C4 and C3, and via C6 and O is 14° and 9°, respectively, and for the unsaturated inhibitor, these angles were 30° and −20°. These dihedral angles were unrestrained during refinement in both cases.

guished crystallographically only by the presence of the glycosidic oxygen, which, because of the distortion at C2 raising the carboxylate to the equatorial position, is axial and directed vertically out of the binding pocket. In this position its closest neighbour on the enzyme is Asp-151, OD2 (Fig. 4). This protein atom is remarkable because of its failure to participate in the complex hydrogen bonding network lining the active site. A parallel with hen egg white lysozyme is suggested, though not yet proven, whereby Asp-151 acts as a general acid in the catalysis by protonating the glycosidic oxygen.

Comparison with the uncomplexed N2 neuraminidase structure[4] suggests that Glu-276 and Asp-151 may undergo small conformational changes on binding substrate. These changes are near the limit of detection at the current resolution and model accuracy, and are not apparent in difference Fourier maps.

DISCUSSION

Glycosidic bond cleavage by lysozyme[18] is facilitated by a pair of carboxylate groups on the enzyme. One of them, Glu-35, is believed to be the general acid and the other, Asp-52, is thought to stabilize the developing positive charge on a carbonium ion intermediate. The bacteriophage T4 lysozyme contains an analogue of Asp-52 but the proton donor is uncertain and probably acts indirectly.[19] There is no carboxylate analogue of Asp-52 in the neuraminidase structure. The closest approach of neuraminidase to the hetero-atom of the pyranose ring is through the phenolic oxygen of Tyr-406 which may direct one of its lone pairs at the ring oxygen. The sialic acid carboxylate itself may also assist in the charge stabilization of the intermediate. The reaction could be completed through the activation of a water molecule by the deprotonated Asp-151, and its attack on the carbonium ion, resulting in the formation of the α anomer of sialic acid. This is the form in which the product is released from the active site of bacterial neuraminidases before the α/β anomeric equilibrium in solution is established.[20]

This "lysozyme-like" mechanism is not consistent with all the available data on the enzyme activity. Some data suggest a bell-shaped pH activity profile for the enzyme[21,22] with critical groups titrating at around pH 4.5 and 9.0. The alkaline pK_a is believed to correspond to a cationic acid on the enzyme interacting with the carboxylate of the sugar.[22] However, there is no evidence in those data for an analogue of Glu-35, pK_a 6.5, in lysozyme. The role of Asp-151 in the neuraminidase mechanism is therefore uncertain. The low pH (4.5) cut off may be a result of either protonation of the substrate or of the glutamates 276 or 277, which are on the floor of the site, the former hydrogen bonded to the glycerol side chain, and the latter hydrogen bonded to Tyr-406. It is unclear how to rationalize data from site directed mutants[21] implicating His-274 in modulating the high pH activity profile. Any role for Tyr-406 seems unlike the general acid/base function proposed for a tyrosine residue in β-galactosidase,[23] although on structural grounds alone the possibility of its involvement in a covalent intermediate cannot be excluded. Rotation around the Cα–Cβ bond of Tyr-406 allows a close approach of the phenolic oxygen to the substrate atom C2 of 2.6 Å for substrate in the observed (boat) conformation. This contact distance would be considerably less if the pyranose ring collapsed back to chair.

A further analogy with lysozyme is the observation of distortion of the bound sugar in the active site,[24] in both cases resulting in structures approximating the proposed oxocarbonium ion transition state. Although there has been some controversy over the geometrical distortion of lysozyme substrates, there is clear evidence of this distortion through the location of the sialic acid carboxylate in the studies reported here. A combination of hydrogen bonding and steric interference is believed to underlie the strained conformation of *N*-acetylmuramic acid in lysozyme.[24] In neuraminidase, the electrostatic effect of the triarginyl cluster (residues 119, 292, and 371) is one influence on the geometry of the bound sialic acid. Unfavorable packing and electrostatic interactions of the carboxylate with the floor of the site are other factors. The contribution of this strained geometry to the cleavage reaction remains to be determined. Molecular mechanics[25] and dynamics[26] calculations both suggest that the role of mechanical strain in the lysozyme hydrolytic mechanism is minimal.

The position of the glycosidic oxygen atom in the bound substrate suggests that there is limited potential for binding interactions between the enzyme and the penultimate sugar from which sialic acid is cleaved. If this is indeed the case, then most of the binding energy must derive from interactions between the enzyme and the product, α-sialic acid. Nevertheless, release of product from the active site is favored by the preferred β conformation for the sugar in solution; β anomers of sialic acid are neither substrates nor inhibitors of the enzyme.[27]

The approach of the substrate into the active site places the aglycon moiety near neuraminidase residues 368–370. However after distortion of the sialic acid, the aglycon lies directly above the site, well removed from residues 368–370. The partial inhibition of enzyme activity on trisaccharide substrates by the monoclonal antibody NC41[28,29] whose binding site includes residues 368–370, may be explained by the proximity of the aglycon to the antibody during the approach of substrate into the active site. Proposed distortion of the 368–370 loop by the antibody[29] has not been confirmed by structure refinement[32] and is not essential to a rationalization of the inhibitory behavior NC41. We are cur-

rently analyzing the structure of subtype N9 neuraminidase–sialic acid complex which, together with solution studies of the reaction, should shed further light on the enzyme mechanism.

ACKNOWLEDGMENTS

We thank Prof. N. Sakabe for assistance with data collection at the Photon Factory, R.G. Webster for the influenza virus strain used, Glaxo Group Research for the gift of compound for the FANA study, W.G. Laver for neuraminidase used in the FANA study, Paul Davis for assistance with computing, A. van Donkelaar and D. Marshall for technical assistance, M. von Itzstein for discussions, and Biota Holdings Ltd. for financial support.

REFERENCES

1. Gottschalk, A. In "The Viruses." Burnet, F.M., and Stanley, W.M. (eds.). New York: Academic, 1959: 51–61.
2. Gottschalk, A. Neuraminidase: The specific enzyme of influenza virus and vibrio cholerae. Biochim. Biophys. Acta 23:645–646, 1957.
3. Varghese, J.N., Laver, W.G., Colman, P.M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature (London) 303:35–40, 1983.
4. Varghese, J.N., Colman, P.M. The three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. J. Mol. Biol. 221:473–486, 1991.
5. Tulip, W.R., Varghese, J.N., Baker, A.T., Van Donkelaar, A., Laver, W.G., Colman, P.M. The refined structures of N9 subtype influenza virus neuraminidase and escape mutants. J. Mol. Biol. 221:487–497, 1991.
6. Colman, P.M., Varghese, J.N., Laver, W.G. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature (London) 303:41–44, 1983.
7. Weis, W., Brown, J.H., Cusack, S., Paulson, J.C., Skehel, J.J., Wiley, D.C. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature (London) 333:426–431, 1988.
8. Meindl, P., Bodo, G., Palese, P., Schulman, J., Tuppy, H. Inhibition of neuraminidase activity by derivatives of 2-deoxy-2,3-dehydro-N-acetylneuraminic acid. Virology 58: 457–463, 1974.
9. Miller, C.A., Wang, P., Flashner, M. Mechanism of *Arthrobacter sialophilus* neuraminidase: The binding of substrates and transition-state analogs. Biochem. Biophys. Res. Commun. 83:1479–1487, 1978.
10. Flashner, M., Kessler, J., Tannenbaum, S.W. The interaction of substrate-related ketals with bacterial and viral neuraminidase. Arch. Biochem. Biophys. 221:188–196, 1983.
11. McKimm-Breschkin, J.L. Caldwell, J.B. Guthrie R.E. Kortt, A.A. A new method for the purification of the influenza A virus neuraminidase. J. Virol. Methods 32:121–124, 1991.
12. Rossmann, M.G. Processing oscillation diffraction data for very large unit cells with an atomatic convolution technique and profile fitting. J. Appl. Crystallogr. 12:225–238, 1979.
13. Sakabe, N. X-ray diffraction data collection system for modern protein crystallography with Weissenberg camera and an imaging plate using synchrotron radiation. Nuc. Inst. Meth. Phys. Res. A303:448–463, 1991.
14. Higashi, T. The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. J. Appl. Crystallogr. 22:9–18, 1989.
15. Brunger, A.T. Crystallographic refinement by simulated annealing: Application to a 2.8 Å resolution structure of Aspartate Aminotransferase. J. Mol. Biol. 203:803–816, 1988.
16. Jones, T.A. Interactive computer graphics: FRODO. Methods Enzymol. 115:157–171, 1985.
17. Quiocho, F. Carbohydrate-binding proteins: Tertiary structures and protein-sugar interaction. Annu. Rev. Biochem. 55:287–315, 1986.
18. Phillips, D.C. The three-dimensional structure of an enzyme molecule. Sci. Am. 215:78–90, 1966.
19. Anderson, W.F., Grutter, M.G., Remington S.G., Weaver, L.H., Matthews, B.W. Crystallographic determination of the mode of binding of oligosaccharides to T4 bacteriophage lysozyme: Implications for the mechanism of catalysis. J. Mol. Biol. 147:523–543, 1981.
20. Friebolin, H., Brossmer, R., Keilich, G., Ziegler, D., Supp, M. ^{1}H-NMR-spektroskopischer nachweis der N-Acetyl-α-D-neuraminsaure als primares spaltprodukt der neuraminidasen. Hoppe-Seyler's Z. Physiol. Chem. 361:697–702, 1980.
21. Lentz, M.R., Webster, R.G., Air, G.M. Site-directed mutation of the active site of influenza neuraminidase and implications for the catalytic mechanism. Biochemistry 26: 5351–5358, 1987.
22. Chong, A.K.J., Pegg, M.S., von Itzstein, M. Characterisation of an ionisable group involved in binding and catalysis by sialidase from influenza virus. Biochem. Int. 24:165–171, 1991.
23. Ring, M., Bader, D.E., Huber, R.E. Site-directed mutagenesis of β-galactosidase (*E. coli*) reveals that TRY 503 is essential for activity. Biochem. Biophys. Res. Commun. 152:1050–1055, 1988.
24. Strynadka, N.C.J., James, M.N.G. Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. J. Mol. Biol. 220:401–424, 1991.
25. Warshel, A., Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J. Mol. Biol. 103:227–249, 1976.
26. Post, C.B., Karplus, M. Does lysozyme follow the lysozyme pathway? An alternate based on dynamic, structural, and stereoelectronic considerations. J. Am. Chem. Soc. 108: 1317–1319, 1986.
27. Meindl, P., Tuppy, H. Ubersynthetische Ketoside der N-Acetyl-D-neuraminsaure 2. Mitt. Mh Chem. 96:816–827, 1965.
28. Webster, R.G., Air, G.M., Metzger, D.W., Colman, P.M., Varghese, J.N., Baker, A.T., Laver, W.G. Antigenic structure and variation in an influenza virus N9 neuraminidase. J. Virol. 61:2910–2916, 1987.
29. Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M., Webster, R.G. Three dimentional structure of a complex of antibody with influenza virus neuraminidase. Nature (London) 326:358–363, 1987.
30. Deizl, C.M., Kamerling, J.P., Vliegenthart, J.F.G. Release of sialic acid from substrates by sialidase in the presence of oxygen-18-labeled water. Carbohydrate Res. 126:338–342, 1984.
31. Luzzatti, P.V. Treatment of statistical errors in the determination of crystal structures. Acta Crystallogr. 5:802–810, 1952.
32. Tulip, W.R., Varghese, J.N., Laver, W.G, Webster, R.G., Colman, P.M. Refined crystal stucture of the influenza virus Ng neuraminidase-NC41 fab complex. J. Mol. Biol., in press, 1992.