Molecular Modeling Studies on Ligand Binding to Sialidase from Influenza Virus and the Mechanism of Catalysis

Neil R. Taylor and Mark von Itzstein*

School of Pharmaceutical Chemistry, Victorian College of Pharmacy, Monash University, 381 Royal Parade, Parkville, 3052, Victoria, Australia

Received September 2, 1993®

A molecular modeling study has been used to investigate the structural and energetic aspects of substrate and inhibitor binding and the mechanism of catalysis of influenza virus sialidase. A detailed analysis of the interactions of both *N*-acetylneuraminic acid (Neu5Ac, 1) and a number of transition-state analogues with the active site of influenza A sialidase at an atomic level is reported. In each case the calculated structures favorably agreed with the results from X-ray studies. A qualitative agreement between the calculated binding energies for inhibitors with positive substituents at the C4 position on the sugar ring and experimental K_i values was observed. We propose that the hydrolysis of sialosides occurs via an S_N1 type mechanism that is facilitated through an activated solvent water molecule which can be expelled upon inhibitor binding. A reaction scheme is presented that is consistent with previously observed crystallographic structures, anomeric products, and isotope effects.

Introduction

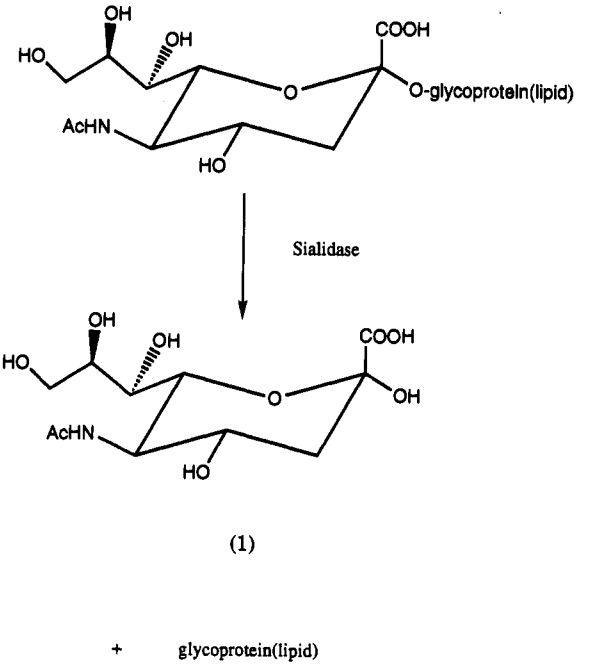
Sialidases (E.C. 3.2.1.18) are glycohydrolases which catalyze the cleavage of terminal sialic acids α-ketosidically linked to glycoproteins, glycolipids, and polysaccharides,[1] and the general reaction is shown in Scheme 1. There are a large number of biological functions ascribed to this enzyme such as cell–cell recognition phenomena and the pathogenicity of some infections by sialidase-bearing microorganisms.[2] In the case of influenza virus, sialidase is one of two surface glycoproteins and is considered to be important for both transporting the virus through mucin[3] and for the elution of virus progeny away from infected cells.[4] The structure of the enzyme has been extensively characterized from sequence studies on numerous antigenic variants[5] and from several crystallographic studies.[6-9] A key feature arising from these investigations is that despite up to 50% sequence variation, the enzymes have similar three-dimensional structures, and the amino acid residues that line the active site are highly conserved in both influenza A and B virus strains.

Sialidase from influenza virus has a molecular weight of approximately 240 kDa and consists of a mushroom-shaped head of four identical subunits in a square-planar arrangement which is anchored to the viral coat at the N-terminus by a long thin stalk.[6-8] Each subunit of the enzyme is made up of six four-stranded β-sheets arranged as if on the blades of a propeller. *N*-Acetylneuraminic acid (1), the product of catalysis, binds in a large pocket on the upper surface of the subunit which is lined with an unusually large number of charged amino acid residues. Two calcium binding sites have been located in the protein, one near the binding pocket[8] and the second on the 4-fold axis of the tetramer.[6] It has been postulated that calcium activation of influenza sialidase may have a physiological role in the positive modulation of enzyme activity during the release of newly synthesized virions from the host cell surface.[10]

Our interest in sialidases,[11] and in particular influenza virus sialidase,[10,12,13] toward the design and synthesis of novel influenza drugs[14] has given us the opportunity to study the mechanism of action of influenza virus sialidase. We have recently published a study[13] which details kinetic isotope effect, nuclear magnetic resonance spectroscopy, and preliminary molecular dynamics simulation experiments of the sialidase–substrate complex. From this study we postulated that the hydrolysis of substrate was solvent mediated and involved a sialosyl cation transition-state intermediate. The binding of substrate induces the sialic acid part of the sialoside to distort from a $^2C_5$ conformation to a boat-like conformation which facilitates the cleavage of the substrate, yielding sialic acid and the aglycon.[13,14]

In order to better understand the interactions between sialidase and substrate, we have now completed a comprehensive series of molecular modeling studies on the complex of *N*-acetylneuraminic acid (Neu5Ac, 1) with influenza virus A/Tokyo/3/67 sialidase. In this paper we elaborate a more refined mechanism for sialidase catalysis and discuss the experimental and theoretical evidence for

Scheme 1



(1)

+ glycoprotein(lipid)

® Abstract published in *Advance ACS Abstracts*, January 15, 1994.

0022-2623/94/1837-0616$04.50/0

the major steps of the reaction. Finally, the geometric and energetic features associated with the binding interactions of several competitive inhibitors of influenza A sialidase are also presented.

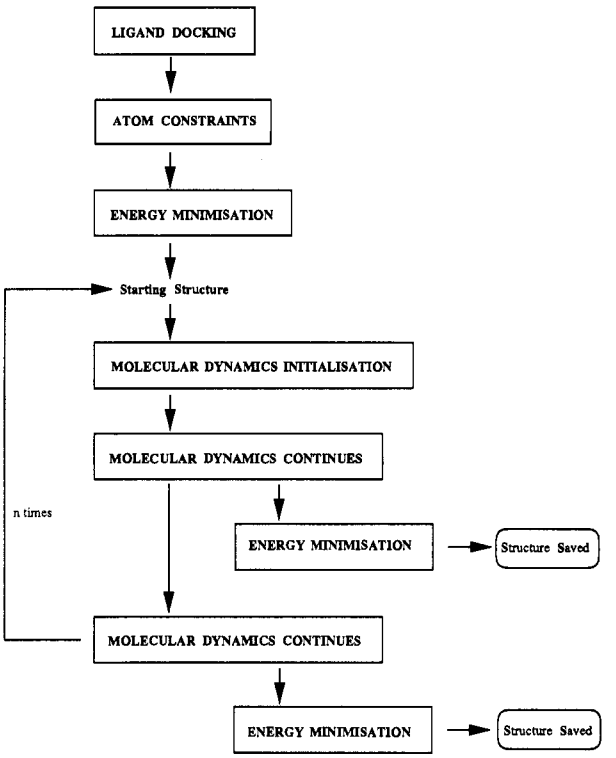
Methods

The experiments in the present study were based upon high-resolution crystal coordinates of the complex of Neu5Ac (1) with a subunit of influenza A sialidase obtained from Dr. Peter Colman (Division of Biomolecular Engineering, CSIRO, Parkville, Australia). The crystal structure of the complex included Neu5Ac (1), 388 protein residues (residues 82 to 469), 117 water molecules, four oligosaccharide groups, and one calcium atom (near the active site of the enzyme). The water molecules for which temperature factors were determined were labeled with the prefix "W" (86 in total) by the crystallographers, and the remaining waters whose temperature factors were not determined were labeled with the suffix "X". This scheme was maintained in our modeling studies. All water molecules external to the protein and not hydrogen bonded to any amino acid residues (31 of type W and 18 of type X) were deleted prior to the calculations. The four oligosaccharide groups were also removed; these groups were assumed to be unimportant in our study. Hydrogen atoms, partial charges, atomic potentials, and bond orders were assigned using the automatic procedures within the Insight II 2.1 package (Biosym Technologies Inc., San Diego, CA). The carboxylic acid of Neu5Ac (1) was ionized with the oxygens assigned equal charges. Arginine, lysine, aspartate, and glutamate amino acid residues were charged while the histidines were uncharged, with hydrogen atoms fixed to the $N\epsilon$ atoms.

From a crystallographic study of influenza B sialidase, a water molecule in the coordination sphere of the calcium ion adjacent to the active site was located.[15] Sequence homology strongly suggested equivalent calcium ion binding sites in influenza A and B sialidases, and so this water molecule, which was not reported for influenza A, was included in the calculations. The charge on this calcium ion was fixed to 1.80 instead of 2.0 to prevent the coordinate bond distances from becoming too short. Finally, a ligated unidentate carboxylate group belonging to residue Asp 324 was polarized with an asymmetry of 120/80 to maintain the metal ion coordination expected from the crystal structures (Dr. R. Doelz, personal communication). Also included in the calculations was water molecule 6X which is adjacent to bound Neu5Ac (1). This water is not included in the recently reported crystal structure in the Protein Data Bank (PDB) at the Brookhaven National Laboratory due to poor resolution (Dr. J. Varghese, personal communication). We have included this "mobile" water in our work because without it a large void exists in the enzyme in an important part of the active site. It is interesting to note that a water molecule at an equivalent position to that of 6X was identified in the crystal structure of influenza B sialidase.[15] The crystal structure of influenza virus B sialidase was obtained from the PDB.

The effects of solvation in an aqueous medium were modeled by partially solvating the enzyme–Neu5Ac complex and by using a distance-dependent dielectric constant. In order to prevent the residues on the protein surface from deviating significantly from their positions in the crystal structure, and to accurately simulate electrostatic energies near the protein surface, the active site of sialidase was solvated by a 15.0-Å layer of water molecules, and all of the remaining charged surface residues were solvated with a 3.0-Å layer of waters using the SOAK option of Insight II. A total of 425 water molecules were added to the system. The distance-dependent dielectric constant ($\epsilon = R$) was incorporated into the Coulombic potential to simulate the electrostatic shielding effects of dipolar reorientation and electronic polarization by the polar bulk solvent. This value for the dielectric constant was considered the most appropriate because it produces the best results for short-range electrostatic forces. Larger value of ϵ (e.g. $4R$) were found to disrupt the hydrogen-bonding networks observed in the crystal structure of the complex to a significant degree. The system studied can be considered to be that of a protein at infinite dilution in an aqueous environment without any counterions.

Scheme 2



Calculations were performed using the Discover 2.7 package (Biosym Technologies Inc., San Diego, CA) on a Silicon Graphics Iris workstation (Silicon Graphics, Mountain View, CA). Energy minimizations were carried out using the consistent valence force field[16] incorporating the simple harmonic function for bond stretching and excluding all nondiagonal terms. The inclusion of nondiagonal terms and the use of the Morse potential were found to give very similar results. A cutoff distance of 12.0 Å was used. A switching function operating over a 1.5-Å and a 2.0-Å buffer region was included. The nonbonded pair list was updated every 20 steps. The crystal structure of the sialidase–Neu5Ac complex was subjected to the well-documented multistage minimization procedure in which the atoms were unconstrained in stages to remove strained interactions.[16] First, solvent molecules were energy minimized with all other atoms fixed, and then all protein and Neu5Ac (1) hydrogen atoms included in the minimization process, followed by the protein side chains, and finally the entire system was free to move. Calculations were done using the algorithm's steepest descents (in all constrained minimizations for up to 500 iterations) and conjugate gradients (down to a maximum atomic root mean square derivative of 0.01 kcal/mol Å). All results were visualized using Insight II.

Conformational space of the enzyme–ligand complex within the active site and in the vicinity of the crystal-structure coordinates was explored using two techniques. The first involved varying the starting conditions in the energy-minimization calculations, such as the nonbonding cutoffs, convergence criteria, and dielectric constant. The second approach involved a combination of constrained molecular dynamics and molecular mechanics calculations. Scheme 2 illustrates the procedure. Beginning with the energy-minimized sialidase–Neu5Ac complex, a ligand was docked into the active site with its position and conformation as similar to that in Neu5Ac (1) as possible. Solvent molecules were then added or subtracted from the area around the docked ligand depending on the space requirements. The complex was constrained throughout all calculations; only the protein residues of the active site (defined in this case to be residues within 16.0 Å of the bound ligand), ligand, calcium ion, and water molecules around the active site, were free to move. The model system thus consisted of the docked ligand, 115 amino acids, 28 crystallographic waters, one calcium ion, and 48 solvent molecules. In addition to the atom constraints, solvent molecules within the active site were subjected to harmonic restraints, with a force constant of 0.05 kcal/mol Å^2, to accurately model solvent

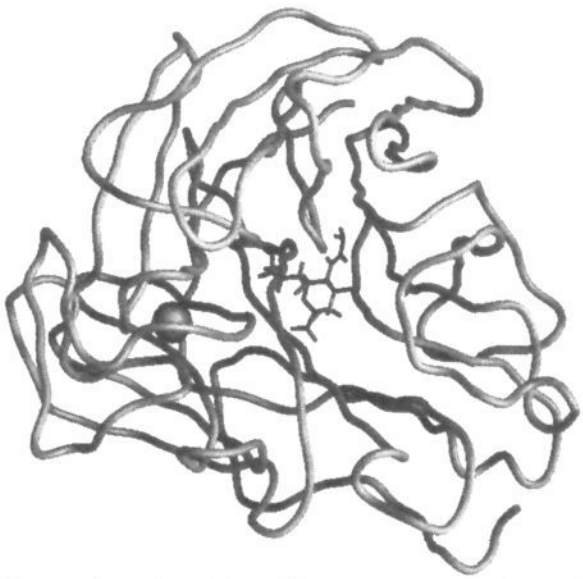


Figure 1. A top view of the sialidase subunit showing the protein backbone, sialic acid bound to the protein loops that define the active site, and the calcium ion adjacent to the binding region (generated using GRASP software, Nicholls and Honig, Columbia University).

mobility.[17] An initial energy minimization was performed to relieve steric strain associated with the newly docked ligand. Molecular dynamics calculations were then performed at 350 K using the leapfrog algorithm in Discover 2.7. Dynamics were equilibrated for 2 ps with time steps of 1 fs and then continued for 2 ps with time steps of 2 fs. The resulting structure was extracted and energy minimized. A further 2 ps of dynamics with time steps of 2 fs were performed, with the final structure again energy minimized. This sequence was repeated 5–10 times beginning at the molecular dynamics initialization stage. This approach of continually resetting the trajectory of the molecular dynamics calculations and using a relatively low temperature was found to be an effective method for exploring conformational space with a protein–ligand complex close to its lowest energy structure. Essentially, the reason for adopting our particular scheme was one of economics: the maximum number of conformations was obtained with the minimum amount of computing.

Results and Discussion

Energy Minimization. The minimized sialidase–Neu5Ac structure exhibiting the closest similarity with the crystal structure had a total of 425 added water molecules. After 4000 steps of steepest descents and 4800 steps of conjugate gradients, the positions of two water molecules, W1 and 14X, adjacent to the bound Neu5Ac (1) were adjusted to maximize hydrogen bonding interactions (see below). Further energy minimization using a few hundred steps of conjugate gradients was performed to obtain the final structure. The all-atom root-mean-square (RMS) deviation between the crystal and calculated protein structures is 1.12 Å. For heavy protein atoms the RMS deviation is 0.97. For heavy atoms of the active site, defined to be all residues within 16.0 Å of the hydroxyl oxygen on residue Tyr 406, the deviation is 0.85 Å.

A global view of the sialidase subunit is shown in Figure 1, and schematic diagram of the active-site, showing the calculated protein resudue–Neu5Ac (1) hydrogen bonding network, is shown in Figure 2. Table 1 lists the distances and angles associated with the interactions, and Table 2 shows the contribution of each protein residue adjacent to Neu5Ac (1) to the binding energy. As has been previously noted,[9,12,13] a salt bridge between the carboxylate

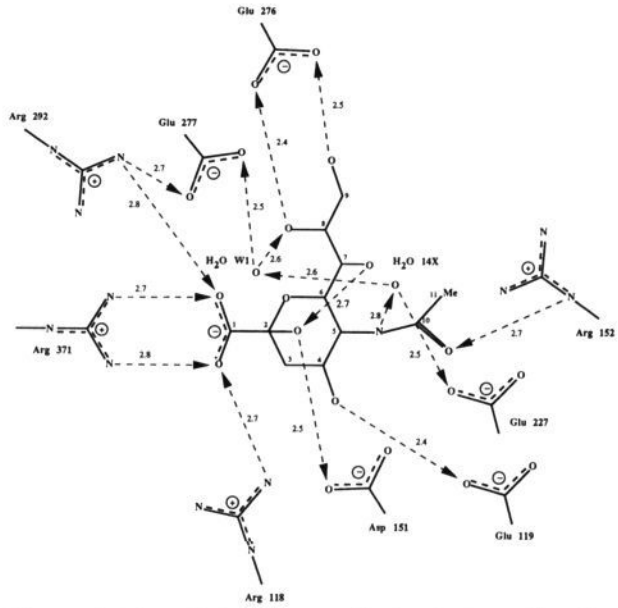


Figure 2. The calculated intermolecular hydrogen bonding network for the influenza virus sialidase–Neu5Ac complex (distances shown in Å).

Table 1. Hydrogen Bonds between Influenza Virus Sialidase and Neu5Ac (1) in the Calculated Lowest Energy Structure

donors (X)	acceptors (Y)	X···Y (Å)	X–H···Y (deg)
Arg 118 Nη2	O1A	2.67	164.4
Arg 371 Nη2	O1A	2.79	170.0
Arg 371 Nη1	O1B	2.68	174.0
Arg 292 Nη2	O1B	2.76	163.5
OH2	Asp 151 Oδ2	2.46	169.2
OH4	Glu 119 Oε1	2.38	173.0
NH5	H_2O 14X	2.84	163.9
Arg 152 Nε	O10	2.73	164.9
OH7	OH2	2.72	151.7
H_2O W1	OH8	2.65	147.6
OH8	Glu 276 Oε2	2.41	174.4
OH9	Glu 276 Oε1	2.46	166.8

Table 2. Intermolecular Binding Enthalpies (kcal/mol) between Influenza Virus Sialidase and Neu5Ac (1)

protein residue	ΔH (kcal/mol)
Arg 371	−48
Arg 292	−32
Arg 118	−28
Glu 276	−28
Arg 152	−19
Glu 119	−13
Arg 224	−8
Tyr 406	−7
Asp 151	−7
H_2O W1	−7
H_2O 14X	−4
Trp 178	−3
Ile 222	−2
total energy	−202

group of 1 and the charged Arg 371 contributes the greatest to binding. The charge–charge interactions between the same carboxylate and arginines 118 and 292 are also important. The hydroxyls on C8 and C9 form an important bidentate interaction to the negatively charged Glu 276. Hydrogen bonds also form between the anomeric hydroxyl group on C2 and Asp 151, between the hydroxyl on C4 and Glu 119, between the acetamido proton on the nitrogen and a structural water molecule, between the acetamido carbonyl oxygen atom and Arg 152, and between the

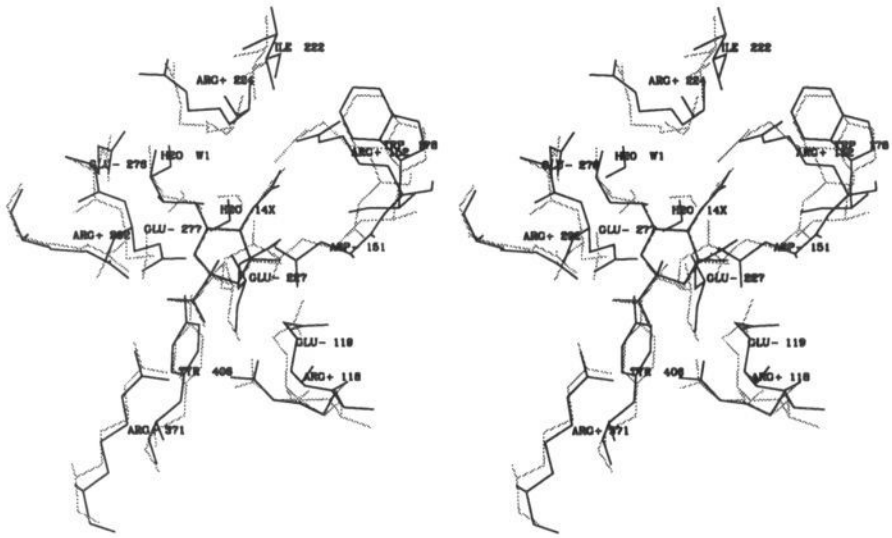


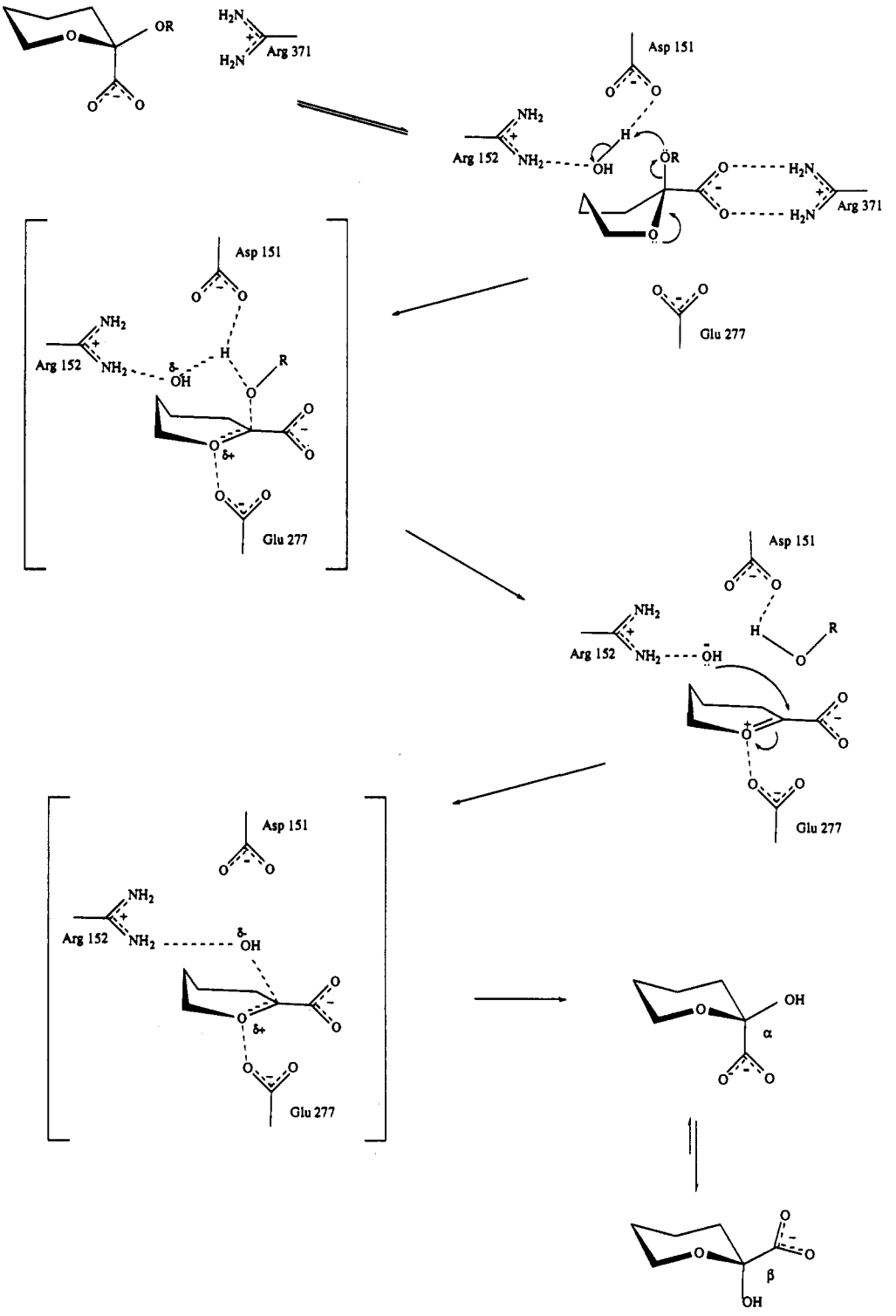
Figure 3. A stereoview of the binding sites of the observed (—) and calculated (···) sialidase–Neu5Ac (1) complexes (obtained from superimposing all ligand atoms).

hydroxyl group on C8 and a structural water. The hydroxyl on C8 is directed into the floor of the active site and is involved in a hydrogen-bonding network involving a structural water molecule and Glu 276. The ring oxygen appears to contribute only marginally to ligand binding through a weak charge–dipole interaction with Arg 292. Each of the hydroxyl groups on C2, C4, C7, and C9 accept protons from bulk solvent rather than from the active-site amino acid residues, so it would appear that they act essentially as hydrogen-bond donor groups in ligand binding. The acetamido carbonyl oxygen (C10) and the carboxylate oxygens also act as partial hydrogen bond acceptors with bulk solvent. Although electrostatic forces in the form of hydrogen bonds dominate the binding energy, a hydrophobic pocket in the active site is very important for molecular recognition. Protein residues Trp 178 and Ile 222 lie close to the methyl carbon (C11) at the terminus of the acetamido fragment and to the hydrocarbon backbone of the C6 glycerol side chain. The hydroxyl group on C7 does not form direct hydrogen bonds with the protein because it is directed out of the active site. There may, however, exist indirect bonds across water molecules to protein residues associated with this hydroxyl group. The calcium ion bound adjacent to the active site does not form important interactions with bound Neu5Ac (1). However, it has been suggested that this calcium might play a role as a biological switch for the activation of the enzyme,[10] and it is possible that it may also be important for providing a favorable electrostatic environment for attracting substrate toward the binding pocket.

A superimposition of the energy-minimized structure and the crystal structure is illustrated in Figure 3. The major differences between the two structures can be traced to difficulties in modeling the water molecules that bridge between active site residues and the bound ligand in the crystal structure of the enzyme–Neu5Ac complex. Approximately eight water molecules in the vicinity of Neu5Ac (1) form a hydrogen-bonded chain (which includes some active-site residues) that begins in the proximity of C3 in the sugar ring, passes around the C4 position, then beneath the acetamido nitrogen atom, and travels into the interior of the protein passing close to the terminus (C9) of the glycerol side chain. At the C4 position, for example, in the crystal structure the hydroxyl group is within hydrogen-bonding distance of two active site residues and one water molecule (OH4···Asp 151 Oδ2 = 3.09 Å, OH4···Glu 119 Oϵ2 = 3.36 Å, and OH4···H_2O 6X = 2.77 Å), whereas in the energy-minimized structure, a short hydrogen bond with residue Glu 119 was observed (see Table 1). In this case, a large displacement in the position of water molecule 6X in the energy-minimized structure relative to its position in the crystal structure (2.18 Å) made way for the strong hydrogen bond to form. Other difficulties arose with water molecule W1 and the adjacent functional groups. Firstly, W1 resides in a large cavity between enzyme and Neu5Ac (1), and secondly, the cavity is lined with many hydrogen bond acceptors and donors. In different energy-minimization calculations, therefore, W1 was found in a variety of positions of similar energy. The position and conformation for water W1 which was considered most reasonable was that which maximized the hydrogen-bonding connectivities for bound Neu5Ac (1). That is, a structure in which the hydroxyl groups of Neu5Ac (1) served simultaneously as hydrogen bond acceptors and donors. The optimization of the hydrogen-bonding network in the calculated structure results in some conformational changes from the crystal structure at active-site residues Glu 227 (hydrogen bonded to 14X, which is in turn bonded to W1), Glu 277 (hydrogen bonded to W1), and Tyr 406 (hydrogen bonded to Glu 277). Assigning a key role for W1 in the hydrogen-bond network between sialidase and Neu5Ac (1) is further justified in light of the recently published crystal structure of the influenza B structure. The calculated position of W1 shows excellent agreement with an experimentally observed water molecule in the active site (in crystal structure: H_2O Wat933···OH8 = 2.8 Å).[15]

The results from numerous energy minimizations under varying starting conditions and from the molecular dynamics conformational searches gave further insight into the physical properties of the enzyme–ligand interactions. Throughout all calculations the salt bridge between the carboxylate group on Neu5Ac (1) and Arg 371 was maintained, as were the interactions to arginines 118 and 292, with the latter residue playing slightly more of a role in binding than the former. The most interesting variation in the results involved Glu 276. Instead of the bidentate interaction observed in the crystal structure, in several

Scheme 3



calculations the hydroxyls on C8 and C9 of Neu5Ac (1) hydrogen bonded to residues Glu 277 and Glu 276, respectively (Glu 277 Oε2···OH8 and Glu 276 Oε1···OH9). The calculated energy of these interactions was essentially equivalent in value to that described previously. In yet another type of bonding arrangement calculated, the hydroxyl groups on C8 and C9 hydrogen bonded to only one of the oxygen atoms in Glu 276 (Oε1). Again, there was energetically no measured loss in this type of interaction. From a dynamics viewpoint, it would be expected from these results that interactions between the hydroxyls on C8 and C9 and the active-site residues would involve connectivities of these types to some extent.

Alternative intermolecular binding modes to that shown in Figure 2 were observed in the molecular dynamics results for all of the hydrogen bond donors in Neu5Ac (1). At the C2 position, for example, the hydroxyl functionality was occasionally found to form an intramolecular hydrogen bond of the type OH2···OH7. Such an interaction clearly does not contribute to the measured binding enthalpy. At the C4 position, the hydroxyl group was calculated to interact with the enzyme in three other modes to that described above, as anticipated from the crystallographic structure. A hydrogen bond from OH4 formed sometimes to water molecule 6X, sometimes to residue Asp 151, and sometimes to Glu 119 Oε2. Energetically the most

favorable of the different types of hydrogen bond are the interactions with Glu 119 (Oε1 or Oε2). The bond to Asp 151 is slightly less stable due to unfavorable electrostatic interactions. With the acetamido fragment on C5, the carbonyl oxygen atom was observed to hydrogen bond to Arg 152 in all experiments, always to Nε and sometimes to Nη1 as well. The NH moiety of the acetamido fragment hydrogen bonded to water 14X in some cases, water W1 in some cases, and water 6X in others, exemplifying the difficulties associated with modeling structural water molecules. The remaining hydroxyl group, OH7, was found to interact either with bulk solvent or in an intramolecular hydrogen bond with OH2, as described above. Clearly neither of these modes of bonding contribute directly to the total measured enzyme–Neu5Ac interaction energy.

Enzyme Mechanism. The catalytic pathway for sialidase can be regarded as consisting of four major steps. The first step is the binding event. The binding of sialoside to sialidase involves considerable distortion of the pyranose ring. In solution the Neu5Ac (1) pyranose ring adopts an expected $^2C_5$ chair conformation[14] whereas in the bound state the ring has a pseudoboat conformation. This conformation is the result of complex ionic, hydrogen-bond, and steric interactions. The carboxylate group transforms from the axial position into a pseudoequatorial position due to the charge–charge interactions to arginines 118, 292, and 371 and steric constraints with residues at the base of the active site, such as Tyr 406. The sp^2-hybridized (planar) nature of the sugar around the carboxylate carbon and the ring atoms O6, C2, and C3 activates the sialoside for hydrolysis.[13] The internal molecular strain energy associated with these distortions would be expected to be energetically fairly costly. However, the extensive hydrogen-bond network binding Neu5Ac (1) to the protein easily compensates for the free energy lost.

The second step of the catalytic reaction involves proton donation from solvent and formation of the endocyclic sialosyl cation transition-state intermediate. One of the most intensely studied and best known enzyme mechanisms is for the glycohydrolase, lysozyme. In this enzyme mechanism, the acidic residue Glu 35 is believed to donate a proton, and a second carboxylate group, Asp 52, is believed to stabilize the glycosyl cation intermediate.[18] Profiles from pH experiments[13,20] indicate that sialidase does not have an equivalent of the acidic Glu 35 residue, and so an alternative mechanism is required which does not involve general acid/base catalysis. From our previously published results of kinetic isotope studies on influenza virus sialidase with the synthetic substrate 4-(methylumbelliferyl)-Neu5Ac and the corresponding [3,3-^{2}H]-substituted substrate, an S_N1-type mechanism with proton donation from an activated water molecule and an endocyclic sialosyl cation transition-state intermediate had been postulated.[13] Two reaction pathways have been considered. In our first proposal we suggested that the crystallographic water molecule W102, which is located in the active site adjacent to the reaction center, could be activated by residue Asp 151.[13] The outcome of one particular molecular dynamics simulation was a hydrogen-bonded network with W102 bridging between the hydroxyl group on C2 and Asp 151 (Asp 151 Oδ2···H_2O W102···OH2). Other similar experiments clearly demonstrated that W102 can escape from its position in the crystal

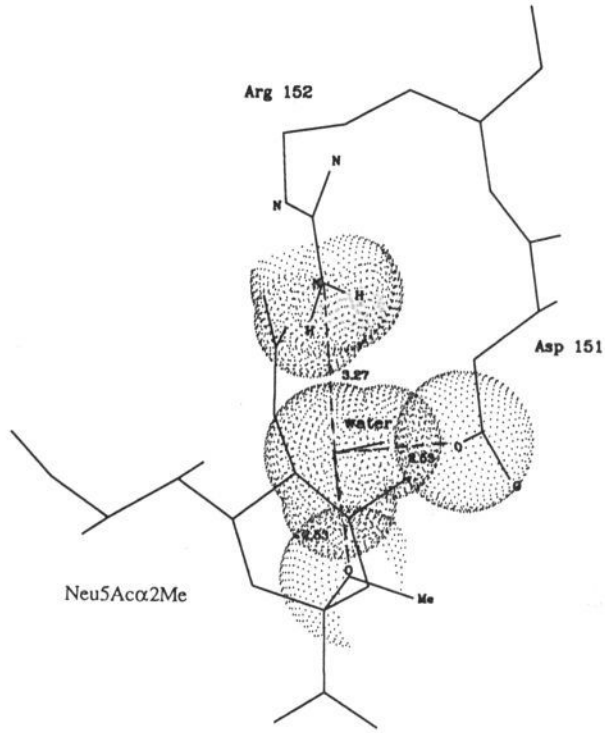


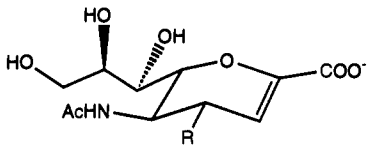
Figure 4. The atomic packing of groups around the reaction center in the active site of sialidase containing the substituted substrate Neu5Acα2Me (distances shown in Å).

structure and migrate toward the reaction center. It is conceivable that a hydrogen-bonding network of water molecules and protein residues leading from a charged group on the protein surface to W102 could facilitate proton donation.[20]

A molecular graphics examination of the fitting of active-site residues around C1 of bound Neu5Ac (1) suggested the involvement of a second residue, Arg 152. It was found that a solvent molecule is capable of packing into the active site of the enzyme–Neu5Ac complex between the charged residues Asp 151 and Arg 152 with a position and orientation optimal for proton donation onto the glycosidic oxygen atom. These reaction steps and the corresponding intermediates are shown in Scheme 3. This refined mechanism is also consistent with the pH properties of the enzyme.[12] Figure 4 illustrates the atomic packing, with the proton on the glycosidic oxygen atom replaced by a methyl group to mimic the glycosidic bond. Enzyme–Neu5Ac bonding distances observed over the course of molecular dynamics simulations indicate that Asp 151 and Arg 152 have terminal side-chain atoms not directly associated with the binding of substrate and the proposed transition-state intermediates (atoms Asp 151 Oδ1 and Oδ2 and Arg 152 Oη1 and Oη2), and similarly for Glu 277, which we believe contributes to the stability of the cationic intermediate.

Interestingly, a recent crystallographic study proposes that sialidase from influenza B virus catalyzes a low-frequency dead-end reaction in which sialic acid (1) is converted to Neu5Ac2en (**2a**).[21] The catalytic mechanism reported in this paper for the cleavage of sialic acid from glycoconjugates is in very good agreement with our proposed mechanism, the only difference being the implication of a tyrosine residue in stabilizing the cationic intermediate in conjunction with the glutamate residue.

The high binding affinity observed for transition-state



(2) (a) R = OH
(b) R = NH_3^+
(c) R = $NHC(NH_2)_2^+$

analogues such as 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en, **2a**)[12] can also be considered as evidence supporting our proposed mechanism. The double bond in Neu5Ac2en (**2a**) constrains the pyranose ring of the sugar into a planar structure around the ring oxygen, C2, and C3, thereby mimicking the proposed transition state, a sialosyl cation intermediate. Results from site-directed mutagenesis studies, in which enzyme activity is lost when Arg 152 is replaced with a lysine group and when Glu 277 is replaced by an aspartate residue,[22] are similarly consistent with our proposed mechanism. The shortened length of these amino acid residues would prevent the charged termini of the side chains from stabilizing the intermediates of the reaction.

The final two steps of the enzyme mechanism are the formation and release of Neu5Ac (1). NMR experiments indicate that Neu5Ac (1) is initially released as the α-anomer,[13] which is consistent with the proposed S_N1 mechanism having a high degree of stereofacial selectivity. It is conceivable that expulsion of product from the active site is favored by the mutarotation of the α-anomer to the thermodynamically more stable β-anomer (Scheme 3) for Neu5Ac (1) in solution.

Inhibitor Binding. The mode of binding of the three influenza sialidase inhibitors Neu5Ac2en (**2a**), 4-amino-Neu5Ac2en (**2b**), and 4-guanidino-Neu5Ac2en (**2c**) have been investigated.[14] Table 3 lists the K_i values for these compounds. The transition-state analogue and nonspecific sialidase inhibitor Neu5Ac2en (**2a**) differs from Neu5Ac (1) in two important ways: (i) removal of the hydroxyl group on C2 and (ii) planarization of the pyranose ring across the ring oxygen, C2, and C3 position. The first of these differences results in a decrease in the calculated intermolecular nonbonding energy, due to the loss of one hydrogen bond from the system (Asp 151 Oδ2···OH2), whereas the second results in insignificant changes in the calculated energies. Consideration of enzyme–ligand nonbonding interactions alone therefore leads to the conclusion that Neu5Ac2en (**2a**) binds weaker than Neu5Ac (1). However, the free energy of complexation is significantly more favorable for the unsaturated sugar because unlike Neu5Ac (1), which distorts from a $^2C_5$ chair conformation to a pseudoboat conformation, there are minimal increases in internal strain energy incurred upon binding of Neu5Ac2en (**2a**). Neu5Ac2en (**2a**) was calculated to bind to the active site through the same set of hydrogen-bonding interactions as Neu5Ac (1) at every functional group site they have in common. Equivalence in binding for the saturated and unsaturated structures was similarly observed in the crystallographic study.[9]

From earlier studies employing the GRID program[23] it was discovered that a positively charged amino group at the 4-position of Neu5Ac (1) (and Neu5Ac2en (**2a**)) could possibly form favorable interactions with active-site residue Glu 119.[14,24] The compound 4-amino-Neu5Ac2en (**2b**) was indeed found to be a potent inhibitor of influenza virus sialidases.[11,14] The energy minimum of the complex of this inhibitor (with a positively charged NH_3^+ group at C4) and the crystal structure of the enzyme, obtained from the multistage optimization procedure, is shown in Figure 5. The expected charge–charge type hydrogen bond with Glu 119 was calculated. In addition, however, active-site residue Asp 151 was found to form an even stronger interaction with the amino substituent. The crystallographic water molecule 6X also interacts with the amino group, thereby completing its coordination sphere. The distances and angles of these bonds are listed in Table 4. Calculated enthalpies of binding between the amino functionality and adjacent enzyme residues are listed in Table 5. It is believed that both of the charge–charge interactions are responsible for the high binding affinity for this molecule.

In all of the energy minima obtained from the molecular dynamics conformational searches of the sialidase–4-amino-Neu5Ac2en complex, the two charge–charge hydrogen bonds were maintained. In some of the minima Glu 119 was found to bond through Oε1 instead of Oε2 (as observed for the 4-hydroxyl functionality), whereas Asp 151 always hydrogen bonded through Oδ2. The hydrogen

Table 3. Inhibition of Influenza A Sialidase by Some Substituted Neu5Ac2en Analogues

inhibitor	K_i (M)[a]
2a	4×10^{-6}
2b	4×10^{-6}
2c	1×10^{-9}

[a] Reference 11.

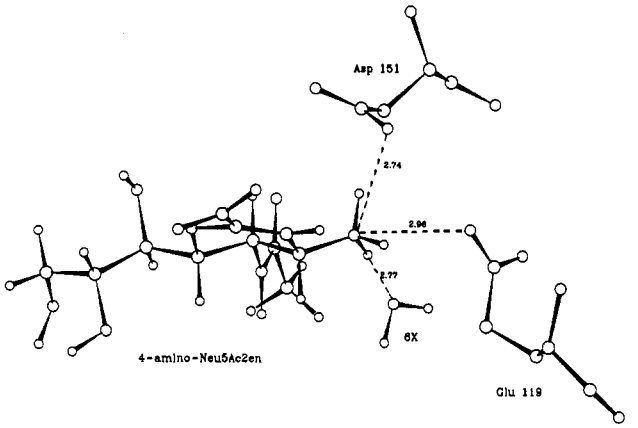


Figure 5. The calculated intermolecular interactions for the sialidase–4-amino-Neu5Ac2en (**2b**) complex in the vicinity of the amino moiety. The hydrogen bonding interactions for the remainder of the inhibitor are the same as in Neu5Ac2en (**2a**) (distances shown in Å).

Table 4. Hydrogen Bonds between Residues in the Active Site of Influenza Virus Sialidase and the Donor Groups 4-Amino (**2b**) and 4-Guanidino (**2c**) of the Substituted Neu5Ac2en Compounds

donors (X)	acceptors (Y)	X···Y (Å)	X–H···Y (deg)
	4-Amino		
N4	Glu 119 Oε2	2.96	131.1
N4	Asp 151 Oδ2	2.74	160.9
N4	H_2O 6X	2.77	162.8
	4-Guanidino		
Nε	Asp 151 Oδ2	2.66	167.8
Nη1	Asp 151 Oδ2	2.93	141.6
Nη1	Glu 119 Oε1	2.70	149.4
Nη2	Glu 227 Oε2	3.54	–
Nη2	H_2O 14X	2.91	161.9
Nη2	Trp 178 O	3.08	159.2

Table 5. Intermolecular Binding Enthalpies (kcal/mol) between Three Inhibitors and Sialidase Residues within the Active Site and Adjacent to C4 on Bound Neu5Ac2en

residue	Neu5Ac2en (**2a**)	4-amino-Neu5Ac2en (**2b**)	4-guanidino-Neu5Ac2en (**2c**)
Glu 119	-12	-22	-23
Asp 151	7	-25	-25
Trp 178	-3	-2	-9
Glu 227	1	-14	-14
Glu 277	8	-7	-8
H_2O 6X	1	-13	
H_2O 14X	-2	-4	-12
total energy	-189	-219	-242

bond to water 6X did not occur in all arrangements; an intramolecular bond to the carbonyl oxygen of the acetamido moiety (N4···O10) was observed in several instances.

Following this rational drug design work, molecular graphics studies on the crystal structure of the sialidase–Neu5Ac complex suggested that the binding pocket accommodating the amino functionality might possibly be able to fit the larger guanidinium ion.[14] An even tighter affinity for the active site was predicted for this substituted species as a result of lateral binding through the terminal nitrogens of the guanidino group with both Glu 119 and Glu 227. The compound 4-guanidino-Neu5Ac2en (**2c**) was synthesized and found to be a more potent inhibitor than compound **2b** (see Table 3).[11,14] Figure 6 illustrates the minimum-energy structure obtained for the complex (again with a positive charge on the group at the 4-position). A strong charge–charge type hydrogen bond between the new group and active-site residue Glu 119 formed, and there is a favorable long-distance interaction with residue Glu 227. As well as these expected interactions, active-site residue Asp 151 was calculated to form a very favorable interaction with the new functionality, as observed in compound **2b**. Water molecule 14X and the carbonyl oxygen of Trp 178 also hydrogen bond to the guanidinyl group. Table 4 lists the distances and angles for all of these bonds. Table 5 shows that the guanidinium moiety forms more favorable interactions with all neighboring active-site residues than both other compounds listed. Net nonbonding energies for the different inhibitors can be observed to agree qualitatively with the K_i values listed in Table 3. The structural water molecule 6X was found to no longer fit in the region, and it is reasonable to speculate that this water is expelled from the active site upon binding of **2c**. This is consistent with the fact that compound **2c** is a slow binding inhibitor;[14] hence significant gains in binding energy are observed upon water being forced out of the binding pocket after the initial binding event. The absence of slow binding between compound **2c** and influenza B sialidase is possibly due to a tighter binding of the equivalent water to 6X in the active site. The fact that the water molecule was clearly resolved in the crystal structure[15] suggests it is tightly bound and cannot be released upon ligand association. In conclusion, a high binding affinity of **2c** to influenza A sialidase is believed to be the result of (i) strong charge–charge hydrogen bonds, (ii) entropy gains from the expulsion of a structural water molecule, and (iii) favorable van der Waals forces between the bulky guanidino group and adjacent enzyme active-site residues.

A crystallographic study of the complex of 4-guanidino-Neu5Ac2en (**2c**) with sialidase[14] confirmed the results from the energy optimization calculation. A close agreement

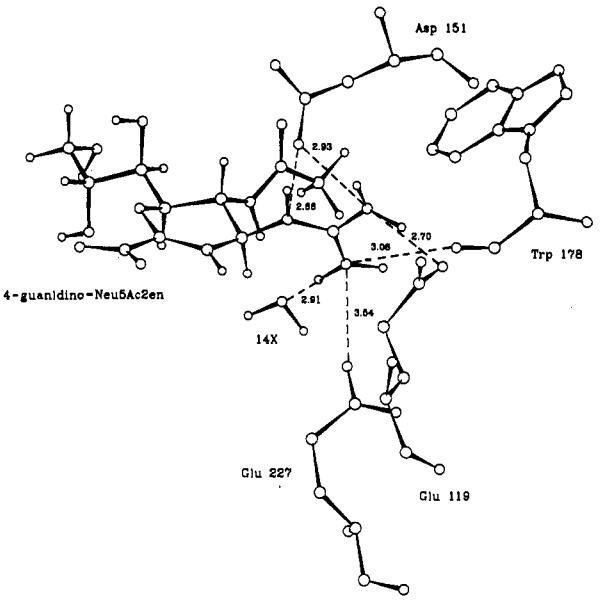


Figure 6. The calculated intermolecular interactions for the sialidase–4-guanidino-Neu5Ac2en (**2c**) complex in the vicinity of the guanidino moiety. The hydrogen bonding interactions for the remainder of the inhibitor are the same as in Neu5Ac2en (**2a**) (distances shown in Å).

between calculated and observed binding modes was found for active-site residues Asp 151 and Trp 178 (from X-ray study: Asp 151 Oδ2···Nε = 3.09 Å and Trp 178 O···Nη2 = 2.77 Å). For residue Glu 119, although the distance to the guanidinium substituent was longer in the crystal structure (Glu 119 Oε1···Nη2 = 3.61 Å), a similar side-on orientation of the carboxylate and guanidinyl moieties was observed. Residue Glu 227 exhibited a slightly worse fit; the hydrogen bonding distance in the crystal structure of the sialidase–inhibitor complex was Glu 227 Oε2···Nη1 = 2.69 Å (cf. Table 4). This difference corresponds to a simple rotation of the carboxylate group in this residue, which, importantly, was successfully modeled in the conformation search experiments (see below). The arrangement of water molecule 14X and residue Glu 277 are highly correlated to the conformation of residue Glu 227 and so, not surprisingly, the hydrogen-bonding pattern for these groups differed in the experimental and calculated structures. The different bonding network, however, corresponded to only small differences in the positions of 14X and Glu 277 in the active site. Overall the calculated structure for the sialidase–4-guanidino-Neu5Ac2en complex closely agreed with experiment.

The strong charge–charge intermolecular hydrogen bonds to residues Glu 119 and Asp 151 were observed in all structures obtained from the molecular dynamics conformational search of the sialidase–4-guanidino-Neu5Ac2en complex. The most important variation in bonding modes occurred with residue Glu 227, as indicated above. In many cases the carboxylate fragment of this active-site residue rotated about 90°, thereby forming a strong hydrogen bond to the guanidinium moiety (for example, Nη2···Glu 227 Oε2 = 2.80 Å). The bond from the carbonyl oxygen of Trp 178 to Nη2 of the guanidinium substituent was not particularly strong and was sometimes replaced by a bond to Nη1. Another weak interaction was Asp 151 Oδ2···Nη1 as suggested by the bond angle listed in Table 4; the bond Asp 151 O···Nη1 sometimes replaced it. Finally, as expected from earlier results, the relatively

high mobility of water molecule 14X results in the hydrogen bond to this water not forming in all minima.

Conclusions

Our molecular modeling studies on ligand binding to the active site of sialidase from influenza virus has considerably furthered our understanding of the specific intermolecular interactions involved. The calculated results agree favorably with the experimentally observed crystal structures. Special attention was required for the crystallographic water molecules in the vicinity of the bound ligands because of the large number of possible orientations for these groups. A conformational search protocol, based on molecular dynamics simulations, was developed to explore the space of intermolecular interaction. The need to examine many minima was found to be vital for obtaining a thorough understanding of the interactions between enzyme and substrate. We have also proposed a detailed enzyme mechanism in which the enzyme plays the roles of (i) activating the sialoside for hydrolysis by inducing ring strain, (ii) facilitating proton donation from solvent, and (iii) providing a favorable electrostatic environment for the charged transition-state intermediates. Our data provide a more complete picture of the molecular interactions of sialic acid-derived substrates and inhibitors with the critical active-site amino acid residues of influenza virus sialidase and should aid in the design of new influenza virus sialidase inhibitors. Finally, the good correlation between the calculated interaction energies for the enzyme–inhibitor complexes and the experimentally determined binding affinities further supports the use of these computational techniques in rational drug design.

Acknowledgment. We thank Glaxo, Australia and Biota Holdings Ltd for financial support.

References

(1) Drzeniek, R. Viral and Bacterial Neuraminidases. *Curr. Top. Microbiol. Immunol.* **1972**, *59*, 35–74.
(2) Schauer, R. Sialic Acids and their Roles as Biological Masks. *Trends Biochem. Sci.* **1985**, *10*, 357–360.
(3) Klenk, H.-D.; Rott, R. Molecular Biology of Influenza Virus Pathogenicity. *Adv. Virus Res.* **1988**, *34*, 247–280.
(4) Palese, P.; Jobita, K.; Ueda, M.; Compans, R. W. Characterization of Temperature Sensitive Influenza Virus Mutants Defective in Neuraminidase. *Virology* **1974**, *61*, 397–410.
(5) Colman, P. M. Neuraminidase: Enzyme and Antigen. In *The Influenza Viruses*; Krug, R. M., Ed.; Plenum: New York, 1989, pp 175–218.
(6) Varghese, J. N.; Laver, W. G.; Colman, P. M. Structure of the Influenza Virus Glycoprotein Antigen Neuraminidase at 2.9 Å Resolution. *Nature* **1983**, *303*, 35–40.
(7) Colman, P. M.; Varghese, J. N.; Laver, W. G. Structure of the Catalytic and Antigenic Sites in Influenza Virus Neuraminidase. *Nature* **1983**, *303*, 41–44.
(8) 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.* **1991**, *221*, 473–486.
(9) Varghese, J. N.; McKimm-Breschkin, J.; Caldwell, J. B.; Kortt, A. A.; Colman, P. M. The Structure of the Complex Between Influenza Virus Neuraminidase and Sialic Acid, the Viral Receptor. *Proteins* **1992**, *14*, 327–332.
(10) Chong, A. K. J.; Pegg, M. S.; von Itzstein, M. Influenza Virus Sialidase: Effect of Calcium on Steady-State Kinetic Parameters. *Biochim. Biophys. Acta* **1991**, *1077*, 65–71.
(11) Holzer, C. T.; von Itzstein, M.; Jin, B.; Pegg, M. S.; Stewart, W. P.; Wu, W.-Y. 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. *Glycoconjug. J.* **1993**, *10*, 40–44.
(12) Chong, A. K. J.; Pegg, M. S.; von Itzstein, M. Characterization of an Ionisable Group Involved in Binding and Catalysis by Sialidase from Influenza Virus. *Biochem. Int.* **1991**, *24*, 165–171.
(13) Chong, A. K. J.; Pegg, M. S.; Taylor, N. R.; von Itzstein, M. Evidence for a Sialosyl Cation Transition-State Complex in the Reaction of Sialidase from Influenza Virus. *Eur. J. Biochem.* **1992**, *207*, 335–343.
(14) von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. Rational Design of Potent Sialidase-Based Inhibitors of Influenza Virus Replication. *Nature* **1993**, *363*, 418–423.
(15) Burmeister, W. P.; Ruigrok, R. W. H.; Cusack, S. The 2.2 Å Resolution Crystal Structure of Influenza B Neuraminidase and its Complex With Sialic Acid. *EMBO J.* **1991**, *11*, 49–56.
(16) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. Structure and Energetics of Ligand Binding to Proteins: Escherichia coli Dihydrofolate Reductase-Trimethoprim, A Drug-Receptor System. *Proteins* **1988**, *4*, 31–47.
(17) Solmajer, T.; Mehler, E. L. Modelling Solvent Effects in Molecular Dynamics Simulations of Proteins. *Int. J. Quantum Chem.* **1992**, *44*, 291–299.
(18) Phillips, D. C. The Three-Dimensional Structure of an Enzyme Molecule. *Sci. Am.* **1966**, *215*, 78–90.
(19) Chong, A. K. J. Influenza Virus Sialidase: A Mechanistic Study. Masters Thesis, Monash University, 1990.
(20) Meyer, E. Internal Water Molecules and H-Bonding in Biological Macromolecules: A Review of Structural Features with Functional Implications. *Prot. Sci.* **1992**, *1*, 1543–1562.
(21) Burmeister, W. P.; Henrissat, B.; Bosso, C.; Cusack, S.; Ruigrok, R. W. H. Influenza B Virus Neuraminidase Can Synthesise its own Inhibitor. *Structure* **1993**, *1*, 19–26.
(22) 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* **1987**, *26*, 5351–5358.
(23) Goodford, P. A. Computational Procedure for Determining Energetically Favourable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
(24) von Itzstein, M.; Dyason, J. C.; White, H. F.; Oliver, S. W. Unpublished results.