A Study of the Active Site of Influenza Virus Sialidase: An Approach to the Rational Design of Novel Anti-influenza Drugs

Mark von Itzstein,* Jeffrey C. Dyason, Stuart W. Oliver, Hume F. White, Wen-Yang Wu, Gaik B. Kok, and Michael S. Pegg

Department of Medicinal Chemistry, Victorian College of Pharmacy, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia

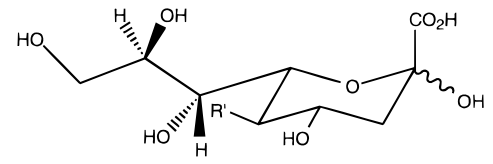
Received April 20, 1995[⊗]

The development of sialidase inhibitor-based potential anti-influenza drugs using rational drug design techniques has been of recent interest. The present study details an investigation of the active site of influenza virus sialidase by using the program GRID in an attempt to design more potent inhibitors in the hope they will eventually lead to anti-influenza drugs. A number of different probes (amino, carboxy, hydroxy, methyl, etc.) have been used in an effort to determine the functional groups most likely to improve the binding of the starting template 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en). The data have correctly predicted the binding regions for the carboxylate, acetamido (NH and methyl), and glycerol (OH) groups of *N*-acetylneuraminic acid. Moreover, the data suggest that the addition of certain functionalities (amino group) at the C-4 position should enhance the overall binding.

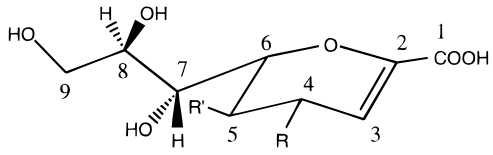
Introduction

Sialidase (EC 3.2.1.18) from influenza virus is an exoglycohydrolase which catalyses the hydrolysis of α-ketosidically-linked terminal sialic acids from glycoproteins, glycolipids, and a variety of oligosaccharide substrates.[1] Our interest in the biochemistry of sialidases[2,3] has led us to propose a mechanism of action for influenza virus sialidase which involves a sialosyl cation intermediate.[4] It has been suggested that among its putative roles sialidase is required for the elution of viral progeny away from the infected cell.[5] The structure of the enzyme has been crystallographically determined from three influenza subtypes, namely A/Tokyo,[6] A/tern,[7] and more recently B/Beijing.[8] Importantly, Varghese *et al.*[9] have very recently described the structure of complexes between influenza virus sialidase and *N*-acetylneuraminic acid (Neu5Ac, **1**) and its analogues. These studies, taken together with the availability of sequence data of sialidases from a large number of antigenic variants,[10] provide the basis for a computer-assisted analysis of the influenza virus sialidase active site with commercially available software.

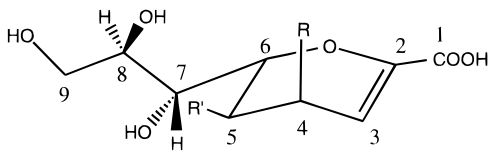
The study of interactions between a series of functional groups (referred to as probes) and a biomolecule has been facilitated by the well-known program GRID.[11] Both electrostatic and steric interactions are taken into account by the program, as it builds up a contour map of the interaction energy versus 3-D position of the probe with respect to the crystal structure of the biomolecule. This information can lead to the prediction of how various functional groups of a ligand will interact in a specific region within the active site of an enzyme. We describe here further studies in the use of the program GRID to analyze the active site of the influenza virus sialidase A/Tokyo.[6] This methodology is verified by comparing the predicted binding site of Neu5Ac with the experimentally determined site as described by Varghese *et al.*[9] Furthermore, we report our efforts to design novel sialic acid analogues based on both the GRID analysis and an eyeball analysis of bound Neu5Ac (**1**) and its analogue, 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en, **2**). This work has led to the synthesis of the most potent influenza virus sialidase inhibitor to be reported,[12,13] as well as providing an insight into the interaction between Neu5Ac and its analogues with influenza virus sialidase.[14]



1 R' = NHAc



2 R = OH, R' = NHAc
3 R = NH_2, R' = NHAc
4 R = NH(C=NH)NH_2, R' = NHAc



5 R = NH_2, R' = NHAc

Methods

1. Crystal Structures. The crystal structures of A/Tokyo/3/67 sialidase with and without Neu5Ac (**1**) and Neu5Ac2en (**2**) bound were kindly supplied by Peter Colman and Jose Varghese. The resolution of the native sialidase structure is 2.2 Å. The residual was 0.21 for data between 6 and 2.2 Å. The resolution of sialidase−Neu5Ac and sialidase−Neu5Ac2en complex was refined to 1.8 and 2.9 Å, respectively. Further information of these structures is detailed by Varghese *et al.*[9]

* Address all correspondence to Dr. Mark von Itzstein, Department of Medicinal Chemistry, Monash University, 381 Royal Parade, Parkville, Victoria, Australia, 3052. Tel: 61-3-9903-9542. Fax: 61-3-9387-9387. E-mail: mark.vonitzstein@vcp.monash.edu.au.

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1995.

0022-2623/96/1839-0388$12.00/0

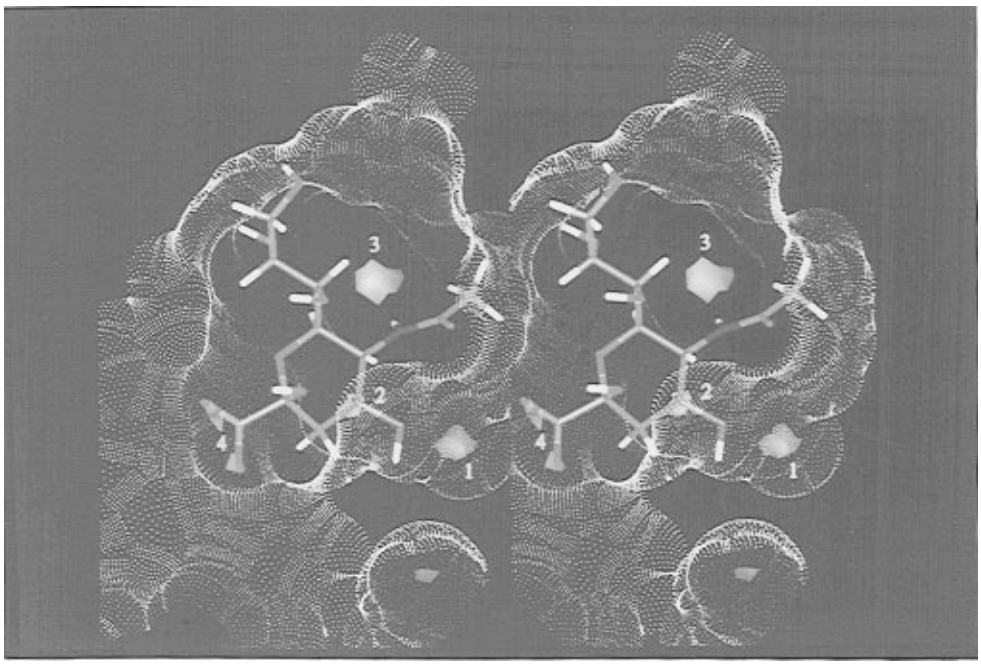


Figure 1. Relaxed stereoview showing the three main regions of interaction between the amino probe (marked 1, 2, and 3 and contoured at an energy of −16 kcal), the carboxy probe (4, contoured at −9.5 kcal), and the active site of influenza virus sialidase.

2. Computing. Initial calculations were performed on a VAX 11/750 using GRID version 4. Subsequent studies were carried out on a variety of Silicon Graphics work stations using GRID version 11. The results and the crystal structure were viewed using Insight II (Biosym Technologies, Inc., San Diego, CA). The GRID calculations presented in this paper were run using version 11 of GRID and the crystal structure of sialidase as mentioned above. The calculations were performed on a cube 35 Å per side, centered on the active site, with a grid spacing of 0.5 Å. The interaction energy between the probe and every atom within the protein structure was evaluated at each grid point. A dielectric constant of 80 was used to simulate a bulk aqueous phase, while areas as determined by GRID to be excluded from solvent were assigned a dielectric constant of 4 (i.e. the interior of the protein). The accompanying program GRIN was used to automatically assign atom types and charges for the protein, using the standard parameter file provided with GRID. The output was converted (using GINS, supplied with GRID) into a form suitable for input to the Biosym utility Contour, and contour maps were built up using steps of 2 kcal/mol. The contour map was then viewed superimposed on the crystal structure of the active site of N2 using Insight II. We have found that the best way of observing the results of GRID was to use the contour map feature of Insight II to overlay the interaction contour map generated by GRID with the crystal structure of the enzyme, which also had present a solvent accessible surface for the active site residues. Further elucidation of the interactions was facilitated by the superimposition of the inhibitor Neu5Ac (**1**), as determined from the crystal structure, onto the active site picture described above. This provides valuable insight into the interactions between the enzyme and substrate-based inhibitors in terms of the inhibitor geometry.

3. Synthesis and Biological Assays. Compounds **2**−**5** have been prepared from *N*-acetylneuraminic acid, and these syntheses are well documented.[12,13] The biological assay used to determine the inhibition constants of the prepared compounds is the well-known fluorometric sialidase assay and is detailed elsewhere.[15] The substrate used for the enzyme assay is 2′-(4-methylumbelliferyl)-*N*-acetylneuraminic acid and is prepared using published procedures.[4]

Results and Discussion

GRID Analysis. Several probes were used to study the active site of influenza virus sialidase in an effort to identify the main areas of interaction between the enzyme and various ligands. Summarized below are the most important enzyme−ligand interaction sites identified by using GRID.

A. *O*-Carboxylate. There was only one significant region of interaction for the carboxylate probe (−9 kcal/mol) in the active site (Figure 1, region 4), and this corresponded very well with the position of the carboxylate group of Neu5Ac (**1**) and highlighted the importance of the presence of the triarginal cluster (Arg 118, Arg 292, Arg 371) and especially Arg 371.

B. NH_3 Cation. The cation form of the amino probe was chosen because a previous pH profile[4] suggested that at the optimum pH for enzyme activity (pH 5.5) an amine group on an inhibitor would be protonated. The regions of greatest interaction proved, perhaps not so surprisingly, to be at sites adjacent to negatively charged amino acids within the 35 Å square box that was studied. Most of these could be ignored as they were buried deep within the protein and were not accessible to inhibitors. The GRID analysis calculated the interaction energies for these sites to start at −26 kcal/mol, with accessible regions in the active site appearing at −18 kcal/mol. There were three main regions of interaction predicted in the active site at this energy and these were (1) adjacent to the 4-OH position of Neu5Ac, (2) beneath the C4 position of the sugar ring, and (3) on the floor of the active site between the glycerol side chain and the 5-*N*-acetyl group. These major sites of interaction are shown in stereo in Figure 1, with the same numbering scheme as mentioned above.

The main residues involved in the above interactions are shown in Table 1 along with some distance information.

C. OH (Hydroxy). This probe showed a region of interaction in the active site at −8 kcal/mol, corresponding to the position of the carboxylate group of Neu5Ac (**1**) (Figure 1, region 4) and its interaction with the triarginal cluster in particular with Arg 371. Since the predicted interaction energy of the hydroxyl group is lower than that of the carboxyl group, it would not be expected to provide an increased binding affinity. At a

Table 1. Summary of the Main Probe Interactions Observed in This Study, Showing the Relevant Residues Involved in a Particular Interaction and Appropriate Distance Information

probe	region[a]	max. energy (kcal)	main residues	distance (Å)[b]
NH_3 cation	1	−16	Glu 119	3.5 (CD)
			Asp 151	4.5 (CG)
			Trp 151	2.9 (O)
			Glu 227	5.1 (CD)
	2	−18	Glu 119	3.7 (CD)
			Glu 227	3.8 (CD)
			Glu 277	4.2 (CD)
			Tyr 406	3.4 (OH)
	3	−18	Ser 179	3.5 (OG)
			Thr 225	2.9 (O)
			Glu 227	3.5 (CD)
			Glu 277	4.0 (CD)
carboxy	4	−9.5	Arg 118	5.6, 3.9 (CZ)
			Arg 292	3.5, 5.7 (CZ)
			Arg 371	3.8, 3.9 (CZ)
methyl	NS[c]	−4	Trp 178	3.6 (CE3)
			Ile 222	4.3 (CG2)
hydroxy	NS	−8	same as carboxy	
phosphate	NS	−9	same as carboxy	

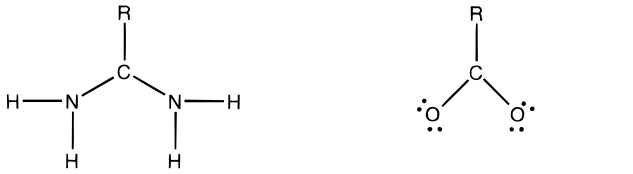
[a] With reference to Figure 1. [b] Atom names in parentheses, according to PDB format. [c] NS = not shown in Figure 1.

contour level of −6 kcal/mol the interaction area involving Arg 371 had grown markedly and there were small areas of interaction showing up in positions corresponding to the glycerol side chain of the template inhibitor. Unexpectedly, there was no interaction at this level predicted for the 4-OH of Neu5Ac (**1**).

D. CH_3 (Methyl). In an effort to study not only the stronger electrostatic interactions but also to try and identify areas of hydrophobic interaction, several of the available hydrophobic probes were tested on the active site. The methyl probe proved to be the most valuable of these hydrophobic probes which showed an area of interaction at a contour level of −4 kcal/mol, centered on a pocket adjacent to the methyl of the 5-*N*-acetyl group of Neu5Ac. This pocket is predominantly due to Trp 178.

E. Other Probes. The other probes provided with GRID were also run and it was found that, in general, oxygen-based probes (for example, O_phenolate) showed good interaction only with the previously mentioned triarginal cluster. None of the oxygen-based probes showed any significant interaction in the region of space occupied by the O6 ring oxygen. The nitrogen-based probes (for example, nh2_amide) showed good interaction at sites similar to those described for the NH_3 cation. An interesting result that suggested some promise was from the use of the phosphate PO_4^{2-} probe. This result was found to be similar to the carboxylate probe in both its region and its energy of interaction. However, the volume of the interaction is larger for the PO_4^{2-} and may provide more latitude for positioning of other templates within this active site.

Multiatom Probes. In later versions of GRID the capability of dealing with multiatom probes has been included. These probes extend the functionality already seen in the "single-atom" probes. For example, the previously mentioned carboxylate probe, although representing a three atom functional group, is actually calculated as if it were a single atom positioned at the grid point. Whereas, the newer multiatom probes place one atom of the probe at the grid point and then perform the energy calculation for all atoms within the probe and assigns that total probe energy to the grid point.[16] The probe is then rotated so that each possible hydrogen-bonding orientation is considered and the most favorable interaction energy is accepted.



Amidine Multi-Atom Probe Carboxylate Multi-Atom Probe

Figure 2. Diagrammatic representation of the two multiatom probes of interest, where R is either an aromatic or aliphatic group.

The most interesting of the multiatom probes investigated were the carboxylate probe and the amidine probe, both attached to either aliphatic or aromatic groups (Figure 2).

The carboxylate multi-atom probe has a charge of −1 and four lone pairs which can accept up to four hydrogen bonds, while the amidine multiatom probe has a charge of +1 and the ability to donate up to four hydrogen bonds. After the preliminary use of the single atom probes, these two multiatom probes were investigated as perhaps better models of the two best types of interaction already observed.

A. Carboxylate Multiatom Probe. Both the aliphatic and aromatic versions of this probe showed similar profiles with a very strong interaction surrounding the region of space where the carboxylate group of Neu5Ac (**1**) resides. These probes predicted a much more intense interaction than that predicted by the single atom probes. Energies of −20 kcal/mol for the aliphatic probe and −17 kcal/mol for the aromatic probe were predicted. Once again, the site of the carboxylate group of Neu5Ac (**1**) was the only one to show up in the active site (Figure 1, region 4).

B. Amidine Multiatom Probe. Both the aliphatic and aromatic versions of this probe showed similar profiles with three main areas of interaction in the active site of influenza virus sialidase. With the aliphatic probe, regions 1 and 3 (Figure 1) were visible at −27 kcal/mol while region 2 (Figure 1) became visible at −23 kcal/mol. The aromatic probe showed region 3 at −27 kcal/mol with regions 1 and 2 showing at −23 kcal/mol. Once again these three regions were the predominant regions in the active site with the interaction volume growing in size as the contouring energy was lowered. This, combined with electrostatic data, clearly shows that the whole floor of the active site is actually quite negatively charged, leaving only the portion around the triarginal cluster being positively charged.

Design. It was clear from the GRID data that replacement of the 4-hydroxyl group by an amino group on the Neu5Ac2en (**2**) template to give 4-amino-Neu5Ac2en (**3**) should produce a substantial increase in affinity of this inhibitor for influenza virus sialidase. Moreover, further modeling studies of other 4-substituted sialic acids indicated that substitution of the 4-hydroxyl group by a more basic 4-guanidinyl group on the parent Neu5Ac2en (**2**) template should even further enhance the binding affinity of this sialic acid analogue. Basic modeling studies were performed by first building and minimizing 4-guanidino-Neu5Ac2en (**4**) and then locating this molecule in the sialidase active

Table 2. List of the Measured Inhibition Constants for Various Inhibitors Using 2′-(4-Methylumbelliferyl)-*N*-acetylneuraminic Acid as the Substrate

compound	K_i (M)
Neu5Ac (**1**)	10^{-3}
Neu5Ac2en (**2**)	4×10^{-6}
4-amino-Neu5Ac2en (**3**)	4×10^{-8}
4-guanidino-Neu5Ac2en (**4**)	3×10^{-11}
4-*epi*-amino-Neu5Ac2en (**5**)	3×10^{-7}

site in a similar position to where Neu5Ac2en (**2**) binds. This showed that there was sufficient room, albeit tight, for the more bulky guanidinyl group to fit into the pocket adjacent to the 4-OH group of the original Neu5Ac (**1**) template. Subsequent work[17] has also shown that 4-guanidino-Neu5Ac2en (**4**) is a slow-binding inhibitor, which may be the result of the expulsion of a water molecule from the C4 binding pocket of the active site, giving rise to an increase in the binding interaction.

The synthesis of **3** and **4** has been achieved[13] from a 4-azido intermediate[18] in reasonable yield. The desired and expected increase in binding affinity and selectivity for influenza virus sialidase (Table 2) is remarkable and can be explained in terms of differences between the C4 binding pocket of sialidases from varying sources.[15]

In order to exploit the predicted favorable interaction with region 2 (Figure 1), the obvious compound to try was the 4-*epi*-amino-Neu5Ac2en (**5**). This compound was synthesized and tested[19] and demonstrated better inhibition (Table 2) than Neu5Ac2en (**2**), but was found not to be as good as 4-amino-Neu5Ac2en (**3**). Modeling studies of this compound showed that the bulky amino group beneath the sugar ring could only just fit into the binding site without causing too much distortion. The closeness of the fit and the relatively small predicted interaction volume preclude any larger substituted compounds from being of use for targeting this region.

The third region of favorable interaction for an amino group (Figure 1) is removed from the inhibitor and situated in a hollow on the floor of the active site. In the crystal structure there are two molecules of water situated in this hollow, and there appears to be an interaction with the amide group attached to the C5 position of Neu5Ac (**1**). Better targeting of this region should increase the interaction between the protein and a ligand and also, the displacement of the crystallographic waters could considerably increase the binding affinity of the ligand.

Conclusion

By making use of the computer program GRID and the crystal structure of the enzyme influenza virus sialidase, we have been able to identify several key regions of the active site that appear to be necessary to the binding of the natural substrate and also several other regions which are of interest in the design of inhibitors for this enzyme. These data correctly predict the binding regions for the carboxylate, acetamido (NH and methyl), and glycerol (OH) groups of Neu5Ac (**1**).

Moreover, the data suggests that the introduction of an amino functionality at the C4 position should enhance the overall binding. Replacement of the carboxylate with a phosphonate group is also predicted to enhance binding. This final option has recently appeared in the literature,[20] where the phosphonate analogue of compound **1** has shown between 1 and 2 orders of magnitude improvement in K_i over compound **1**, for sialidases from two different influenza virus serotypes.

Acknowledgment. We thank the Australian Government, through an Industry Research and Development Board generic grant and the National Health & Medical Research Council, and Biota Holdings Pty Ltd and Glaxo, Australia, for their financial support. We would also like to thank one of the reviewers for their helpful comments.

References

(1) Drzeniek, R. Viral and Bacterial Neuraminidases. *Curr. Top. Microbiol. Immunol.* **1972**, *59*, 35−74.
(2) 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.
(3) 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.
(4) 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.
(5) Palese, P.; Jobita, K.; Ueda, M.; Compans, R. W. Characterization of Temperature Sensitive Influenza Virus Mutants Defective in Neuraminidase. *Virology* **1974**, *61*, 397−410.
(6) Varghese, J. H.; 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.
(7) Bossart-Whitaker, P.; Carson, M.; Babu, Y. S.; Smith, C. D.; Laver, W. G.; Air, G. M. Three-dimensional Structure of Influenza A N9 Neuraminidase and Its Complex with the Inhibitor 2-Deoxy-2,3-Dehydro-*N*-Acetyl Neuraminic Acid. *J. Mol. Biol.* **1993**, *232*, 1069−1083.
(8) 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.
(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*, 327−32.
(10) Colman, P. M. Neuraminidase: Enzyme and Antigen. In *The Influenza Viruses*; Krug, R. M., Ed.; Plenum: New York, 1989; pp 175−218.
(11) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.
(12) von Itzstein, M.; Wu, W. Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M.; 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.
(13) von Itzstein, M.; Wu, W. Y.; Jin, B. The Synthesis of 2,3-didehydro-2,4-dideoxy-4-guanidinyl-*N*-acetylneuraminic acid: a potent influenza virus sialidase inhibitor. *Carbohydr. Res.* **1994**, *259*, 301−305.
(14) Taylor, N. R.; von Itzstein, M. Molecular Modeling Studies on Ligand Binding to Sialidase from Influenza Virus and the Mechanism of Catalysis. *J. Med. Chem.* **1994**, *37*, 616−624.
(15) 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. *Glycoconjugate J.* **1993**, *10*, 40−44.
(16) Goodford, P. J. Personal communication.
(17) Pegg, M. S.; von Itzstein, M. Slow-Binding Inhibition of Sialidase from Influenza Virus. *Biochem. Mol. Biol. Int.* **1994**, *32*, 851−858.
(18) von Itzstein, M.; Jin, B.; Wu, W. Y.; Chandler, M. A convenient method for the introduction of nitrogen and sulfur at C-4 on a sialic acid analogue. *Carbohydr. Res.* **1993**, *244*, 181−185.
(19) von Itzstein, M.; Wu, W. Y.; Van Phan, T.; Danylec, B.; Jin, B. International Patent Appl. WO9116320-A92.10.31 (9147), 1991; *Chem. Abstr.* **1991**, *117*, 49151y.
(20) White, C. L.; Janakiraman, M. N.; Laver, W. G.; Philippon, C.; Vasella, A.; Air, G. M.; Luo, M. A Sialic Acid-derived Phosphonate analog inhibits different strains of Influenza virus Neuraminidase with different efficiencies. *J. Mol. Biol.* **1995**, *245*, 623−634.

JM950294C