Optimal Charges in Lead Progression: A Structure-Based Neuraminidase Case Study

Kathryn A. Armstrong,[†,‡] Bruce Tidor,*[,†,‡,§] and Alan C. Cheng*[,||]

Biological Engineering Division, Computer Science and Artificial Intelligence Laboratory, and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Research Technology Center, Pfizer Global Research & Development, Cambridge, Massachusetts 02139

Collective experience in structure-based lead progression has found electrostatic interactions to be more difficult to optimize than shape-based ones. A major reason for this is that the net electrostatic contribution observed includes a significant nonintuitive desolvation component in addition to the more intuitive intermolecular interaction component. To investigate whether knowledge of the ligand optimal charge distribution can facilitate more intuitive design of electrostatic interactions, we took a series of small-molecule influenza neuraminidase inhibitors with known protein cocrystal structures and calculated the difference between the optimal and actual charge distributions. This difference from the electrostatic optimum correlates with the calculated electrostatic contribution to binding ($r^2 = 0.94$) despite small changes in binding modes caused by chemical substitutions, suggesting that the optimal charge distribution is a useful design goal. Furthermore, detailed suggestions for chemical modification generated by this approach are in many cases consistent with observed improvements in binding affinity, and the method appears to be useful despite discrete chemical constraints. Taken together, these results suggest that charge optimization is useful in facilitating generation of compound ideas in lead optimization. Our results also provide insight into design of neuraminidase inhibitors.

Introduction

Ligand−receptor affinity is affected by shape and charge complementarity as well as other factors including strain, entropy, titration effects, and conformational change. Collective structure-based drug design (SBDD) experience has found that molecular visualization and chemical intuition are often sufficient for successful optimization of shape-based hydrophobic and van der Waals interactions, while charge-based hydrogen-bond and salt-bridge interactions are more difficult to optimize in the same fashion. Davis and Teague[1] conclude from a large number of SBDD case studies that optimization of hydrogen bonds and salt bridges is unpredictable, and they advise medicinal chemists to focus on optimizing the more predictable hydrophobic interactions. They further point out that optimization of electrostatic interactions may be more of an issue with the advent of high-throughput screens, which preferentially select lipophillic compounds that can require addition of polar surface area for bioavailability. Thus, there is substantial need for approaches for discovering electrostatic enhancements to binding affinity.

The nonintuitive nature of electrostatics in ligand optimization is largely due to the net electrostatics being composed of ligand and receptor desolvation in addition to the intermolecular charge−charge interaction that can be "seen" in molecular visualization of the bound complex. While desolvation is difficult to visualize, being a difference in properties in the unbound and bound states, it can be calculated using continuum electrostatic methods such as those that solve the Poisson−

Boltzmann equation.[2−6] Reasonably accurate affinity estimates with inclusion of desolvation can be calculated using these methods, despite the positional sensitivity noted in some docking studies.[7] In applying these methods, a library of possible analogues might be designed around a cocrystallized lead and then prioritized computationally before synthesis of a subset, thereby decreasing the number of compounds needing to be made in lead optimization. However, this is an iterative trial-and-error approach, and ideally one wants to address charge optimization in the same fashion as shape-based optimization, where one can see the ligand and binding pocket shapes and design changes with moderate success.

Quantitative charge optimization techniques[8,9] provide an approach to facilitate idea generation by computing the optimal charge distribution for a given scaffold taking into account the ligand and receptor desolvation as well as the intermolecular electrostatic interaction. The method has been used to design chemical modifications to improve charge interactions.[10−12] Designing synthetically feasible molecules that conform to these optimal charges can be difficult, because discrete chemical modifications will rarely produce the exact desired charges,[12] and chemical modifications can result in shifts in the ligand binding mode as well as ligand and receptor conformations. The designed improvement in electrostatics must compete with possibly unfavorable changes to the other energetic components, which can result in unpredictable tradeoffs between electrostatics, van der Waals, entropy, and conformational change as compared to the starting co-complex structure used in the design analysis.

Here, we examine a series of known small molecule neuraminidase inhibitors with available co-complex structures in a simulated lead-optimization study to understand the behavior and utility of charge optimization in a drug discovery setting. We ask whether adjusting a molecular charge distribution such that it more closely matches the calculated optimum

* Corresponding authors. B.T.: Phone, 1-617-253-7258; Fax, 1-617-252-1816; E-mail, tidor@mit.edu. A.C.C.: Phone, 1-617-551-3166; Fax, 1-617-551-3178; E-mail, alan.c.cheng@pfizer.com.
   † Biological Engineering Division, Massachusetts Institute of Technology.
   ‡ Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
   § Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.
   || Pfizer Global Research & Development.

improves the electrostatic portion of the binding affinity and whether these changes translate to improvements in the experimental binding affinity. This is a nontrivial correspondence due to many factors, including the discrete constraints of chemistry and minor shifts in the ligand binding mode and receptor conformation.

Neuraminidase is an important antiviral target for influenza, with efforts to develop inhibitors starting in the late 1960s. It was only with the elucidation of the protein crystal structure in 1983[13] and SBDD efforts over the following decade that inhibitors with therapeutically efficacious potencies were first developed. These efforts led to the marketing of the first approved drugs, oseltamivir and zanamivir, in 1998. We investigate the utility of charge-optimization within the context of neuraminidase and present analyses that may aid further optimization and design of new inhibitors.

Methods

Structure Preparation. N9-subtype neuraminidase co-complex structures were identified through a BLAST search of proteins in the Protein Data Bank (PDB).[14] All selected structures are atomic-resolution structures with crystallographic resolutions in the 1.6–2.1 Å range. Eleven receptor–ligand complex structures were used as starting points for the neuraminidase calculations (PDB identifiers 1BJI, 1F8B, 1F8C, 1F8D, 1F8E, 1L7F, 1MWE, 1NNC, 2QWI, 2QWJ, 2QWK).[15–20]

Cocrystal structures were prepared using Pfizer in-house software that automates standard protein preparation tasks. Manual preparation or publicly available tools can also be used. We specify here our protein preparation procedure. Ends of protein regions with missing density were capped with either an N-terminal acetyl group or C-terminal amide group, and incomplete residues were automatically built using rotamer replacement. All cofactors were removed except the buried calcium ion that is about 11 Å from the binding site, which was included and assigned a charge state of $+2e$. Hydrogens were added to the molecules, titrated at neutral pH, and rotationally optimized in the presence of the cocrystallized ligand partner. Histidines were assigned the most favorable protonation and flip states, and glutamines and asparagines were assigned their most favorable flip states. Water molecules in the structure exposing more than 1 Å^2 of surface area to other waters or to solvent accessible regions were removed, and the remaining crystallographic water molecules were each rotated spherically about the oxygen atom to optimize hydrogen-bonding networks.

The charges and radii of all receptor atoms were assigned according to the CHARMM PARAM19 parameter set and all ligands were assigned CHARMM PARAM19 radii.[21] As a test of parameter sensitivity, parallel calculations were run with all charges and radii reassigned using the PARSE parameter set.[22] Reference charges for all ligands were determined using Gaussian 03[23] and the RESP procedure.[24,25] Ligands were first geometry-optimized in the gas phase using the Hartree–Fock method with the 6-31G* basis set, and then their electrostatic potentials were calculated using the same level of theory. Partial atomic charges for all atom centers were then calculated using a two-stage RESP fit to these potentials. The overall charge used for each compound was zero, except for compounds **1** and **4**, which were each assigned $-1e$ overall charge, and **3**, which was assigned a $+1e$ overall charge.[16]

General Electrostatic Calculations. The continuum electrostatic approximation was used to calculate electrostatic potentials. Calculations were made with a locally developed Poisson–Boltzmann solver (M. D. Altman and B.T., unpublished) that uses the finite-difference approach to solve the linearized Poisson–Boltzmann equation[26,27] with a multigrid implementation.[28,29] These potentials were calculated using a solvent dielectric of 80 with a bulk ionic strength of 0.145 M, a solute dielectric of 4, and a Stern layer of 2.0 Å.[30] A molecular surface generated with a 1.4-Å radius

probe defined the dielectric boundary. The averaged results of 10 translations on a 129 × 129 × 129 cubic grid were used in all cases.

Charge Optimization. The optimal charge distribution for each ligand was calculated by optimizing a variable point charge at each atom center of the ligand. Charge-optimization theory, described in detail elsewhere,[8–11,31–33] enables calculation of the optimal charge distribution for a given ligand bound to a receptor. This distribution lies at the minimum electrostatic free energy of binding for those atom centers, which is computed by adding the unfavorable desolvation penalties of both the ligand and receptor and the favorable intermolecular interactions between the molecules. The vector of optimal charges is a function of the receptor charges, $\mathbf{Q}_R$, the matrix of ligand–receptor interactions, $\mathbf{C}$, and the ligand desolvation matrix, $\mathbf{L}$:

$$\mathbf{Q}_L{}^{opt} = -\frac{1}{2}\mathbf{L}^{-1}\mathbf{C}^T\mathbf{Q}_R.$$

The individual partial atomic charges computed for each ligand were constrained using LOQO[34] to be no more than 1.0e in magnitude, except in the lead progression analysis, where no atomic charge constraints were used. Additionally, the total charge on each ligand was constrained to be no more than 2.0e in magnitude, except in the lead progression analysis, where no total charge constraints were used. These constraints limit the optimization to a reasonable charge space for medicinal chemistry. This theory has been applied to several ligand–receptor systems, such as barnase–barstar,[10] chorismate mutase,[11] and granulocyte colony stimulating factor,[35] as well as protein kinase A,[12] CDK2,[12] CDK5,[36] cathepsin B,[37] PSD-95 PDZ domain,[38] bovine trypsin,[38] Gln-tRNA synthetase,[39] and HIV-1 gp41 ectodomain.[40]

Results and Discussion

For the set of N9-subtype neuraminidase inhibitors with known cocrystal structures, we calculated both the actual ligand atomic charges and the theoretically optimal ligand atomic charges that would lead to the most favorable electrostatic binding free energy for each ligand. Comparison of the reference and optimal charges suggests nonoptimal ligand atoms and functional groups whose modification or substitution may lead to improved potency. One way to represent the neuraminidase inhibitor binding pocket is to divide it into five regions[41] based on common inhibitor functional group interactions, as shown in Figure 1A. The carboxylate at position 2 makes salt bridges to three arginines (Arg118, Arg371, Arg292) in the S1 pocket and is conserved in all known inhibitors. The acetamide at position 5 making interactions in the S3 pocket is also conserved in most inhibitors, while the substituent at position 4 occupying the S2 pocket and the glycerol at position 6 occupying the S4 and S5 pockets can be substantially optimized, as can be seen in Figure 2.

Electrostatic Distance from the Optimum. For each of the 11 ligands shown in Figure 2, we find that the electrostatic component of the computed binding free energy improves with chemical modifications that bring the charge distribution of each molecule closer to its optimum. This distance from the optimum, denoted $\Delta\Delta G_{elec}^{opt}$, is defined as the difference between ΔG_{elec}^{ref}, the electrostatic component of the free energy of binding the ligand with reference charges, and ΔG_{elec}^{opt}, the electrostatic component of the free energy of binding the ligand with optimal charges. Table 1 gives these values for each of the ligands. Experimentally measured binding affinities for the series of small molecule inhibitors were also collected from the literature[15,16,20,41,42] to facilitate comparisons and aid understanding of the tradeoffs being made in each active site between the electrostatics and other contributors to binding.

Table 1. Computed Optimal and Reference Electrostatic Components for Neuraminidase−Ligand Binding[a]

PDB ID	D_L^{ref}	D_L^{opt}	D_R	I^{ref}	I^{opt}	ΔG_{elec}^{ref}	ΔG_{elec}^{opt}	$\Delta\Delta G_{opt}$	ΔG_{expt}
1MWE	12.4	27.9	27.7	−19.0	−59.5	21.1	−4.0	−25.1	−4.1
1F8B	12.0	22.2	26.9	−15.1	−48.9	23.7	0.2	−23.5	−7.3
1F8C	16.8	24.9	27.5	−32.4	−54.5	11.9	−2.1	−14.0	−10.1
1F8D	14.2	23.8	26.4	−23.0	−50.9	17.5	−0.7	−18.2	−4.6
1F8E	25.0	25.7	27.5	−42.1	−55.4	10.3	−2.2	−12.5	−6.6
1BJI	15.8	24.7	28.3	−35.4	−53.3	8.6	−0.4	−9.0	−10.7
1L7F	16.3	29.0	28.1	−35.4	−61.6	9.1	−4.5	−13.6	−12.2
1NNC	19.5	29.7	30.8	−35.1	−63.3	15.2	−2.8	−18.0	−11.8
2QWI	17.6	25.9	27.3	−36.7	−55.9	8.1	−2.7	−10.8	−11.4
2QWJ	15.3	22.8	26.6	−35.1	−49.7	6.8	−0.3	−7.0	−9.0
2QWK	14.6	26.4	28.2	−37.0	−56.7	5.7	−2.2	−8.0	−11.8

[a] Energetics (kcal/mol) computed using the CHARMM PARAM19 parameter set. All experimental values contain some reported error as well as inaccuracy related to conversions using $\Delta G = -RT \ln(K_i \text{ or } IC_{50})$. D_L^{ref}, D_L^{opt}, and D_R are the calculated desolvation of the reference ("wild-type") ligand, the optimal ligand, and the receptor, respectively. I^{ref} and I^{opt} are the calculated electrostatic component of the interaction energy of the reference and optimal ligands with each receptor. ΔG_{elec}^{ref} and ΔG_{elec}^{opt} are the calculated electrostatic free energy of binding the reference and optimal ligands. $\Delta\Delta G_{opt}$ is the distance from the optimum as described earlier, and ΔG_{expt} is the experimental binding affinity, converted into kcal/mol. The IC_{50}s for compounds 1MWE and 1L7F varied widely over different papers and trials, so we have used an average value for 1L7F and an approximate value of 1 mM for 1MWE, following the precedence of Taylor et al. (ref 15).
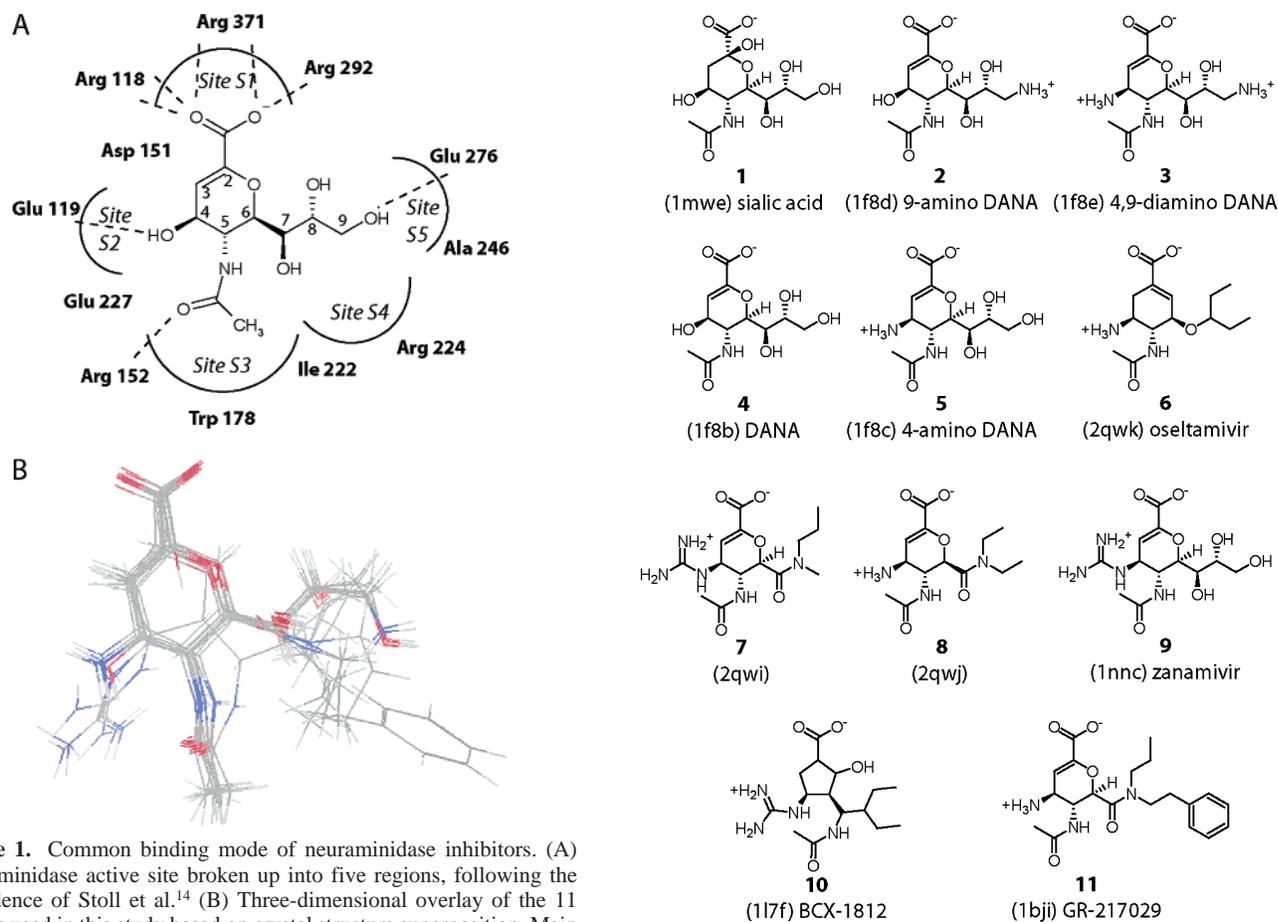
Figure 1. Common binding mode of neuraminidase inhibitors. (A) Neuraminidase active site broken up into five regions, following the precedence of Stoll et al.[14] (B) Three-dimensional overlay of the 11 ligands used in this study based on crystal structure superposition. Main chain atoms of the protein backbone were superposed using MOE (version 2004.03; Chemical Computing Group, Montreal, Quebec, Canada).

Figure 2. Chemical structures of neuraminidase N9-subtype small molecule ligands used in this study. The PDB IDs of the relevant cocrystal structure are given in parentheses, followed by applicable common names.

All 11 ligands have similar five- or six-membered ring scaffolds and bind in very similar modes, as shown in Figure 1B. Even though the optimal electrostatic free energies of binding vary in value between −4.5 and +0.2 kcal/mol, there is good correlation ($r^2 = 0.94$) between the distance from the optimum, $\Delta\Delta G_{elec}^{opt}$, and the reference electrostatic free energy contribution, ΔG_{elec}^{ref} (see Table 1 and Figure 3A). These results from the full charge optimization suggest that the electrostatic gain in binding free energy as the charge distribution of the ligand moves closer to its optimum is significantly greater than the amount that the electrostatics of the optimum changes due

to the new chemical structure and positioning in the active site. The sensitivity of these results to parameter sets has been tested by repeating all calculations using PARSE radii and partial atomic charges (Supporting Information, Table 1). The trends from the PARSE results track closely with the CHARMM PARAM19 results.

The Neuraminidase Series Case Study. Full charge optimization analysis of the neuraminidase substrate sialic acid is presented in Figure 4A. The full optimization suggests substantial room for electrostatic improvement and indicates that
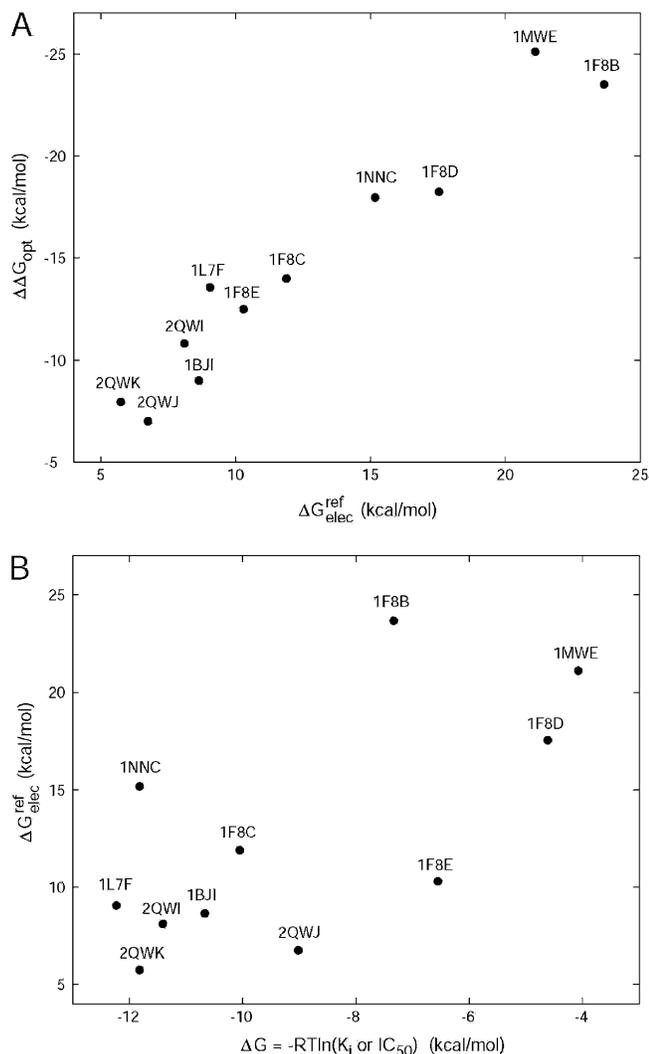
Figure 3. Electrostatic contributions to ligand binding relative to electrostatic differences from the optimum and experimentally measured binding affinities. (A) Scatter plot of the differences from the optimal charge distribution, $\Delta\Delta G_{opt}$, versus calculated electrostatic component of the binding affinity, ΔG_{elec}^{ref}, for the 11 ligands. (B) Scatter plot of calculated electrostatic components, ΔG_{elec}^{ref}, versus experimentally measured binding affinities, $\Delta G = -RT \ln(K_i \text{ or } IC_{50})$, for the 11 ligands. Free energy terms are defined and discussed in the text. Positive values of ΔG_{elec}^{ref} and ΔG represent unfavorable net contributions to binding affinity, while negative values represent favorable contributions.

sialic acid (compound **1**) is a promising starting point for inhibitor design given that its IC_{50} against neuraminidase is about 1 mM.[15] With 18 of the 39 atoms in compound **1** calculated to be greater than 0.5e from the optimal overall charge distribution, a natural question is where to begin applying chemical modifications in order to increase binding affinity. In lead optimization, chemical modification is often focused on a few R groups. To understand the utility of charge optimization in this setting, we focus on modifications at the position 2, 4, 7, 8, and 9 hydroxyl groups of compound **1**. The series of small molecules deriving from compound **1** increase in potency through compounds **2**, **3**, **4**, **5**, and **6**, following a simulated forward lead progression effort.

Performing charge optimization on only the R groups as a set and comparing these optimal charges to the reference partial charges on compound **1** suggests changes to improve binding. The group with the largest charge difference from the optimum is the hydroxyl at position 7, where the oxygen and hydrogen have reference charges of $-0.6e$ and $+0.4e$, respectively.

Comparison with calculated optimal charges of $+0.1e$ and $-1.2e$ for the oxygen and hydrogen atom positions, respectively, suggests conversion of the hydroxyl to a more negative group. The second largest charge differences are found on the hydroxyl groups at positions 2 and 4, where charge optimization suggests (1) converting the hydroxyl group at position 2 with reference charges of $-0.6e$ and $+0.4e$ for the oxygen and hydrogen atoms, respectively, to a more neutrally charged substituent with optimal charges of about $+0.2e$ at both positions and (2) substituting the hydroxyl group at position 4, which has reference charges of $-0.7e$ and $+0.4e$ for the oxygen and hydrogen atoms, respectively, with a positively charged substituent that has optimal charges of $+0.4e$ and $-0.05e$ at the respective positions. At the position 9 hydroxyl group, which has reference charges of $-0.7e$ and $+0.4e$ for the oxygen and hydrogen atoms, respectively, the analysis suggests replacement with a slightly positive group with optimal charges of $+0.3e$ and $-0.1e$ at the respective atomic positions. Finally, neutralizing the hydroxyl group at position 8 from reference charges of $-0.7e$ and $+0.4e$ at the oxygen and hydrogen atom positions to optimal charges of $-0.1e$ and $+0.2e$, respectively, is computed to improve binding.

The suggestions at positions 2 and 9 are in some sense incorporated into compound **2**, with the replacement of the position 9 hydroxyl group with a positive amine group and deletion of the position 2 hydroxyl group. Other modifications are possible that closely match the optimal charge distribution, including, for instance, alkyl substitutions. However, the actual favorable changes are along the lines of the charge optimization suggestions, indicating that charge optimization provides a useful guide. Although the greatest charge differences between reference and optimal are not found at positions 2 and 9, modifications consistent with the optimal charge distribution at these positions do contribute to a more potent compound with a measured K_i of 400 μM. The change in bond order and resulting change in stereochemistry at the 2-position carbon is not accounted for in the calculation of the optimal charges, although comparing the two cocrystal structures (1MWE and 1F8D) shows that these changes do not substantially change the position and conformation of the molecule (see Figure 1B). Analysis of compound **2** suggests converting the hydroxyl group at position 7 to a negatively charged substituent, converting the hydroxyl group at position 4 to a positive group, and converting the hydroxyl at position 8 to a neutral or slightly positive group. These three suggestions are nearly identical to those found for compound **1**. The group at position 9 is now closer to the electrostatic optimum, as can be seen in Figure 4B, although the analysis suggests that bringing the group closer to neutral can improve binding.

Compound **3** can be derived from compound **2** by conversion of the 4-position hydroxyl group to a protonated amine as shown, appearing to satisfy one of the charge optimization suggestions and resulting in a 27-fold improvement in affinity. The amine at position 4 has an overall charge of $+0.6e$ and is very close to the calculated optimal overall charge of $+0.5e$ for the group. A closer look at the atomic charges, however, suggests that the reference charges of $-0.3e$ for the nitrogen and $+0.3e$ for each of the three hydrogens on the amine would more optimally be inverted so that the nitrogen position has a charge of $+0.4e$ and the hydrogens have charges of $-0.1e$, $+0.1e$, and $+0.1e$. This charge inversion would decrease the desolvation penalty while a similar electrostatic interaction energy would be maintained, because the overall charge would
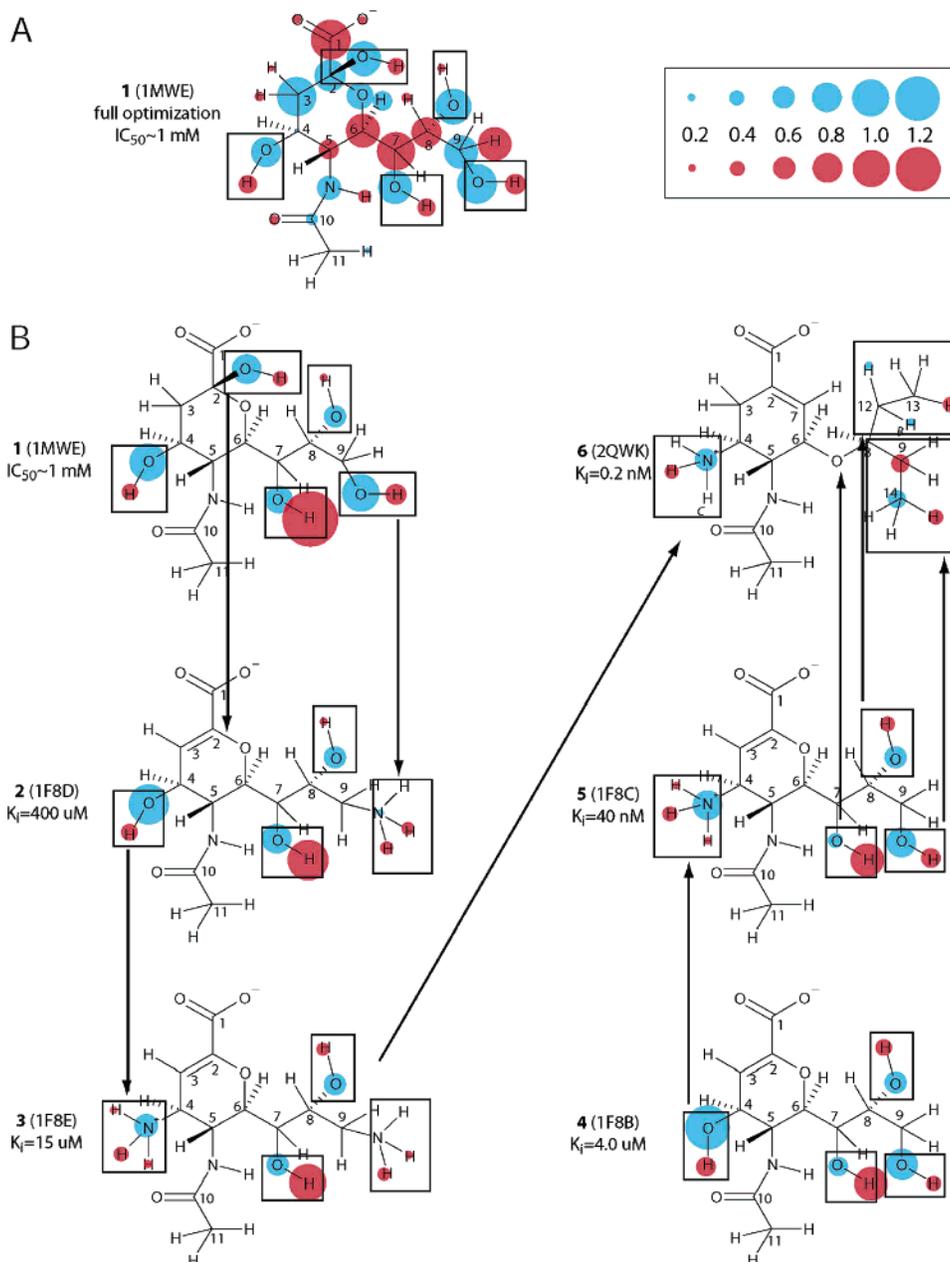
Figure 4. Structure−activity relationship analysis in a subset of neuraminidase ligands showing the suggested charge modifications for each compound. (A) Unconstrained optimization as described in the Methods for 1MWE. (B) Optimization of only the R groups. For each ligand, all groups in boxes were optimized together without constraints. The magnitude of the difference between the real and optimal charge for each atom is depicted by the size and color of the circle centered on each atom, as shown in the box in the upper right corner. Larger differences are indicated by larger circles, and circles are colored blue if the optimal charge is more positive and red if it is more negative. Numbers given in the key are in units of electron charge, and charge differences less than $0.15e$ are not shown.

be the same but the exterior atomic positions would be less charged. Unfortunately, this change appears chemically difficult.

The hydroxyl group at position 7 is still suboptimal in compound **3**, as is, to a lesser degree, the hydroxyl group at position 8. At the position 7 hydroxyl group, substitution of reference charges of $-0.6e$ and $+0.4e$ at the oxygen and hydrogen positions, respectively, with "inverted" optimal charges of $+0.04e$ and $-0.7e$ is predicted to increase binding affinity. One idea that is consistent with the charge optimization suggestion is conversion of the hydroxyl to a carbonyl. An interesting alternative approach to realizing the optimal distribution is found in compound **6.** The hydroxyl group at position 7 is deleted, and in the cocrystal structure (PDB ID 2QWK), a crystallographic water molecule sits near where the hydroxyl group sat in compound **3**. The water hydrogen likely sits at the

hydroxyl oxygen position, thereby satisfying the calculated optimal charge distribution. The ether oxygen of compound **6** probably makes a hydrogen bond to the water in 2QWK and could thus play an important role in stabilizing the water's position. The crystallographic water is not seen in cocrystal structures with compounds **1**−**5**. This analysis together with the ether being calculated to be nearly optimal in compound 6 (see Supporting Information, Figure S1) suggests a contributory role for the crystallographic water.

The 4-fold increase in K_i in moving from compound **3** to compound **4** is not explained by the charge-optimization results, suggesting that other contributions are responsible for the increase in potency. Other studies have pointed out difficulties in explaining the structure-activity relationship (SAR) between compounds **3** and **4**,[43] and possible explanations include

differences in binding site protonation states,[16] differences between binding states in low-temperature crystallographic and room-temperature assay conditions,[43] and differences in internal strain contributions to binding.[44] Compound **4** has most of the same R groups as compound **1**, with the difference being the removal of the hydroxyl group at position 2 and a change in bond order, as described previously. The difference between the reference and optimal charges at the four R groups is essentially the same as found in compound **1**, except for at the hydroxyl group at position 7, where the group is significantly closer to the optimum (see Figure 4B). The optimal charge distribution with $-0.7e$ and $+0.04e$ at the oxygen and hydrogen, respectively, is closer to the reference charges of $-0.6e$ and $+0.4e$, respectively, than found with compound **1**. The optimal charge distribution appears to be more sensitive than the reference charge distribution to the modifications in moving from compound **1** to **4**.

Compound **4** is electrostatically suboptimal at four positions; the hydroxyl group at position 4 should be more positive, the hydroxyl groups at positions 8 and 9 should be neutral, and the hydroxyl group at position 7 should be negative, if the optimal distribution is used as a guide. In compound **5**, an amine replaces the position 4 hydroxyl, and the modification, which is consistent with the first suggestion, results in a 100-fold increase in potency. The overall charge on the amine at position 4 is now $+0.53e$ and is very close to the overall charge of $+0.55e$ found from optimization. Optimal charges for the hydroxyl groups at positions 8 and 9 of compound **5** are both approximately neutral at $+0.1e$, while the optimal charges for the hydroxyl group at position 7 is as described above for compound **3**. Making the hydroxyl groups at positions 8 and 9 neutral by conversion to alkyl groups along with other changes (see Figure 4B) results in compound **6**, which further improves the affinity 200-fold. The change at position 7, described in detail above, appears to involve a crystallographic water. The central ring is also modified, though this change was not made as much for SAR reasons as it was for chemical stability and versatility reasons.[45] Results of full optimization of compound **6** (see Supporting Information) suggest that the R groups are close to their charge optimum, while further charge optimization is possible on the central scaffold.

Application of Charge Optimization in Lead Progression. Medicinal chemistry optimization of neuraminidase inhibitors has focused on the five sites as defined in Figure 1A, and it is interesting to note that compound **6** has a nearly optimal charge distribution at all of these sites, suggesting that lead optimization in this case corresponds to charge optimization. Charge optimization does not provide modifications per se and partly requires chemical intuition to design actual substitutions that satisfy the computed optimal charge distribution. However, it does appear to retrospectively explain the structure−activity relationships seen in the series of neuraminidase inhibitors. Two potential lessons can be taken from this study. The first is that strong suggestions do result in substantial inhibitor potency improvements. On the other hand, slight modification of charges, such as those suggested for compound **6** R groups, may in practice be difficult to make given discrete chemical constraints. A second lesson suggested by the data here is that individual optimizations of these subsites are relatively independent from one another. A lead progression analysis, where all atoms of each ligand were optimized simultaneously, suggests changes to parts of the molecule that are "fixed" in the analysis here (shown in Supporting Information, Figure S1). Even so, the suggested changes to the variable subsites coincide very well

with those suggested in Figure 4B. Likewise, when each of these subsites is optimized individually instead of as a set (shown in Supporting Information, Figure S2), the suggested changes coincide very well with those presented here.

One arising question is whether charge optimization must be applied in a sequential step basis. Comparison of compounds **1** and **6** suggests that the recommended modifications are largely applicable despite small changes such as the rotation of Glu276.[45] In addition, because optimization of these groups is largely independent, changes could potentially have been made concurrently to improve affinity. However, the magnitudes of the charge optimization suggestions do change as we follow the series, as seen in Figure 4B, indicating that the subsites are not completely independent. Since each chemical modification shifts the charge distribution of the entire molecule, and because each ligand sits in the active site in a slightly different conformation, it is likely that charge optimization will be most useful when applied sequentially. In a drug discovery team setting, we envision that a computational chemist applies this method to cocrystal structures of leads and works with a medicinal chemist in finding suitable replacements according to charge optimization suggestions. Quantitative charges for R group replacements can be computed using the RESP methodology (see Methods) before synthesis. In this way, visualization and chemical intuition facilitates optimization of charge-based hydrogen-bond and salt-bridge interactions in a fashion similar to what is possible with shape-based interactions.

Conclusion

The trends associated with the neuraminidase ligands examined here indicate that in a polar active site, calculation of the electrostatic optimum of a ligand can help to suggest useful chemical modifications that generally improve the overall binding affinity. Small modifications in the neuraminidase ligands cause their charge distributions to move significantly with respect to their optima, and the modifications result in large changes in overall binding affinities. These changes are not only due to electrostatics, however. As shown in Figure 3B, the electrostatic component of the free energy of binding does not dominate the total binding affinity. Therefore, the van der Waals, entropy, and molecular strain changes that occur with each new ligand overpower, in some cases, the electrostatic effects in determining the final binding affinity. Because the active site of neuraminidase is quite polar, these effects are not expected to decrease in other targets.

Despite this, charge optimization can be a useful tool, because it can be used to make and evaluate molecular modifications. Charge optimization can suggest useful changes to a molecule in lead progression that consistently improve the electrostatic portion of the binding affinity. We note that neuraminidase is a challenging drug target, and over three decades were required to find a compound that was potent enough for physiological interference. In comparison, inhibitors binding in more hydrophobic pockets, such as those found in kinases and other "drugable" targets, can often be progressed in terms of potency without the aid of cocrystal structures, suggesting that while medicinal chemists can intuitively optimize compounds that are driven by van der Waals packing, it is much more difficult to electrostatically optimize compounds, and computational analysis such as that presented here can be particularly useful in these cases. Pharmaceutical lead progression involves optimization of properties other than potency, such as selectivity and ADME profiles, requiring multifactorial optimization and decision-making. Because charge optimization provides intuitive direc-

tion, it can be incorporated into human generation of compound ideas that balance these various requirements.

Supporting Information Available: Table of electrostatic components computed using the parse parameter set and figures of SAR analyses for full charge optimization of ligands and charge optimization of selected R groups corresponding to changes made in the series. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Davis, A. M.; Teague S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem., Int. Ed.* **1999**, *38*, 736−749.

(2) Kuntz, I. D.; Shoichet, B. K.; Leach, A. R. Ligand solvation in molecular docking. *Proteins* **1999**, *34*, 4−16.

(3) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **1987**, *9*, 327−335.

(4) Gilson, M. K.; Honig, B. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins* **1988**, *4*, 7−18.

(5) Sharp, K. A.; Honig, B. Electrostatic interactions in macromolecules: Theory and applications. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301−332.

(6) Sharp, K. A.; Honig, B. Calculating total electrostatic energies with the nonlinear Poisson−Boltzmann equation. *J. Phys. Chem.* **1990**, *94*, 7684−7692.

(7) Price, D. J.; Klebe, G.; Brooks, C. L., 3rd; Ferrara, P.; Gohlke, H. Assessing scoring functions for protein−ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032−3047.

(8) Lee, L. P.; Tidor, B. Optimization of electrostatic binding free energy. *J. Chem. Phys.* **1997**, *106*, 8681−8690.

(9) Kangas, E.; Tidor, B. Optimizing electrostatic affinity in ligand−receptor binding: Theory, computation, and ligand properties. *J. Chem. Phys.* **1998**, *109*, 7522−7545.

(10) Lee, L. P.; Tidor, B. Optimization of binding electrostatics: Charge complementarity in the barnase−barstar protein complex. *Protein Sci.* **2001**, *10*, 362−377.

(11) Kangas, E.; Tidor, B. Electrostatic complementarity at ligand binding sites: Application to chorismate mutase. *J. Phys. Chem. B.* **2001**, *105*, 880−888.

(12) Sims, P. A.; Wong, C. F.; McCammon, J. A. Charge optimization of the interface between protein kinases and their ligands. *J. Comput. Chem.* **2004**, *25*, 1416−1429.

(13) 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.

(14) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(15) Taylor, N. R.; Cleasby, A.; Singh, O.; Skarzynski, T.; Wonacott, A. J.; Smith, P. W.; Sollis, S. L.; Howes, P. D.; Cherry, P. C.; Bethell, R.; Colman, P.; Varghese, J. Dihydropyrancarboxamides related to zanamivir: A new series of inhibitors of influenza virus sialidases. 2. Crystallographic and molecular modeling study of complexes of 4-amino-4*H*-pyran-6-carboxamides and sialidase from influenza virus types A and B. *J. Med. Chem.* **1998**, *41*, 798−807.

(16) Smith, B. J.; Colman, P. M.; vonItzstein, M.; Danylec, B.; Varghese, J. N. Analysis of inhibitor binding in influenza virus neuraminidase. *Protein Sci.* **2001**, *10*, 689−696.

(17) Smith, B. J.; McKimm-Breshkin, J. L.; McDonald, M.; Fernley, R. T.; Varghese, J. N.; Colman, P. M. Structural studies of the resistance of influenza virus neuraminidase to inhibitors. *J. Med. Chem.* **2002**, *45*, 2207−2212.

(18) Varghese, J. N.; Colman, P. M.; van Donkelaar, A.; Blickand, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L. Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11808−11812.

(19) Varghese, J. N.; Epa, V. C.; Colman, P. M. Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. *Protein Sci.* **1995**, *4*, 1081−1087.

(20) Varghese, J. N.; Smith, P. W.; Sollis, S. L.; Blick, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L.; Colman, P. M. Drug design against a shifting target: A structural basis for resistance to inhibitors in a variant of influenza virus neuraminidase. *Structure* **1998**, *6*, 735−746.

(21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(22) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.

(23) *Gaussian 03 Software*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(24) Bayly, C. I.; Geplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(25) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620−10280.

(26) Nicholls, A.; Honig, B. A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson−Boltzmann equation. *J. Comput. Chem.* **1991**, *12*, 435−445.

(27) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Electrostatics and diffusion of molecules in solutions−Simulations with the University of Houston Brownian dynamics program. *Comput. Phys. Comm.* **1991**, *62*, 187−197.

(28) Oberoi, H.; Allewell, N. M. Multigrid solution of the Poisson−Boltzmann equation. *Biophys. J.* **1993**, *65*, 48−55.

(29) Holst, M.; Saied, F. Multigrid solution of the nonlinear Poisson−Boltzmann equation and calculation of titration curves. *J. Comput. Chem.* **1994**, *98*, 1978−1988.

(30) O'M Bockris, J.; Reddy, A. K. N. *Modern Electrochemistry*; Plenum Press: New York, 1993.

(31) Chong, L. T.; Dempster, S. E.; Hendsch, Z. S.; Lee, L. P.; Tidor B. Computation of electrostatic complements to proteins: A case of charge stabilized binding. *Protein Sci.* **1998**, *7*, 206−210.

(32) Kangas, E.; Tidor, B. Charge optimization leads to favorable electrostatic binding, *Phys. Rev. E.* **1999**, *59*, 5958−5961.

(33) Kangas, E.; Tidor, B. Electrostatic specificity in molecular ligand design. *J. Chem. Phys.* **2000**, *112*, 9120−9131.

(34) Vanderbei, R. J. LOQO: An interior point code for quadratic programming. *Optimization Methods Software* **1999**, *11*, 451−484.

(35) Sarkar, C. A.; Lowenhaupt, K.; Horan, T.; Boone, T. C.; Tidor, B.; Lauffenburger, D. A. Rational cytokine design for increased lifetime and enhanced potency using pH-activated "histidine switching". *Nature Biotechnol.* **2002**, *20*, 908−913.

(36) Ahn, J. S.; Radhakrishnan, M. L.; Mapelli, M.; Choi, S.; Tidor, B.; Cuny, G. D.; Musacchio, A.; Yeh, L. A.; Kosik, K. S. Defining CDK5 ligand chemical space with small molecule inhibitors of tau phosphorylation. *Chem. Biol.* **2005**, *12*, 811−823.

(37) Sulea, T.; Purisima, E. O. Optimizing ligand charges for maximum binding affinity. A solvated interaction energy approach. *J. Phys. Chem. B.* **2001**, *105*, 889−899.

(38) Sulea, T.; Purisima, E. O. Profiling charge complementarity and selectivity for binding at the protein surface. *Biophys. J.* **2003**, *84*, 2883−2896.

(39) Green, D. F.; Tidor, B. *Escherichia coli* glutaminyl-tRNA synthetase is electrostatically optimized for binding of its cognate substrates. *J. Mol. Biol.* **2004**, *10*, 435−452.

(40) Green, D. F.; Tidor, B. Design of improved protein inhibitors of HIV-1 cell entry: Optimization of electrostatic interactions at the binding interface. *Proteins* **2005**, *60*, 644−657.

(41) Stoll, V.; Stewart, K. D.; Maring, C. J.; Muchmore, S.; Giranda, V.; Gu, Y. Y.; Wang, G.; Chen, Y.; Sun, M.; Zhao, C.; Kennedy, A. L.; Madigan, D. L.; Xu, Y.; Saldivar, A.; Kati, W.; Laver, G.; Sowin, T.; Sham, H. L.; Greer, J.; Kempf, D. Influenza neuraminidase inhibitors: Structure-based design of a novel inhibitor series. *Biochemistry* **2003**, *42*, 718−727.

(42) Babu, Y. S.; Chand, P.; Bantia, S.; Kotian, P.; Dahghani, A.; El-Kattan, Y.; Lin, T.; Hutchison, T. L.; Elliott, A. J.; Parker, C. D.; Ananth, S. L.; Horn, L. L.; Laver, G. W.; Montgomery J. A. Bcx-1812 (RWJ-270201): Discovery of a novel, highly potent, orally active, and selective influenza neuraminidase inhibitor through structure-based drug design. *J. Med. Chem.* **2000**, *43*, 3482−3486.

(43) Fornabaio, M.; Cozzini, P.; Mozzarelli, A.; Abraham, D. J.; Kellog, G. E. Simple, intuitive calculations of free energy of binding for protein−ligand complexes. 2. Computational titration and pH effects in molecular models of neuraminidase−inhibitor complexes. *J. Med. Chem.* **2003**, *46*, 4487−4500.

(44) Masukawa, K. M.; Kollman, P. A.; Kuntz, I. D. Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. *J. Med. Chem.* **2003**, *46*, 5628−5637.

(45) Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681−690.

JM051105L