J|A|C|S COMMUNICATIONS

Published on Web 02/15/2010

Role of Secondary Sialic Acid Binding Sites in Influenza N1 Neuraminidase

Jeffrey C. Sung,[†] Adam W. Van Wynsberghe,*[,‡] Rommie E. Amaro,*[,§,||] Wilfred W. Li,[||] and J. Andrew McCammon*[,†,||,⊥]

Departments of Chemistry & Biochemistry and Pharmacology and NSF Center for Theoretical Biological Physics, University of California—San Diego, La Jolla, California 92093, Department of Chemistry, Hamilton College, Clinton, New York, 13323, Department of Pharmaceutical Sciences and Computer Science, University of California—Irvine, Irvine, California 92697, National Biomedical Computation Resource, University of California—San Diego, La Jolla, California 92093, and Howard Hughes Medical Institute, University of California—San Diego, La Jolla, California 92093

Received August 31, 2009; E-mail: avanwyns@hamilton.edu; ramaro@uci.edu

Within influenza viral particles, the intricate balance between host cell binding and sialic acid receptor destruction is carefully maintained by the hemagglutinin (HA) and neuraminidase (NA) glycoproteins, respectively.[1] A major outstanding question in influenza biology is the function of a secondary sialic acid binding site on the NA enzyme (Figure 1). This second site was first discovered through HA activity exhibited by the N9 NA in avian isolates.[2] Subsequent characterization of this strain through escape mutant mutagenesis experiments localized the HA activity to residues on the outer rim of the actual NA binding pocket that play no role in NA catalytic activity.[3] This activity, also termed hemabsorbing (HB) for its ability to bind red blood cells, was shown to be transferable to the human N2 subtype by site-directed mutagenesis.[4] Structural proof for the existence of this secondary sialic acid binding site was later established through X-ray crystallography experiments.[5] The conservation of residues comprising this secondary site in the NA from avian strains, but not human or swine, suggested that this feature may play an unknown biological role, and a careful analysis of individual strains further indicated positive selection pressure in the HB site within the avian strains only.[6] Through a series of Brownian dynamics (BD) simulations of the avian N1 and human N2 enzymes, we have probed the role of this secondary sialic acid binding site in the avian N1 subtype. Our results suggest that electrostatic interactions at the secondary and primary sites in avian NA may play a key role in the recognition process of the sialic acid receptors and catalytic efficiency[7] of NA.

Eight separate sets of BD simulations were carried out using the natural substrate sialic acid and the oseltamivir inhibitor (Supporting Information (SI) Figure S1) and the full tetrameric crystal structures of avian N1 (A/Vietnam/1203/04)[8] and human pandemic N2 (A/Tokyo/3/67).[9] BD simulations are able to model the relative diffusion of two or more biomolecules on biologically relevant time scales[10,11] and allow the calculation of the second-order association rate constant of a ligand to a predefined binding site.[12] In these simulations, the proteins and ligands are free to translate and tumble through an implicit solvent model of water that is viscous and contains random fluctuations in accordance with the fluctuation−dissipation theorem. Intermolecular forces are exclusively electrostatic in nature and are modeled with a modified form of Poisson−Boltzmann electrostatics.[13] In addition, the condition that the molecular surfaces of the two diffusing particles cannot overlap is strictly enforced. As is common in BD simulations, internal degrees of freedom are ignored. Given this treatment of molecular interactions, BD simulations can reliably calculate association rate constants (k_{on}) that are dominated by long-ranged electrostatic steering effects but are not well-suited for calculating dissociation rate constants (k_{off}) that are dominated by shorter ranged (near van der Waals contacts) and more detailed interactions such as the hydrophobic effect. Additionally, our model assumes a rigid body approximation for both the diffusing small molecules as well as the NA enzymes. This is important to consider in light of the fact that the avian N1 strain has been shown both crystallographically[8] and via all-atom molecular dynamics simulations[14] to exhibit remarkable flexibility in the 150- and 430-loop regions. Still, the rigid body approximations present in our model allow us to address the relative association rate of the two molecules to the disparate sites, for a particular atomic configuration. All simulations were carried out using the SDA 5 package.[15] Using simple geometric criteria, we are able to define separate encounter complexes and monitor the kinetics of association to both the active site and the secondary binding site. A close examination of the crystal structures that contained sialic acid bound to the secondary site allowed us to define a set of residue contacts that were used to define the reaction criteria in the BD simulations (SI Figure S2). All simulations were performed under experimental NA reaction conditions: 150 mM NaCl, 310 K, with requisite calcium ions bound;[16] convergence data for the simulations are presented in the SI (SI Figures S4 and S5). As a control, our computational k_{on} for oseltamivir using a reaction criterion of 8.0 Å to the active site of N1 compares well to the experimental 2.52 μM^{-1} s^{-1} for N1.[17] While both the 7.5 Å and 8.0 Å criteria are close to the experimental value, we chose to utilize the 8.0 Å cutoff because many more encounter events were observed at this distance, providing much better statistics. This correspondence with experimental data provides support that these simulations, although highly coarse-grained, are reasonably modeling the physics relevant to ligand binding. The secondary site holds an inherent advantage to binding ligands over the active site given its relatively surface-exposed position. We have estimated this at approximately 20%, as the ratio of the secondary site/active site association rate constants at long reaction criteria (22.5 Å) is approximately 1.2.

We have determined that the fastest k_{on} is for sialic acid binding to the secondary site in the avian N1 and that it binds to this site 3 orders of magnitude faster than to the active site (Table 1, SI Figure S7). This is in contrast to the behavior exhibited by the human N2 enzyme, which has a less than 4-fold difference in

[†] Departments of Chemistry & Biochemistry and Pharmacology and NSF Center for Theoretical Biological Physics, University of California—San Diego.
[‡] Hamilton College.
[§] University of California—Irvine.
[||] National Biomedical Computation Resource, University of California—San Diego.
[⊥] Howard Hughes Medical Institute, University of California—San Diego.

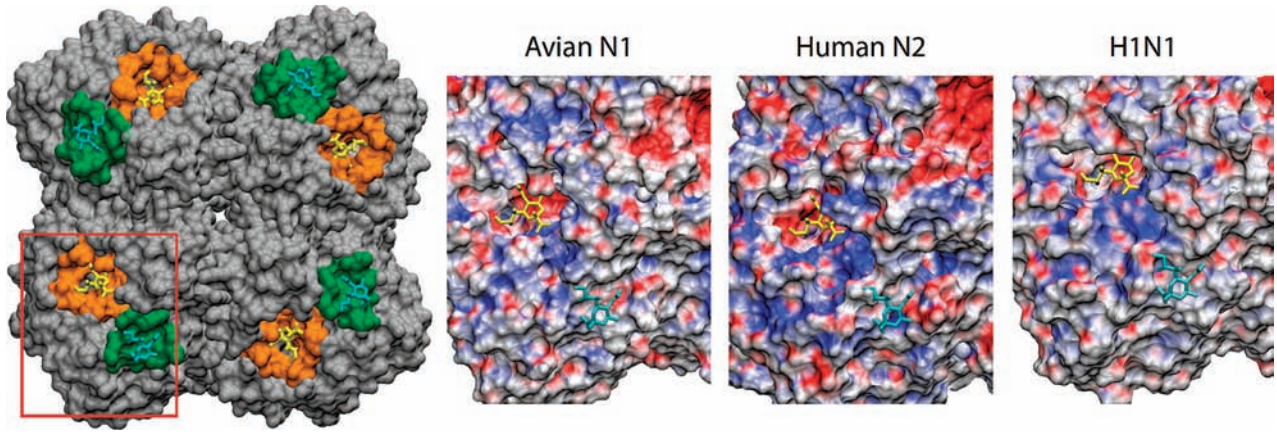

Figure 1. Top view of the N1 tetramer (PDB: 2HTY) depicted in solvent exposed surface area representation; active sites are shown in orange, and secondary sialic acid binding sites are in green (left panel). Electrostatic potential surfaces of the area outlined in the red box are shown in the remaining panels for avian N1 (PDB: 2HTY), human N2 (PDB: 1NN2), and the currently circulating pandemic H1N1 strain (homology model presented in ref 19). Within these surfaces, red represents -10 kcal mol^{-1} e^{-1} and blue represents $+10$ kcal mol^{-1} e^{-1}. Sialic acid molecules are shown as licorice bound in the active (yellow) and secondary (cyan) sites.

association kinetics for the secondary site over the active site (Table 1, SI Figure S7). These results suggest that electrostatic forces largely drive the binding of sialic acid to this external secondary site. The human N2 secondary sialic acid site is more negative in electrostatic potential, compared to the avian N1, which may account for the reduced attraction of the negatively charged sialic acid moieties to this secondary site in the human strain (Figure 1 and SI Figure S3-B). The k_{on} of sialic acid binding to the human N2 active site is more than an order of magnitude faster than that of the avian N1 (Table 1, SI Figure S7). This may explain why the secondary binding site is conserved in avian NA, as it may compensate for the rather weak k_{on} of sialic acid to the avian N1 active site. Recently, loss of the secondary sialic acid binding sites has been shown to correlate with reduced hydrolysis of substrates;[7] therefore this additional sialic acid binding site may facilitate the association of productive encounter complexes with sialic acid moieties on host cell receptors. If and how ligands are transferred between the secondary and active sites is difficult to ascertain in this work given the approximations of the BD simulations, but it is clear that the natural substrate is initially attracted much more rapidly to the secondary site. Alternatively, no transfer is required if the macromolecular substrates are biantennary or are closely juxtaposed on the host cell surface.

The known NA inhibitor oseltamivir has the fastest k_{on} for the secondary site of human N2 enzyme (Table 1, SI Figure S7). The association of oseltamivir to the N2 active site is more than two times faster than to the N1 site. In the N1 enzyme, oseltamivir forms a more frequent encounter complex with the secondary site over the active site, with an almost 2-fold difference (Table 1, SI Figure S7), whereas an almost 7-fold rate difference is observed in N2 (Table 1, SI Figure S7). In both strains, oseltamivir exhibits a larger k_{on} rate constant to the active site than sialic acid. These results are consistent with the fact that the inhibitor was designed using the group-2 strains.[18] They also highlight a potential role of the secondary site in the association with inhibitors, where the loss of this feature may be a potential mechanism of drug resistance.

Table 1. Association Rates (k_{on}) of Ligands to the Active Site and Secondary Site of Neuraminidases from Three Influenza Strains[a]

	$k_{on,active\ site}$ ($\mu M^{-1}\cdot s^{-1}$)	$k_{on,secondary\ site}$ ($\mu M^{-1}\cdot s^{-1}$)	$k_{on,secondary\ site}/k_{on,active\ site}$
Oseltamivir-N1-a	5.17 (8×10^{-4})	9.73 (1×10^{-3})	1.88
Sialic Acid-N1-a	9.41×10^{-2} (1×10^{-4})	208 (5×10^{-3})	2210
Oseltamivir-N2-h	12.1 (2×10^{-3})	84.2 (6×10^{-3})	6.96
Sialic Acid-N2-h	0.503 (5×10^{-4})	1.78 (1×10^{-3})	3.54
Sialic Acid-H1N1-h	0.168 (8×10^{-4})	12.6 (1×10^{-3})	75.0

[a] "-a" and "-h" denote avian- or human-derived strains, respectively. Standard errors of the mean are shown in parentheses.

Although the secondary sialic acid binding site has previously been shown to be nonconserved in swine NA strains,[6] the currently circulating pandemic H1N1 strain (A/California/04/2009) with a proposed swine origin appears to have retained some of the key features of the secondary sialic acid binding site. A sequence analysis showing the residues responsible for sialic acid binding in the secondary site (Figure 2) indicates that 5 of the 6 residues responsible for sialic acid binding are conserved in the current H1N1 strain. Moreover, examination of a recently published H1N1 NA homology model structure[19] shows that the lone nonconserved residue, a serine to asparagine mutation, preserves the hydrogen bonding interaction present in the avian N1 strain (SI Figure S6). The sequence conservation and high number of conserved interactions with the secondary binding site in the pandemic H1N1 strain indicate that this strain may have retained some features of this secondary sialic acid binding site. With this in mind, we pursued similar computational experiments on the pandemic N1 structure.[19]

	366-373	399-403	430-433
Avian N1	KSTNSRSG	AITDWS	RPKE
Human N2	ISKDLRSG	DSDNRS	RKQE
H1N1	KSISSRNG	GINEWS	RPKE

Figure 2. A sequence comparison of critical secondary sialic acid binding site residues in the avian N1, human N2, and currently circulating pandemic strain with possible swine origin is shown. Key conserved residues that make contact with sialic acid are underlined, as derived from ref 6.

Our results indicate that the currently circulating H1N1 strain presents an intermediary case of secondary sialic acid association kinetics, exhibiting a 75-fold difference between the k_{on} rate to the secondary site and the active site (Table 1, SI Figure S7). This is a result of both a faster k_{on} to the active site and a slower k_{on} to the

secondary site than in the avian N1 strain. Despite retaining 5 of 6 key sialic acid binding residues in the secondary site, the k_{on} to the secondary site for the circulating H1N1 strain is roughly 16 times lower than that for the avian N1 strain. This result indicates possible lowered HB activity for this secondary sialic acid site and may be an important event in the emergence of the current pandemic strain.

The human N2 NA from the epidemic strain has a sialic acid k_{on} to the active site that is ~8 times faster than that of the pandemic N1 NA (Table 1, SI Figure S7), whereas the pandemic N1 NA has a k_{on} to the secondary site that is more than 2 times faster than that of the human N2 NA. The lower k_{on} for the latter may be due to the additional charged residues missing in the pandemic strain. Thus, a complex interplay of structural topology and electrostatics seems to be important in driving the association of the small molecules to the various sites on NA.

This BD analysis allows us to speak to the nature of encounter complex formation events for the terminal sialic acid moieties on the host cell surface receptors. Understanding that the actual sialic acid receptors on target host cells are more complex than a single sialic acid molecule, but recognizing that this terminal group plays a crucial role in the recognition process, we find these results quantitatively suggest that this secondary binding site may play an important role in the recognition of these receptors by N1. The secondary sialic acid binding site, which is adjacent to the actual NA active site, appears to facilitate the formation of complexes with the NA protein and the sialic acid receptors and may even provide supplemental HA activity. Moreover, this site is able to steer inhibitor binding as well, albeit with a reduced capacity in N1, and may have potential implications for drug resistance or optimal inhibitor design. Specifically, more realistically accounting for detailed kinetic information, in addition to traditional thermodynamic considerations, could be advantageous for the rational design of small molecule inhibitors, where the design of inhibitors that have a faster k_{on} to the active site as compared to the competing sialic acid substrate would be advantageous.

The exact role of this secondary sialic acid site in the emergence of a pandemic virus remains to be experimentally determined. The lack of secondary site conservation in human NA strains to date may suggest that loss of this feature may actually promote transmissibility and viral survival among humans, and this has very recently been shown for the 1918, 1957, and 1968 pandemic strains.[7] The avian N1, with its strong secondary sialic acid binding site character, is not highly transmissible among humans, although when contracted it is highly virulent, with a fatality rate of ~65%.[20,21] In contrast, the human pandemic N2 is highly transmissible but exhibits reduced virulence, with a fatality rate of ~1%. The currently circulating pandemic N1 strain is indeed highly transmissible but less lethal. Yet unanswered is whether the enhancement of the secondary sialic acid binding would promote virulence in the human host. It seems conceivable that the secondary sialic acid binding site feature could potentially be related to increased severity in human-contracted avian strains, through the promotion of modified recognition events in the respiratory tract, where avian-like α2,3 linkages on the terminal sialic acid receptors are found present in humans.[22–24]

Acknowledgment. The authors thank Razif Gabdoulline for helpful discussions and an advance release of SDA 5. This work was funded in part by NIH F32-GM084595 to A.V.W.; NIH F32-GM077729 and MRAC CHE060073N to R.E.A.; and NIH GM31749, NSF MCB-0506593, and MCA93S013 to J.A.M. W.W.L. is supported by NIH P41 RR08605. Additional support from the Howard Hughes Medical Institute, the National Biomedical Computation Resource, and the Center for Theoretical Biological Physics is gratefully acknowledged.

Supporting Information Available: Seven additional figures, five additional tables, and computational methodology details. This material is available free of charge via the Internet at http://pubs.acs.org

References

(1) Wagner, R.; Matrosovich, M.; Klenk, H. D. *Rev Med Virol* **2002**, *12*, 159–66.
(2) Laver, W. G.; Colman, P. M.; Webster, R. G.; Hinshaw, V. S.; Air, G. M. *Virology* **1984**, *137*, 314–23.
(3) Webster, R. G.; Air, G. M.; Metzger, D. W.; Colman, P. M.; Varghese, J. N.; Baker, A. T.; Laver, W. G. *J. Virol.* **1987**, *61*, 2910–6.
(4) Nuss, J. M.; Air, G. M. *Virology* **1991**, *183*, 496–504.
(5) Varghese, J. N.; Colman, P. M.; van Donkelaar, A.; Blick, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11808–12.
(6) Matrosovich, M. N.; Krauss, S.; Webster, R. G. *Virology* **2001**, *281*, 156–62.
(7) Uhlendorff, J.; Matrosovich, T.; Klenk, H. D.; Matrosovich, M. *Arch. Virol.* **2009**, *154*, 945–57.
(8) Russell, R. J.; Haire, L. F.; Stevens, D. J.; Collins, P. J.; Lin, Y. P.; Blackburn, G. M.; Hay, A. J.; Gamblin, S. J.; Skehel, J. J. *Nature* **2006**, *443*, 45–9.
(9) Varghese, J. N.; Colman, P. M. *J. Mol. Biol.* **1991**, *221*, 473–86.
(10) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329–341.
(11) McGuffee, S. R.; Elcock, A. H. *J. Am. Chem. Soc.* **2006**, *128*, 12098–12110.
(12) Wade, R. C.; Gabdoulline, R. R.; Ludemann, S. K.; Lounnas, V. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5942–5949.
(13) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868–3878.
(14) Amaro, R. E.; Minh, D. D.; Cheng, L. S.; Lindstrom, W. M., Jr.; Olson, A. J.; Lin, J. H.; Li, W. W.; McCammon, J. A. *J. Am. Chem. Soc.* **2007**, *129*, 7764–5.
(15) Gabdoulline, R. R.; Wade, R. C. *J. Am. Chem. Soc.* **2009**, *131*, 9230–9238.
(16) Chong, A. K.; Pegg, M. S.; von Itzstein, M. *Biochim. Biophys. Acta* **1991**, *1077*, 65–71.
(17) Collins, P. J.; Haire, L. F.; Lin, Y. P.; Liu, J.; Russell, R. J.; Walker, P. A.; Skehel, J. J.; Martin, S. R.; Hay, A. J.; Gamblin, S. J. *Nature* **2008**, *453*, 1258–61.
(18) Colman, P. *Structure-Based Drug Discovery: An Overview*; Royal Society of Chemistry, 2006.
(19) Maurer-Stroh, S.; Ma, J.; Lee, R.; Sirota, F.; Eisenhaber, F. *Biology Direct* **2009**, *4*, 18.
(20) Normile, D. *Science* **2008**, *319*, 1178–1179.
(21) World Health Organization, *Communicable disease surveillance and response*, 2008.
(22) Shinya, K.; Ebina, M.; Yamada, S.; Ono, M.; Kasai, N.; Kawaoka, Y. *Nature* **2006**, *440*, 435–436.
(23) van Riel, D.; Munster, V. J.; de Wit, E.; Rimmelzwaan, G. F.; Fouchier, R. A. M.; Osterhaus, A. D. M. E.; Kuiken, T. *Am. J. Pathol.* **2007**, *171*, 1215–1223.
(24) van Riel, D.; Munster, V. J.; de Wit, E.; Rimmelzwaan, G. F.; Fouchier, R. A. M.; Osterhaus, A. D. M. E.; Kuiken, T. *Science* **2006**, *312*, 399.

JA9073672