Glycan Microarrays

DOI: 10.1002/anie.201200596

Recognition of Sialylated Poly-*N*-acetyllactosamine Chains on *N*- and *O*-Linked Glycans by Human and Avian Influenza A Virus Hemagglutinins**

*Corwin M. Nycholat, Ryan McBride, Damian C. Ekiert, Rui Xu, Janani Rangarajan, Wenjie Peng, Nahid Razi, Michel Gilbert, Warren Wakarchuk, Ian A. Wilson, and James C. Paulson**

The initial stages of an influenza A virus infection are mediated by the binding of the viral hemagglutinin (HA) to sialylated glycan receptors on host epithelial cells.[1] The specificity of the HA is believed to be a key determinant of viral host range.[2] While all 16 influenza HA subtypes are found in avian viruses, only three are found in viruses adapted to humans (H1, H2, and H3), each resulting in a major pandemic. HAs from avian and human viruses are characterized by their preference for α2-3 and α2-6-linked sialic acids, respectively. Studies now suggest that other elements of sialoglycan sequence are also important factors of HA specificity that contribute to the species barrier.[3] Recently, human and swine respiratory epithelial cells were shown to express sialylated N-linked glycans with extended poly-*N*-acetyllactosamine (poly-LacNAc) chains.[4] Poly-LacNAc chains are Galβ1-4GlcNAcβ1-3 tandem repeats that extend *N*- and *O*-linked glycans of glycoproteins and contribute to the biology mediated by glycan-binding proteins.[5] Sasisekharan and co-workers have suggested that human HAs bind preferentially to extended α2-6 sialosides and may be critically important for viral adaptation to humans.[4a,6]

Studies on the preference of influenza HAs for extended glycans have employed synthetic sialosides that are linear terminal fragments of natural *N*- and *O*-linked glycans, which differ in their core structure and are often branched.[4a,7] To more fully address the influence of poly-LacNAc chains on HA specificity in the context of natural glycans, we have synthesized a series of sialylated poly-LacNAc structures on intact *O*- (**4**–**9**, **16**–**21**) and *N*-linked glycan (**10**–**12**, **22**–**24**) cores (Figure 1). These sialosides were incorporated into a custom glycan microarray alongside the linear terminal fragments (**1**–**3**, **13**–**15**) for analysis of specificities of human and avian influenza HAs.

Several groups have reported chemical and chemoenzymatic syntheses of poly-LacNAc structures.[8] For synthesis of extended natural *N*- and *O*-linked glycans, our strategy relied on enzymatic elaboration of advanced core structures. The α2-3 and α2-6 sialoside targets comprised *O*-linked cores

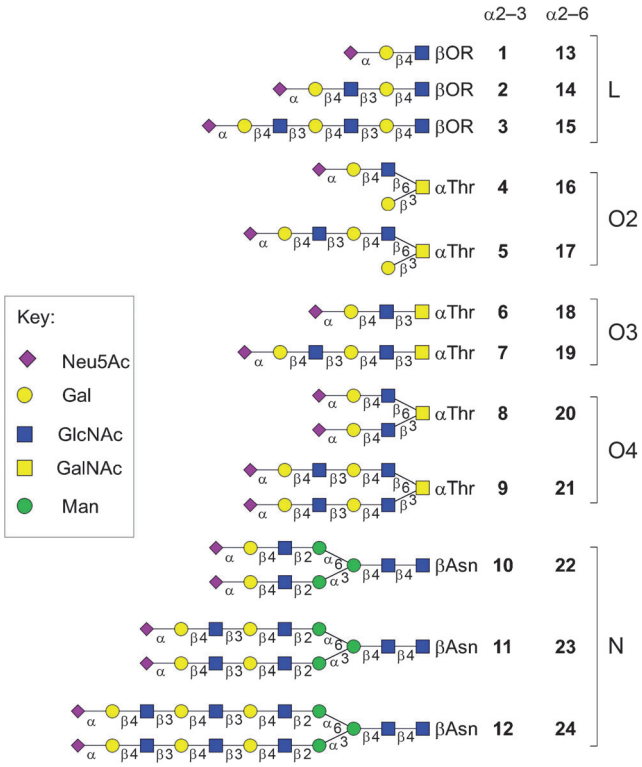


Figure 1. Structures of sialylated poly-LacNAc linear (L) fragments (**1**–**3**, **13**–**15**) and the same sequences elaborated on *O*-linked glycan cores 2–4 (O2, O3, O4; **4**–**9**, **16**–**21**) and *N*-linked (N) glycan cores (**10**–**12**, **22**–**24**). α2-3 and α2-6 indicate the linkage of the sialic acid moiety. Neu5Ac = sialic acid, GlcNAc = *N*-acetylglucosamine, GalNAc = *N*-acetylgalactosamine, R = ethyl amine.

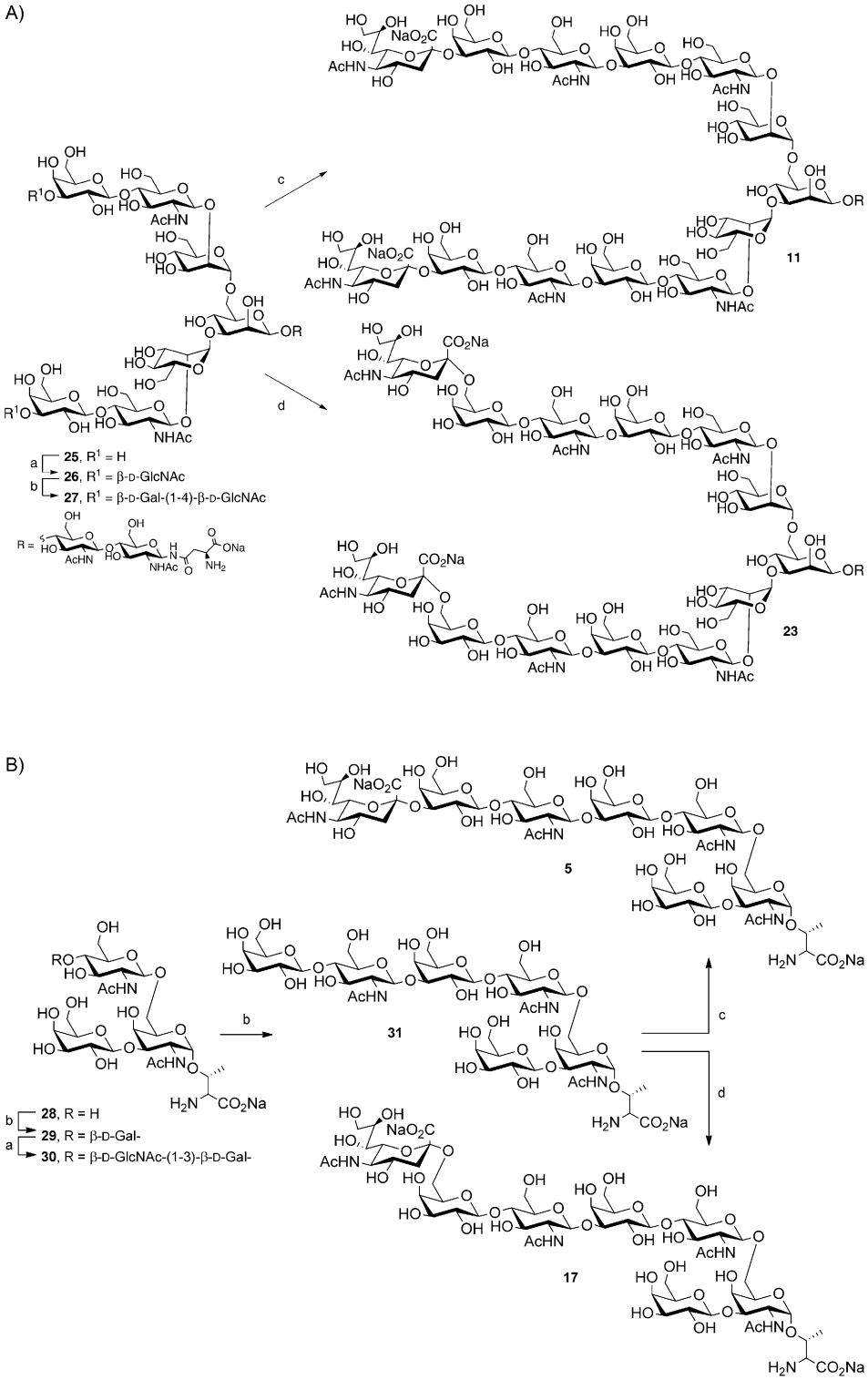
[*] Dr. C. M. Nycholat, R. McBride, J. Rangarajan, Dr. W. Peng, Dr. N. Razi, Prof. J. C. Paulson
Department of Chemical Physiology, The Scripps Research Institute
10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)
E-mail: jpaulson@scripps.edu
Dr. D. C. Ekiert, Dr. R. Xu, Prof. I. A. Wilson
Department of Molecular Biology, The Scripps Research Institute
10550 N. Torrey Pines Road, La Jolla, CA 92037 (USA)
Dr. M. Gilbert, Dr. W. Wakarchuk
Institute for Biological Sciences, National Research Council Canada
Ottawa, ON K1A 0R6 (Canada)

[**] We thank the Otsuka Chemical Co., Ltd for the generous gift of N-linked glycans (**10** and **22**), and the Consortium for Functional Glycomics (GM62116) for providing the linear sialosides used in these studies. This work was supported in part by NIH grants AI058113 (to J.C.P. and I.A.W.), GM62116 (to J.C.P.), a predoctoral fellowship from the Achievement Rewards for College Scientists Foundation (to DCE); grant GM080209 from the NIH Molecular Evolution Training Program (to D.C.E.), and the Skaggs Institute (I.A.W.).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201200596.

(cores 2–4) and *N*-linked cores with up to two and three LacNAc repeats, respectively. Representative syntheses for *N*-linked glycans (**11** and **23**) and *O*-linked glycans with core 2 (**5** and **17**) are described in Scheme 1. Key LacNAc extensions were attained by alternating reactions using recombinant *Helicobacter pylori* β1-3-*N*-acetylglucosaminyltransferase (β1-3GlcNAcT)[8b] and the bacterial β1-4-galactosyltransferase/UDP-4′-Gal-epimerase fusion protein (GalT-GalE).[9] Reaction of *N*-linked glycan **25** with UDP-GlcNAc (4 equiv) using enzyme β1-3GlcNAcT and subsequent treatment with UDP-Glc (4 equiv) and GalT-GalE allowed efficient construction of LacNAc on both antennae affording **27** (Scheme 1 A). Divergent sialylation of **27** using rat α2-3-sialyltransferase (rST3Gal-III) or human α2-6-sialyltransferase (hST6Gal-I), with CMP-Neu5Ac gave the desired α2-3-linked sialoglycan **11** and α2-6-linked sialoglycan **23**, respectively. The synthesis of *O*-linked cores 3–4 and of the tri-LacNAc *N*-linked glycans were conducted using similar conditions (Schemes S1–S6 in the Supporting Information).

The β1-6 branch of core-2 *O*-linked glycans are commonly extended with poly-LacNAc. Initial galactosylation of **28** added Galβ1-4 to GlcNAc giving **29** (Scheme 1 B). As both branches of **29** present terminal Gal residues, two sites were potentially reactive for GlcNAc addition. Regioselective reaction on the β1-6 branch was anticipated, because the enzyme β1-3GlcNAcT demonstrates higher selectivity for Galβ1-4GlcNAc substrates. Thus, under controlled conditions using UDP-GlcNAc (2 equiv), selective elongation of the β1-6 branch was achieved to afford branched glycan **30**.[10] NMR spectroscopy and MS analysis confirmed addition of a single GlcNAc unit. The asialo di-LacNAc structure **31** was prepared by reaction of **30** with UDP-Glc catalyzed by fusion protein GalT-GalE. Finally, selective sialylation of **31** was performed with either enzyme rST3Gal-III or enzyme hST6Gal-I and CMP-Neu5Ac (2 equiv). Both sialyltransferases show preference for Galβ1-4GlcNAc substrates and



Scheme 1. A) Enzymatic transformations of **25** to **11** and **23**: a) enzyme β1-3GlcNAcT, UDP-GlcNAc; b) fusion protein GalT-GalE, UDP-Glc; c) sialyltransferase rST3Gal-III, CMP-Neu5Ac; d) sialyltransferase hST6Gal-I, CMP-Neu5Ac. B) Enzymatic transformation of **28** to **5** and **17**. See Scheme 1 A for conditions. UDP = uridine-5′-diphospate, CMP = cytidine monophosphate.

gave compounds **5** and **17**, respectively. The mono-sialylated products were confirmed by NMR spectroscopy and MS analysis.

The 24 glycans in the sialoside library (Figure 1) contain either the terminal Neu5Acα2-3Gal (**1**–**12**) or Neu5Acα2-6Gal (**13**–**24**) sequence. A glycan microarray was constructed from this library to study the binding properties of influenza A virus HA.[7a,11] The aglycone of each sialoside was equipped with a free amino group for direct printing on slides activated with *N*-hydroxy succinimide (Figure S1 in the Supporting Information). Recombinant HAs from selected avian and human influenza A viruses were then screened to assess the effects on HA binding of both length and presentation of sialylated poly-LacNAc.

As expected, the avian HAs preferentially recognized α2-3-linked sialosides (Figure 2 and Figure S2 in the Supporting Information). However, while hemagglutinin H4 (A/duck/Czech/1/56) bound strongly to nearly all α2-3 structures, other avian HAs showed more selective binding patterns. For instance, hemagglutinin H3 (A/duck/Ukr/1/63), a progenitor of the 1968 Hong Kong pandemic,[12] only bound the linear glycans **2** and **3** and the *O*-linked glycan **8** and *N*-linked glycan

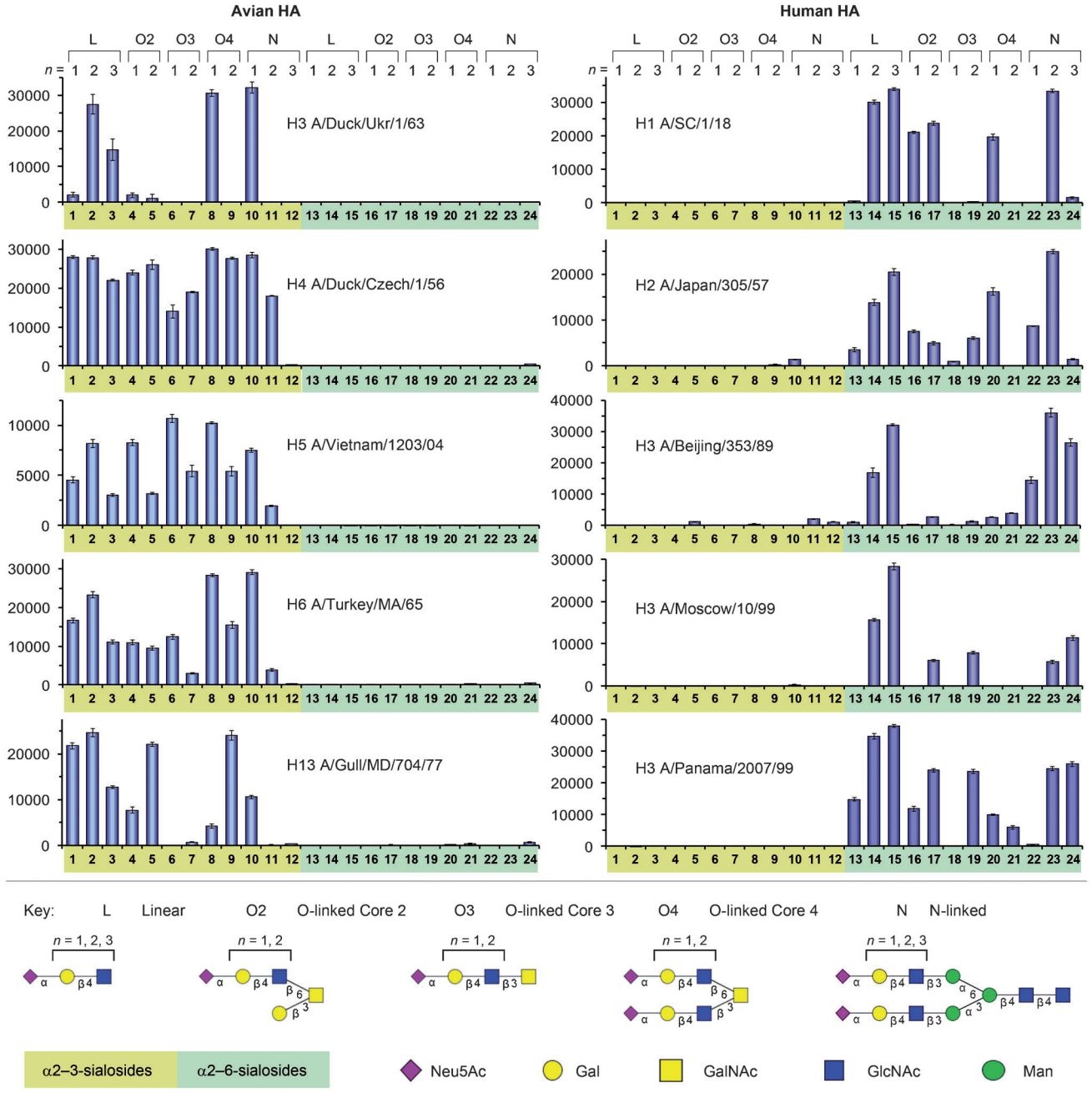


Figure 2. Glycan microarray binding analyses as measured by fluorescence intensity for avian and human influenza A recombinant hemagglutinins. All HAs were evaluated at 15 μg mL^{-1} except for A/SC and A/Beijing, which were evaluated at 150 μg mL^{-1}. See additional details in the Supporting Information.

10. Remarkably, all avian HAs, including H5 (A/Vietnam/1203/04), a highly pathogenic human isolate of the bird flu,[13] showed strong preference for short *N*-linked structures, thus binding strongly to **10**, and showed reduced or no binding to the longer glycans **11** and **12**.

Although human HAs demonstrated classic preference for α2-6 sialosides, they exhibited varied fine specificity for the extended *N*- and *O*-linked glycans (Figure 2 and Figure S2 in the Supporting Information). As reported, the human HAs bound best to the linear sialosides with di- and tri-LacNAc extensions (**14**–**15**).[4] Significantly, however, the same sequences were not uniformly recognized when presented on *N*- and *O*-linked glycan cores. For instance, while the H1 (A/SC/1/18) and the H2 (A/Japan/305/57) HAs bound strongly to the linear sialoside with the di-LacNAc extension (**14**), they bound poorly to the same sequence presented on glycans of core 3 (**19**) and core 4 (**21**). Surprisingly, these same two HAs exhibited strong binding to *N*-linked glycans with the di-LacNAc sequence (**23**) but dramatically reduced binding to the same sequence with the tri-LacNAc repeat (**24**).

In summary, we have synthesized a panel of novel glycans containing sialylated poly-LacNAc on intact *N*- and *O*-linked glycan cores as candidates of the natural glycan receptors of influenza viruses. While all avian and human virus HAs retained their basic specificity for α2-3 and α2-6 linkages, respectively, the *N*- and *O*-linked glycan cores differentially impacted the ability of individual HAs to recognize the sialic acid as a receptor. The lack of a consistent recognition pattern for human HAs suggests that the fine specificity of the virus for receptor(s) may drift under antigenic selective pressure, while the ability to bind to a subset of α2-6 sialosides is sufficiently retained to mediate infection and transmission. It should also be noted that the branched *N*-linked glycans and *O*-linked glycans with core 4, which were produced with our synthetic strategy, are symmetric di-sialylated glycans. However, glycans extended on a single branch also occur in nature.[4a,b] Thus, it will also be of interest to investigate the role of asymmetric glycans on influenza receptor biology.

Received: January 21, 2012
Published online: April 13, 2012

Keywords: glycans · hemagglutinins · influenza viruses · microarrays · sialic acids

[1] a) G. Neumann, Y. Kawaoka, *Emerging Infect. Dis.* **2006**, *12*, 881 – 886; b) R. Salomon, R. G. Webster, *Cell* **2009**, *136*, 402 – 410.
[2] R. J. Connor, Y. Kawaoka, R. G. Webster, J. C. Paulson, *Virology* **1994**, *205*, 17 – 23.
[3] a) M. N. Matrosovich, A. S. Gambaryan, H. D. Klenk in *Avian Influenza*, *Vol. 27* (Eds.: H. D. Klenk, M. N. Matrosovich, J. Stech), Karger, Basel, **2008**, pp. 134 – 155; b) J. M. Nicholls, R. W. Chan, R. J. Russell, G. M. Air, J. S. Peiris, *Trends Microbiol.* **2008**, *16*, 149 – 157.
[4] a) A. Chandrasekaran, A. Srinivasan, R. Raman, K. Viswanathan, S. Raguram, T. M. Tumpey, V. Sasisekharan, R. Sasisekharan, *Nat. Biotechnol.* **2008**, *26*, 107 – 113; b) A. C. Bateman, R. Karamanska, M. G. Busch, A. Dell, C. W. Olsen, S. M. Haslam, *J. Biol. Chem.* **2010**, *285*, 34016 – 34026; c) T. R. Maines, A. Jayaraman, J. A. Belser, D. A. Wadford, C. Pappas, H. Zeng, K. M. Gustin, M. B. Pearce, K. Viswanathan, Z. H. Shriver, R. Raman, N. J. Cox, R. Sasisekharan, J. M. Katz, T. M. Tumpey, *Science* **2009**, *325*, 484 – 487.
[5] a) M. J. Cho, R. D. Cummings, *Trends Glycosci. Glycotechnol.* **1997**, *9*, 47 – 56; b) R. Renkonen, R. Niemela, J. Natunen, M. L. Majuri, H. Maaheimo, J. Helin, J. B. Lowe, O. Renkonen, *J. Biol. Chem.* **1998**, *273*, 4021 – 4026; c) K. Sasaki, K. Kurata-Miura, M. Ujita, K. Angata, S. Nakagawa, S. Sekine, T. Nishi, M. Fukuda, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 14294 – 14299.
[6] K. Viswanathan, A. Chandrasekaran, A. Srinivasan, R. Raman, V. Sasisekharan, R. Sasisekharan, *Glycoconjugate J.* **2010**, *27*, 561 – 570.
[7] a) J. Stevens, O. Blixt, J. C. Paulson, I. A. Wilson, *Nat. Rev. Microbiol.* **2006**, *4*, 857 – 864; b) R. A. Childs, A. S. Palma, S. Wharton, T. Matrosovich, Y. Liu, W. Chai, M. A. Campanero-Rhodes, Y. Zhang, M. Eickmann, M. Kiso, A. Hay, M. Matrosovich, T. Feizi, *Nat. Biotechnol.* **2009**, *27*, 797 – 799.
[8] a) O. Blixt, N. Razi, *Methods Enzymol.* **2006**, *415*, 137 – 153; b) B. Sauerzapfe, K. Krenek, J. Schmiedel, W. W. Wakarchuk, H. Pelantova, V. Kren, L. Elling, *Glycoconjugate J.* **2009**, *26*, 141 – 159; c) T. K. Mong, C. Y. Huang, C. H. Wong, *J. Org. Chem.* **2003**, *68*, 2135 – 2142.
[9] O. Blixt, J. Brown, M. J. Schur, W. Wakarchuk, J. C. Paulson, *J. Org. Chem.* **2001**, *66*, 2442 – 2448.
[10] W. Peng, J. Pranskevich, C. M. Nycholat, M. Gilbert, W. Wakarchuk, J. C. Paulson, N. Razi, **2012**, unpublished results..
[11] J. Stevens, O. Blixt, L. Glaser, J. K. Taubenberger, P. Palese, J. C. Paulson, I. A. Wilson, *J. Mol. Biol.* **2006**, *355*, 1143 – 1155.
[12] Y. Ha, D. J. Stevens, J. J. Skehel, D. C. Wiley, *Virology* **2003**, *309*, 209 – 218.
[13] J. Stevens, O. Blixt, T. M. Tumpey, J. K. Taubenberger, J. C. Paulson, I. A. Wilson, *Science* **2006**, *312*, 404 – 410.

