Tetrahedron Letters 50 (2009) 6130–6132

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

Recognition between a divalent sialyl molecule and wheat germ agglutinin

Yi-Ping Yu [a], An-Tai Wu [b], Wei Zou [c], Chien-Sheng Chen [d,*], Shih-Hsiung Wu [d,*]

[a] Department of Food, Health and Nutrition Science, Chinese Culture University, Taipei 11114, Taiwan
[b] Department of Chemistry, National Changhua University of Education, Changhua 50058, Taiwan, ROC
[c] Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario, Canada K1A 0R6
[d] Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan

ARTICLE INFO

Article history:
Received 8 May 2009
Revised 14 August 2009
Accepted 21 August 2009
Available online 26 August 2009

Keywords:
Sialic acid
Wheat germ agglutinin
Surface plasmon resonance
Isothermal titration calorimetry
Molecular modeling

ABSTRACT

Structural divalency between a designed *N*-acetyl-neuraminic acid (NeuAc)-containing molecule and lectin wheat germ agglutinin (WGA) is investigated. The sialyl molecule was designed based on the NeuAc–WGA complex in the Protein Data Bank and featured polyethylene glycol linkers connecting to an aromatic scaffold. Our results elucidate the divalent recognition association constant between WGA and the multivalent-NeuAc molecules to be 10^7 by surface plasmon resonance.

Wheat germ agglutinin (WGA), a lectin consisting of two identical subunits with four binding domains (A–D) in total, specifically binds to either four *N*-acetyl-neuraminic acid (NeuAc)[1] or eight *N*-acetyl-glucosamine (GlcNAc)[2–5] residues. The intrinsic monovalent affinities of both were determined by isothermal titration calorimetry (ITC), in which the association constant[1,6,7] of GlcNAc–WGA is 10^3 M^{-1} and of NeuAc–WGA is 172 M^{-1}. Interestingly, the association constant of the divalent GlcNAc–WGA was measured as 10^7 M^{-1}, suggesting a 10,000-fold increase in binding affinity compared to monovalence.[8] The association constant of divalent NeuAc–WGA interaction remains elusive, in which traditional affinity purification of sialyl glycoprotein with immobilized WGA technique was applied to determine the binding action.[9] Here, we report the divalent binding properties of the NeuAc–WGA by using surface plasmon resonance (SPR).

Regarding the structural stability of WGA related to its functionality, the effective multivalent recognition of sialyl-containing molecules is more flexible and worth considering in designing new molecules.[10–12] The crystal structure of WGA was solved with potential substrate-binding sites A, B, C, and D. The primary sequences of sites A and C are identical to those of sites B and D. The distances between sites C and D, A and C, and A and B are around 20, 20, and 40 Å, respectively. Based on the crystal structure of NeuAc–WGA complex, Wright and colleagues showed that sites A and C of WGA could recognize two sialoglycopeptides[13,14], and sites B and D were predicted to have sialyl-binding abilities using a hydropathic interaction modeling program (HINT).[2] Appropriate distance between a pair of binding sites is absolutely essential for divalent recognition. If the distances are too long, it may cause hindrance or internal rotation between ligand-containing molecules and receptors. Such assumption is based on the trimeric binding activities of hemagglutinin (HA) trimer[15–17], where each NeuAc-recognition is located in the inner grooves of each HA subunit and physically separated up to 46 Å. Glick and Knowles had synthesized two classes of divalent sialosides with a linker span of up to 65 Å[18], and these divalent molecules associate with HA in a monovalent fashion.[19] Therefore, we chose to study the sites C and D of NeuAc–WGA divalent recognition because of the shorter distance and less potential hindrance between the binding pair. The A–C site is not a good candidate because of internal hindrance, and the distance between sites A and B is too long for divalent recognition.

We designed and constructed molecules **1m**, **1d**, and **1t**, which consist of different number of sialyl residues connected by a linker to an aromatic scaffold, a spacer, and then a biotinyl group (Fig. 1a). The biotinyl group allows the molecules (**1m**, **1d**, and **1t**) to attach to a BIAcore streptavidin-coated sensor chip and subsequent kinetic study by surface plasmon resonance. The linker adjusts the distance of the sialyl residue to the divalent C and D sites in WGA. To determine the optimal length of the linker, we constructed molecule **7** at the divalent C–D site of WGA by modifying the WGA–NeuAc crystal structure with molecular dynamic (MD) simulation (Fig. 1b). The optimal length of the linker, consisting of glycolate and four ethylene glycol residues, is about 20 Å, that

* Corresponding authors. Tel.: +886 27855696x7101; fax: +886 26539142.
E-mail addresses: jasonhemlot@gmail.com (C.-S. Chen), shwu@gate.sinica.edu.tw (S.-H. Wu).

0040-4039/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2009.08.073

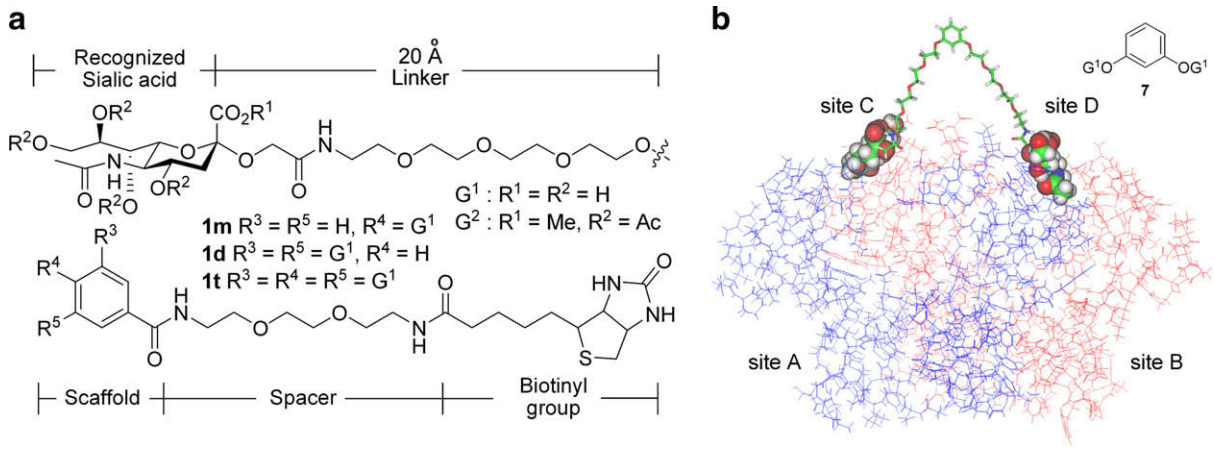


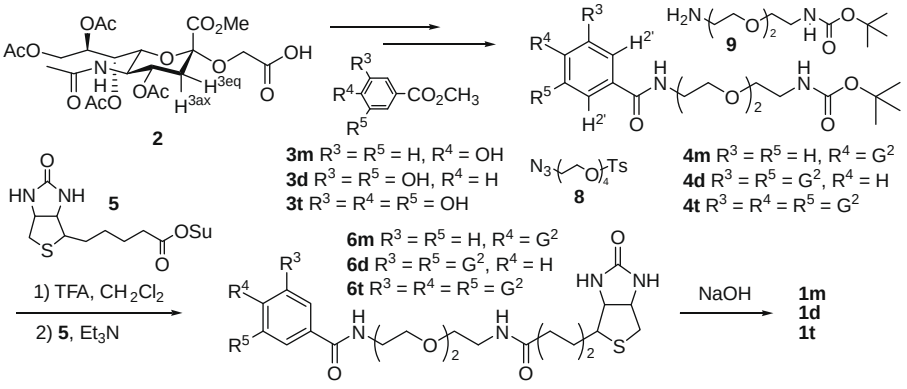
Figure 1. (a) The structures of designed molecules **1m**, **1d**, and **1t**, which contain different number of sialyl residues, linkers, and a scaffold linked to a biotinyl group. (b) Structure of the WGA–**7** complex generated by molecular dynamic simulation. The sites in the grooves of WGA recognized by NeuAc are shown. Bound NeuAc residues in sites C and D are constrained by a force of 100 kcal mol^{-1}, as observed in the crystal structure of the NeuAc–WGA complex.

is, approximately the distance between sites C and D. The simulated scaffold is an *ortho*-linked resorcinol, which can present each vicinal group of the linker in the *anti*-configuration during the simulation and reduce the internal hindrance.

Molecules **1m**, **1d**, and **1t** were synthesized from α-sialyl anomer **2**,[20,21] following the procedures reported previously[22] (Scheme 1). Briefly, the multivalent scaffolds **4m**, **4d**, and **4t** were prepared from the aromatic moieties **3m**, **3d**, and **3t** by extension of the linkers through O-substitution with the azido tosylate **8**, by extension of the spacers through amide bond formation with amine **9**, and by subsequent extension of sialyl residues through amide bond formation with α-sialyl anomer **2**. The sialyl residues in scaffolds **4m**, **4d**, and **4t** were identified by ^{1}H-NMR integrations of H^{3eq} (δ 2.59–2.61 ppm)/$H^{2'}$ (δ 6.90–6.75 ppm), which had relative values of 2/1, 1/1, and 3/2 in **4m**, **4d**, and **4t**, respectively. The *t*-butoxycarbonyls in scaffolds **4m**, **4d**, and **4t** were deprotected with TFA in methylene chloride, followed by amide formation with the active *O*-succinyl biotin **5** to yield biotinylates **6m**, **6d**, and **6t**. O-Deacetylation of the biotinylates **6m**, **6d**, and **6t** in alkaline sodium hydroxide solution afforded the multivalent molecules **1m**, **1d**, and **1t**.

The binding properties of the **1m**–, **1d**–, or **1t**–WGA complexes were examined by SPR. Molecules **1m**, **1d**, and **1t** were immobilized individually on a BIAcore streptavidin-coated chip. The sensorgrams revealed concentration and valence dependence (Fig. 2a and Table 1). The observed association magnitude (10^5) for molecule **1m**–WGA might be an equilibrium value of the monovalent and divalent affinities, as illustrated in Figure 2b. This magnitude is identical to that of immobilized $(GlcNAc)_n$ – WGA (n = 2–5), in which two active sites of WGA might be involved in the recognition.[23] The adjacent biotinyl monovalences of molecule **1m** on the SPR chip could donate two adjacent sialic acid residues for recognition by one WGA lectin and enhance the binding response; this phenomenon is known as the glycol-cluster effect[24,25] or ligand-promoted oligomerization.[26] The association constants of the divalent molecule **1d** and the trivalent molecule **1t** have the same magnitude (10^7); the association rate constants (k_a, 10^4) and dissociation rate constants (k_d, 10^{-3}) of the molecules are also similar. Similar magnitudes of the constants indicate that two sialyl valences of molecules **1** recognize the divalent NeuAc–WGA efficiently (Fig. 2c). This divalent interaction is tighter than the monovalent interaction, largely due to a decrease in dissociation rate.[27] The primary reason is that the polyethylene glycol linker is not a rigid tether, and the interaction between the WGA and molecule **1d**/**1t** will be a sequential, rather than concert process. For a divalent molecule with 20 Å linkers, the binding interaction of one sialyl residue would little affect that of the other sialyl recognition, and this kind of distance-dependent divalent recognition was explored by Bundle and co-workers[12]. The formation of a divalent interaction in WGA would enhance the association constant by decreasing the kinetic dissociation (k_d). As a result, the same magnitude of k_a and divalent dependence of k_d in monovalent and divalent molecules **1** were observed.

The identical association constants of the divalent and trivalent NeuAc-immobilized molecules **1d** and **1t** indicated that two of the three sialyl residues *ortho*- or *meta*-linked to the aromatic scaffold could only provide two valences. The third valence of molecule **1t** is not recognized efficiently or causing apparent steric hindrance.



Scheme 1. Synthesis of the *N*-acetyl-neuraminic-acid-containing molecules **1m**, **1d**, and **1t**.

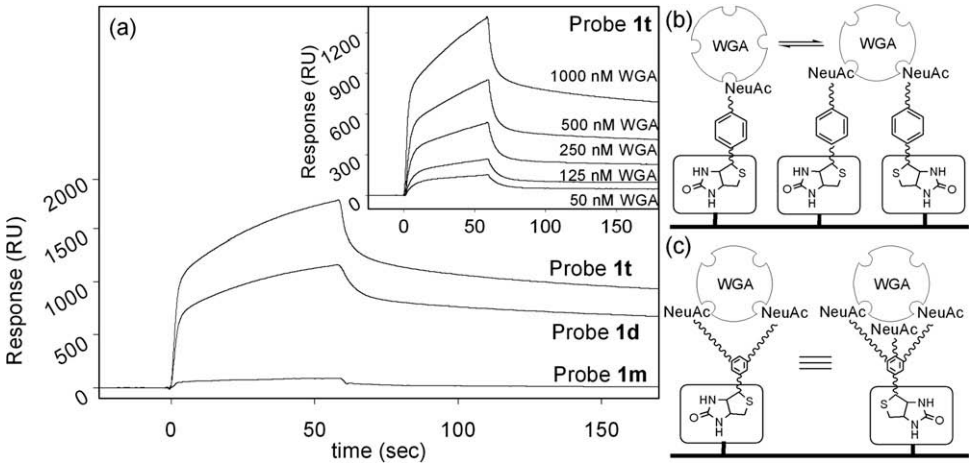


Figure 2. Sialyl recognition of the complex of molecules **1m**, **1d**, or **1t** with WGA. (a) SPR sensorgrams. The interaction of immobilized NeuAc-containing molecules **1m**, **1d**, and **1t** with 1 mM WGA (relative response). Inset: sensorgram showing binding of different concentrations of WGA to the immobilized trivalent molecule **1t**. (b) Illustration representing the recognition of monosialyl molecules **1m**–WGA and (c) the divalent molecules **1d**–WGA and **1t**–WGA.

Table 1
Kinetic parameter for the binding affinity between NeuAc-containing molecules **1** with WGA by surface plasmon resonance[a]

Probe	k_a (M^{-1} s^{-1})	k_d (M^{-1} s^{-1})	K_a (M^{-1})[b]
1m	$1.62 \pm 0.130 \times 10^4$	$2.25 \pm 0.057 \times 10^{-2}$	7.20×10^5
1d	$1.30 \pm 0.003 \times 10^4$	$1.11 \pm 0.059 \times 10^{-3}$	1.17×10^7
1t	$1.63 \pm 0.032 \times 10^4$	$1.26 \pm 0.054 \times 10^{-3}$	1.29×10^7

[a] BIA evaluation version 3.0 software (BIAcore) was applied. Data evaluation from five different concentrations of probes (50, 125, 250, 500, and 1000 nM), and the 1:1 Langmuir binding model was chosen.
[b] Association constant, $K_a = k_a/k_d$.

From the complexation model (Fig. 1b), the third valence could skew from the active recognition sites or from the surface of WGA lectin. Similar divalent association constants have been found for HA and an immobilizing sialyl glycoprotein, which was aimed to mimic the virus/cell affinity.[28] We do not know whether virus/cell adhesion involves divalent recognition or enhancement through the glycol-cluster effect. However, if the sialyl recognition between molecules **1d**/**1t**–WGA is a result of the glycol-cluster effect, then a similar association constant should be found in the recognition of immobilized molecules **1m**–WGA.

The affinity between WGA and the monovalent oligosialic acids, including α-methoxyethoxy-NeuAc and α2,8-$(NeuAc)_n$ ($n = 2–3$), was studied using ITC (see Supplementary data); the titration curves indicated a dilution effect (<0.2 μcal s^{-1}). These molecules contain at least one anomeric α-linkage in the non-reducing end, which is responsible for sialyl recognition. However, the observed isotherms were different from the titration isotherm of $(GlcNAc)_n$ – WGA, which shows a residual dependence.[7] In our opinion, the monovalent recognition of α2,8-$(NeuAc)_n$ – WGA is too weak ($K_a < 10^3$) to be observed in ITC experiments, and the sialyl affinity in WGA lectin could be enhanced by structural divalency.

In summary, we verified the sialyl divalent recognition of WGA lectin. Our study indicated that the structural valency between target lectin and ligand-containing molecule can be predicted if the ligand–receptor complex structure is available in the Protein Data Bank. Structurally, the binding pairs that recognize the sialyl groups are separated into two in WGA (dimer), and sialyl divalency might be the maximum capacity in WGA (dimer). Eventually, a multivalent recognition-combined NeuAc–WGA and GlcNAc–WGA could be achieved with delicate design on the basis of the eight GlcNAc-binding sites and four NeuAc-binding sites in WGA lectin.

Acknowledgment

This work was supported by the National Science Council of Taiwan and Academia Sinica, Taipei, Taiwan.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2009.08.073.

References and notes

1. Lins, R. J.; Flitsch, S. L.; Turner, N. J.; Irving, E.; Brown, S. A. *Tetrahedron* **2004**, *60*, 771–780.
2. Wright, C. S.; Kellogg, G. E. *Protein Sci.* **1996**, *5*, 1466–1476.
3. Wittmann, V.; Seeberger, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 4348–4352.
4. Wittmann, V.; Seeberger, S. *Angew. Chem., Int. Ed.* **2004**, *43*, 900–903.
5. Maierhofer, C.; Rohmer, K.; Wittmann, C. *Bioorg. Med. Chem.* **2007**, *15*, 7661–7676.
6. Muraki, M.; Ishimura, M.; Harata, K. *Biochim. Biophys. Acta* **2002**, *1569*, 10–20.
7. Bains, G.; Lee, R. T.; Lee, Y. C.; Freire, E. *Biochemistry* **1992**, *31*, 12624–12628.
8. Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387–429.
9. James, W. M.; Emerick, M. C.; Agnew, W. S. *Biochemistry* **1989**, *28*, 6001–6009.
10. Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371–386.
11. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669–672.
12. Kitov, P. I.; Shimizu, H.; Homans, S. W.; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 3284–3294.
13. Wright, C. S.; Jaeger, J. *J. Mol. Biol.* **1993**, *232*, 620–638.
14. Wright, C. S. *J. Mol. Biol.* **1990**, *215*, 635–651.
15. Wilson, I. A.; Skehel, J. J.; Wiley, D. C. *Nature* **1981**, *289*, 366–373.
16. Weis, W.; Brown, J. H.; Cusack, S.; Paulson, J. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1988**, *333*, 426–431.
17. Ha, Y.; Stevens, D. J.; Skehel, J. J.; Wiley, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11181–11186.
18. Glick, G. D.; Knowles, J. *J. Am. Chem. Soc.* **1991**, *113*, 4701–4703.
19. Choi, S. K. *Synthetic Multivalent Molecules*; John Wiley & Sons: Hoboken, New Jersey, 2004.
20. Ren, C. T.; Chen, C. S.; Wu, S. H. *J. Org. Chem.* **2002**, *67*, 1376–1379.
21. Ren, C. T.; Chen, C. S.; Yu, Y. P.; Tsai, Y. F.; Lin, P. Y.; Chen, Y. J.; Zou, W.; Wu, S. H. *Chem. Eur. J.* **2003**, *9*, 1085–1095.
22. Roy, R.; Park, W. K. C.; Wu, Q. Q.; Wang, S. N. *Tetrahedron Lett.* **1995**, *36*, 4377–4380.
23. Lienemann, M.; Paananen, A.; Boer, H.; Fuente, J. M. d. l.; Garcia, I.; Penades, S.; Koivula, A. *Glycobiology* **2009**, *19*, 633–643.
24. Lee, Y. C. *Carbohydr. Res.* **1978**, *67*, 509–514.
25. Miura, Y.; Yamauchi, T.; Sato, H.; Fukuda, T. *Thin Solid Films* **2008**, *516*, 2443–2449.
26. Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555–578.
27. Kramer, R. H.; Karpen, J. W. *Nature* **1998**, *395*, 710–713.
28. Takemoto, D. K.; Skehel, J. J.; Wiley, D. C. *Virology* **1996**, *217*, 452–458.