Bioorganic Chemistry

DOI: 10.1002/anie.200903328

DNA-Templated Homo- and Heterodimerization of Peptide Nucleic Acid Encoded Oligosaccharides that Mimick the Carbohydrate Epitope of HIV**

*Katarzyna Gorska, Kuo-Ting Huang, Olivier Chaloin, and Nicolas Winssinger**

Dedicated to Professor Jean-Marie Lehn on the occasion of his 70th birthday

Homo- and heterooligomerization of receptors is a fundamental feature of cellular recognition and signal transduction.[1–3] Protein–carbohydrate interactions often achieve high avidity by the cooperative interaction of multiple units.[4] An example of therapeutic significance is the interaction of the cholera toxin with cell-surface carbohydrates, for which high avidity was achieved (10^6-fold enhancement over the monomer) by using a synthetic oligomer with an architecture matching the geometry of the toxin receptor.[5] While this example illustrates the importance of accessing oligomeric carbohydrate structures with controlled topology, the synthetic challenges of accessing complex architectures have hampered progress. Herein we report a simple method to obtain oligomeric carbohydrates with controlled topology in a combinatorial fashion by the self-assembly of oligosaccharides tagged with peptide nucleic acids (PNAs) onto DNA templates. The potential of this method is demonstrated with different supramolecular architectures which mimic multiple copies of the carbohydrates (high mannose nonasaccharide, **1**, Figure 1) that shield gp120 and are known to interact in a multivalent fashion with 2G12, an antibody that shows broad-spectrum activity against HIV.[6–8]

The crystal structure of 2G12 with the high-mannose sugar **1** showed that 2G12 assembled into an interlocked dimer, thereby resulting in two additional binding sites at its dimerization interface (Figure 2).[9,10] This observation not only provided a rationale for the high affinity of the antibody for its target (HIV's glycoprotein 120–gp120) by virtue of the highly cooperative binding mode, but also for its selectivity for gp120 bearing high-mannose carbohydrates compared to the host. Indeed, the 2G12 antibody displays appropriately

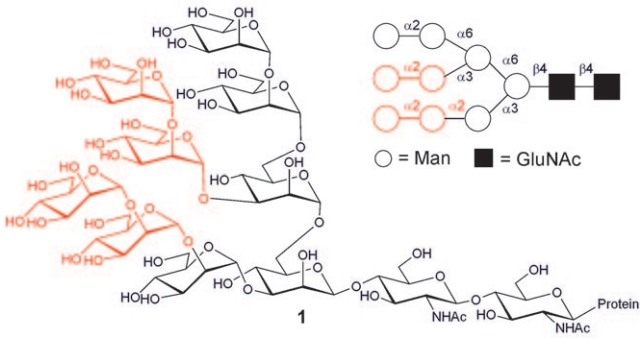


Figure 1. Structure of the high-manose nonasaccharide present on gp120. Carbohydrate units shown in red interact directly with 2G12.[9]

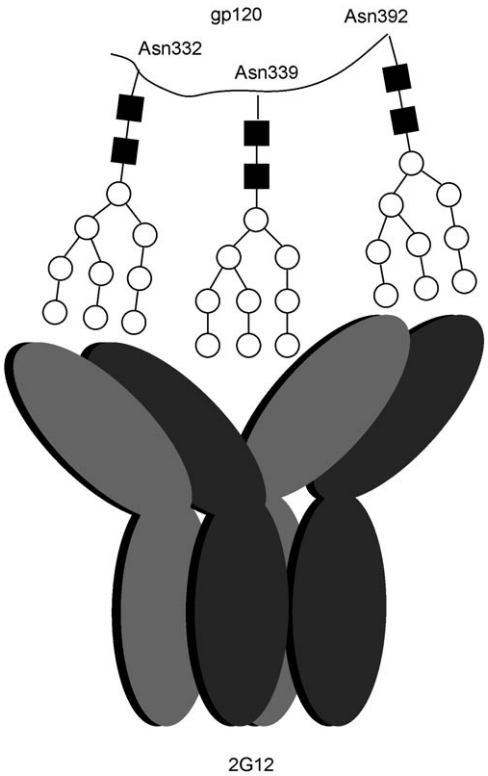


Figure 2. Proposed binding mode of 2G12 dimer with gp120.[9]

[*] K. Gorska, K.-T. Huang, Dr. O. Chaloin, Prof. N. Winssinger
Institut de Science et Ingénierie Supramoléculaires (ISIS)
Université de Strasbourg—CNRS (UMR 7006)
8 allée Gaspard Monge, 67000 Strasbourg (France)
Fax: (+33) 3-9024-5112
E-mail: winssinger@isis.u-strasbg.fr
Homepage: http://www-isis.u-strasbg.fr/winssinger/

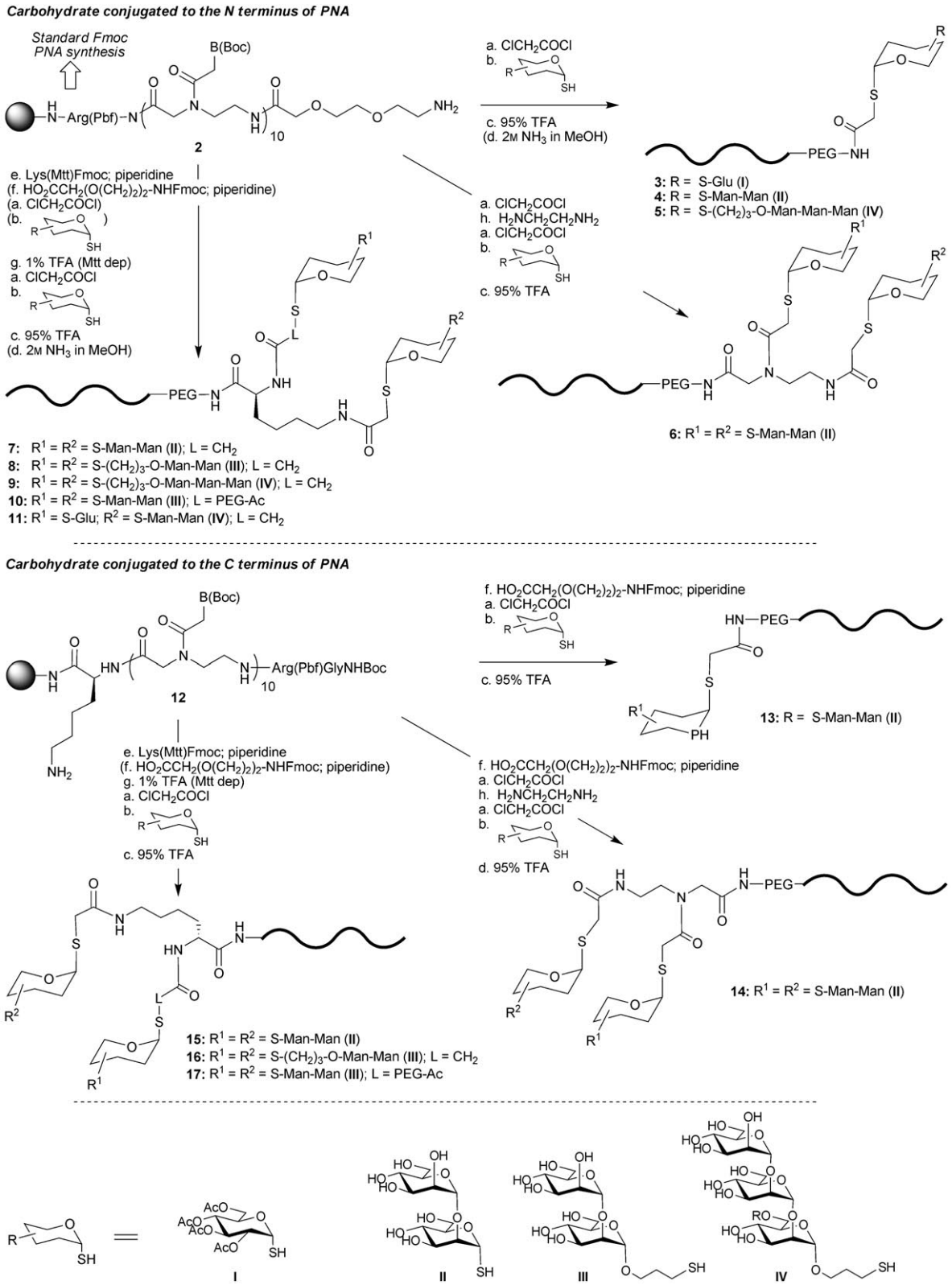
[**] K.G. and K.-T.H. contributed equally to this work. This work was supported by a grant from the European Research Council (ERC 201749). A fellowship from Boehringer Inglheim Fonds (to K.G.) and the French Ministry of Research and Education (K.-T.H.) is gratefully acknowledged. 2G12 was obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH; and the HIV-1 gp120 monoclonal antibody (2G12) from Dr. Hermann Katinger. We thank Laurence Oswald for her technical assistance.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200903328.

spaced binding sites that match the spacing of these structures on the viral surface. Interestingly, the crystallographic information suggests that only the terminal mannose moieties shown in red (Figure 1) are involved in the interaction with the antibody. Corroborating these results, it has been shown that carbohydrate **1** or fragments thereof bearing the terminal mannose presented at high density on a surface have a higher affinity for 2G12 than their monomeric counterparts.[9,10] While the nonasaccharide **1** has no notable affinity for 2G12, Danishefsky and co-workers have shown that a trimer of **1** displayed on a rigid scaffold[11] binds 2G12 with moderate affinity,[12] while Wang et al. showed that a tetramer of a mannose tetrasaccharide on a similar scaffold also binds 2G12 with micromolar affinity.[13] The distance between the two primary binding sites in 2G12 is 30 Å (measured from PDB 1OP5[9]), and while some level of cooperativity has been achieved with previously reported oligomers, a more systematic investigation of the optimal spacing geometry between the ligands to maximize the cooperativity has not been reported. In fact Danishefsky and co-workers have shown that the oligomer conjugated to an immunogenic protein is able to elicit an antibody response that recognizes the oligomer of **1** but fails to neutralize HIV, thereby suggesting that it is not an optimal mimic of the epitope of gp120.[14] Dendrons of **1** have also been shown to bind 2G12 cooperatively.[15]

Peptide nucleic acids[16,17] are attractive tags for programming self-assembled structures, since their chemistry is significantly more permissive than that of natural oligonucleotides. Furthermore, the higher affinity of PNAs for natural oligonucleotides allows for shorter tags, which are more specific and less sensitive to the ionic strength of the solution.[18] For example, we have shown that they could be used to encode combinatorial libraries of peptides.[19–22] In fact, only monosaccharide–DNA conjugates have been reported thus far for microarraying[23] and to study lectin interactions,[24,25] whereas larger oligosaccharides with more complex branching patterns have never been reported. As shown in Scheme 1, carbohydrates can be efficiently coupled to polymer-bound PNAs at the C terminus or the N terminus by coupling a thiol (thioacetals or carbohydrates bearing a thioalkyl group at the anomeric position) to a chloroacetamide with a mild base (such as Hunig′s base or DBU). For the purpose at hand, a 10 mer PNA was deemed appropriate as it would provide a melting temperature (T_m) of greater than 50 °C and would present adjacent ligands on the same face of the helix. Thus 10 mer PNA **2** obtained by standard Fmoc chemistry and bearing a short polyethyleneglycol (PEG) spacer (10 Å) at the N terminus was coupled to chloroacetyl chloride and treated with commercially available tetraacetyl glucothiolactol **I** (Scheme 1) in the presence of Hünig's base. Analysis of the cleavage product by LC–MS and MALDI MS indicated complete conversion and showed a single signal for product **3**. The acetyl groups on the glucose moiety were removed by treatment with ammonia in MeOH, which was found to be equally efficient prior to or after cleavage from the resin with TFA. The same procedure was applied for the coupling of unprotected disaccharide **II**[26] to afford conjugate **4** after cleavage. To our gratification, no trace of glycosidic cleavage was observed upon treatment with acid. Saccharides labeled with a thioalkyl group at the anomeric position[26] could also be efficiently coupled to the chloroacetamide-substituted PNA, but using DBU rather than Hünig's base, thus affording conjugate **5** after cleavage. Starting from PNA **2**, a similar strategy was used to install two units of the carbohydrate by treatment of the chloroacetamide product with ethylenediamine. The resulting PNA–diamine conjugate was then treated with chloroacetyl chloride followed by a thiosaccharide to obtain, after cleavage, product **6**, which bears two copies of the carbohydrate. Assuming a *trans* or *anti* conformation for all the bonds, the distance between the anomeric centers in the two carbohydrate units is 11.5 Å.[27] Conversely, PNA **2** can be coupled to an orthogonally protected lysine (Fmoc, Mtt). Cleavage of both the Fmoc and Mtt[28] protecting groups followed by chloroacetylation and coupling with thioglycosides **II** and **III** affords products **7** and **8**, respectively, in which the same carbohydrate units are separate by a maximum distance of 15 and 19 Å repectively. To obtain a greater distance between the carbohydrate units, the Fmoc group can be selectively removed and a 10 Å PEG spacer added prior to the chloroacetylations to obtain **10** after glycoside conjugation and release from the resin. Finally, to obtain different carbohydrate units on each side of the lysine residue, a first chloroacetylation/carbohydrate coupling is performed after cleavage of the Fmoc group but prior to removal of the Mtt group. A second chloroacetylation/carbohydrate coupling can then be performed after cleavage of the Mtt group to obtain heterodimeric conjugates such as **11**. Similar sequences were utilized to conjugate carbohydrates at the C teminus of the PNA, starting from resin **12**, to afford conjugates **13**–**17** bearing a single unit of α-1,2-mannose dimers or two units separated by distances ranging from 11.5 to 25 Å. The PNA–carbohydrate conjugates were then assembled into dimers and oligomers by hybridization to the appropriate DNA template. Considering the high persistent length of the double helix and the fact that all the PNA sequences are 10 mers, it can be anticipated that carbohydrate units in architectures such as those in entries 2, 3, and 15 of Figure 3 will be separated by 30, 60, and 90 Å respectively. However, the inclusion of a PEG linker between the PNA and the carbohydrate allows for some degree of flexibility.

A pilot library of over 30 architectures was tested for their affinity to 2G12 by surface plasmon resonance (SPR). The antibody was immobilized accordingly to a previously described protocol.[12] Under these conditions, no notable binding was observed for nonasaccharide **1**, as previously reported by Danishefsky and co-workers,[12,29] and in agreement with the binding mode reported by Wilson[9,10] which involves four units of mannose disaccharide from multiple units of **1**. Significant binding (μM) was observed for conjugates having the key α-1,2-mannose disaccharide units (Figure 4). However, the distance between the two carbohydrate units of this disaccharide was critical for the binding, thus attesting to the importance of their cooperativity for avidity. Only conjugates bearing an 11 atom spacer between the two disaccharide units (structure **7** and **15**) displayed any binding (entries 12–18, Figure 3). It is interesting to note that the distance between the two mannose units involved in the



Scheme 1. Synthesis of PNA–oligosaccharide conjugates. a) $ClCH_2COCl$, 2,6-lutidine, DMF; b) saccharide (10 equiv), for **I** or **II**: Et*i*Pr$_2$N (15 equiv) or DBU (10 equiv) for **III** and **IV**, DMF; c) TFA/H_2O 95:5; d) 2 M H_3N in MeOH; e) HO-Lys(Mtt)Fmoc, Et*i*Pr$_2$N, HCTU, DMF; 20% piperidine in DMF; f) $HO_2CCH_2(O(CH_2)_2)_2NH_2$, Et*i*Pr$_2$N, HCTU, DMF; 20% piperidine in DMF; g) ethylenediamine (2×10 equiv), DMF. DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene, TFA = trifluoroacetic acid, Mtt = 4-methyltrityl, Fmoc = 9-fluorenylmethoxycarbonyl, HCTU = 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate, Boc = *tert*-butoxycarbonyl, dep = deprotection.

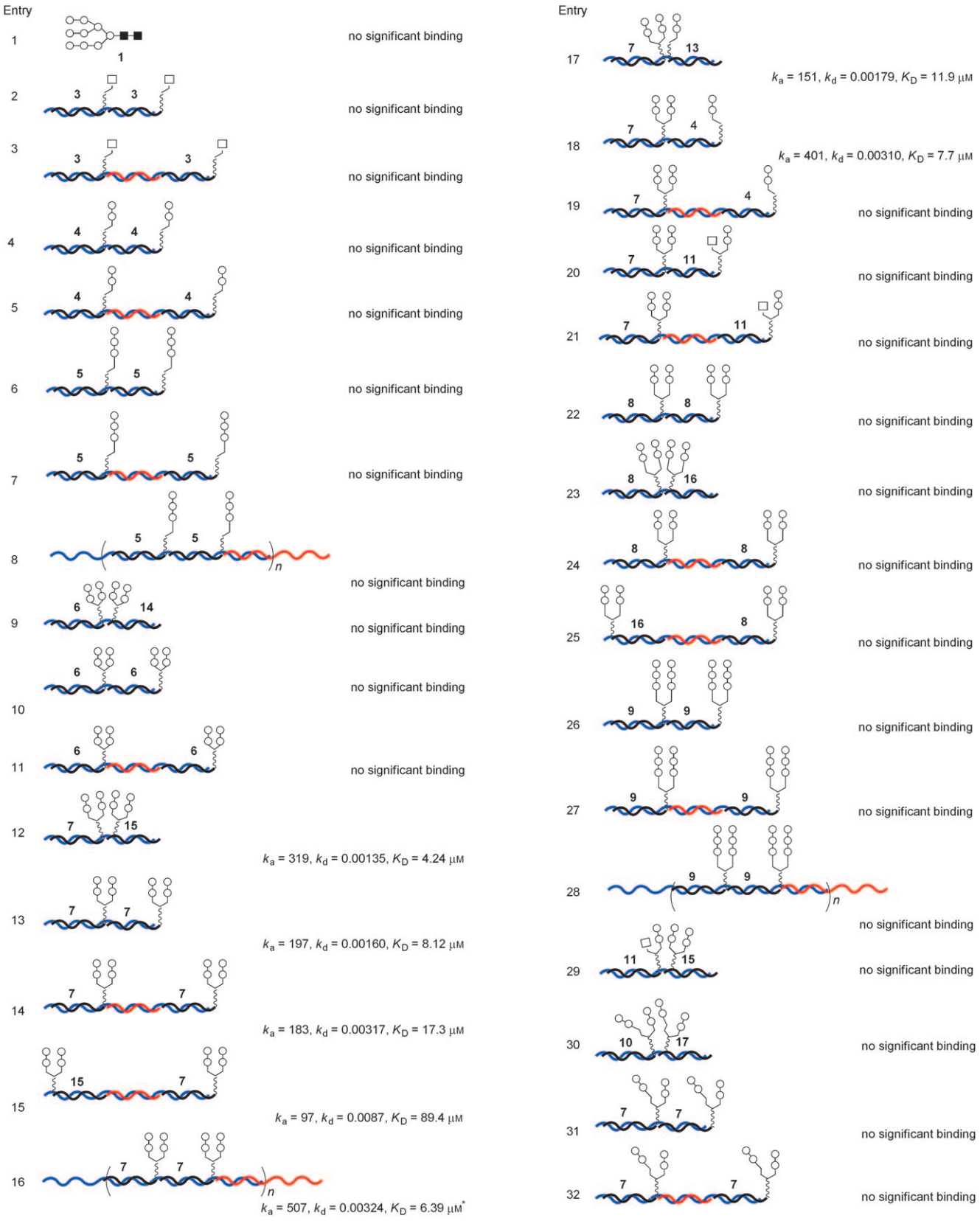


Figure 3. Affinity of supramolecular complexes to 2G12 measured by SPR (k_a (M^{-1} s^{-1}) and k_d (s^{-1})). Solutions were prepared by mixing 1 equivalent of each PNA with their respective template. Binding kinetics were obtained using a Langmuir 1:1 model ($R_{max}/\chi^2 > 50$). * denotes the concentration of the template.

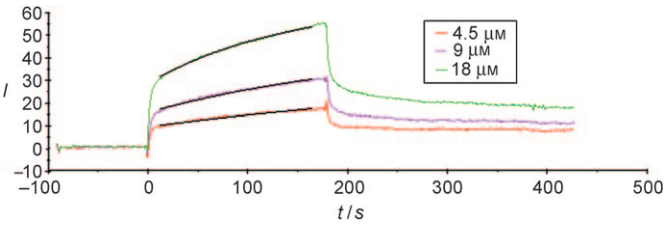


Figure 4. Measurment of the affinity for the architecture shown in entry 12 of Figure 3 by SPR (green: 18 μM, magenta: 9 μM, red: 4.5 μM).

binding of nonasaccharide **1** is also 11 atoms. Since the through-space distance between the anomeric substituents is 10 Å (measured from PDB 1OP5), the shorter spacer present in structures **6** and **14** clearly do not adequately replicate this geometry (entries 9–11), whereas longer spacers in structures **8**, **10**, **16**, and **17** (entry 22–25, 30–32; Figure 3) do not provide the adequate level of cooperativity. Following the same argument, architectures based on the mannose trisaccharide fail to provide significant binding (entries 26–28; Figure 3). The topology of the supramolecular architecture also has a significant impact on binding. While the PEG spacer between the carbohydrate moiety and the PNA does provide a certain level of flexibility, a clear trend emerges on comparison of entries 12 to 15 of Figure 3, with the shortest distance (entry 12) being best. Considering that a distance of about 30 Å is required for cooperative binding to 2G12, the architecture in entry 12 (the maximum distance between the branch points joining the carbohydrate units is 38.5 Å) would be most suitable, whereas the architecture of entry 15 shows a lower level of cooperative binding. The architecture in entry 16 has complimentary "sticky ends" and is anticipated to polymerize extensively.[30] The value of 6.39 μM reported in Figure 3 is for the DNA template, and thus underestimates the affinity of the oligomeric polymer.

The novelty of the approach reported herein is that it exploits the programmability of hybridization to generate a library of architectures which emulate the topologies of complex carbohydrates. The gp120 epitope has stimulated tremendous efforts towards the production of vaccines.[12,15,31–33] While this example illustrates the importance of multimeric recognition with controlled topology, the generality of this concept extends far beyond pathogen recognition[34] since it is a recurrent feature in cellular recognition and communication, with many signal transduction pathways also being regulated by multimeric interactions.[2] Better insight into the mechanisms that govern these two hierarchical levels requires methods to prepare well-organized oligomers with controlled topology. Last but not least, these principles should be applicable beyond carbohydrate recognition to generate small molecules or drug fragments in a combinatorial fashion and provide an encoding strategy for dynamic combinatorial libraries.[35,36]

Received: June 19, 2009

Keywords: bioorganic chemistry · carbohydrates · DNA · HIV · peptide nucleic acids

[1] J. P. Pin, R. Neubig, M. Bouvier, L. Devi, M. Filizola, J. A. Javitch, M. J. Lohse, G. Milligan, K. Palczewski, M. Parmentier, M. Spedding, *Pharmacol. Rev.* **2007**, *59*, 5–13.
[2] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew. Chem.* **2006**, *118*, 2408–2429; *Angew. Chem. Int. Ed.* **2006**, *45*, 2348–2368.
[3] M. Mammen, S. Choi, G. Whitesides, *Angew. Chem.* **1998**, *110*, 2908–2953; *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.
[4] J. Lundquist, E. Toone, *Chem. Rev.* **2002**, *102*, 555–578.
[5] P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Armstrong, H. Ling, N. S. Pannu, R. J. Read, D. R. Bundle, *Nature* **2000**, *403*, 669–672.
[6] J. R. Mascola, M. G. Lewis, G. Stiegler, D. Harris, T. C. VanCott, D. Hayes, M. K. Louder, C. R. Brown, C. V. Sapan, S. S. Frankel, Y. Lu, M. L. Robb, H. Katinger, D. L. Birx, *J. Virol.* **1999**, *73*, 4009–4018.
[7] B. Etemad-Moghadam, Y. Sun, E. K. Nicholson, G. B. Karlsson, D. Schenten, J. Sodroski, *J. Virol.* **1999**, *73*, 8873–8879.
[8] J. M. Crawford, P. L. Earl, B. Moss, K. A. Reimann, M. S. Wyand, K. H. Manson, M. Bilska, J. T. Zhou, C. D. Pauza, P. W. Parren, D. R. Burton, J. G. Sodroski, N. L. Letvin, D. C. Montefiori, *J. Virol.* **1999**, *73*, 10199–10207.
[9] D. A. Calarese, C. N. Scanlan, M. B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M. R. Wormald, R. L. Stanfield, K. H. Roux, J. W. Kelly, P. M. Rudd, R. A. Dwek, H. Katinger, D. R. Burton, I. A. Wilson, *Science* **2003**, *300*, 2065–2071.
[10] D. A. Calarese, H. K. Lee, C. Y. Huang, M. D. Best, R. D. Astronomo, R. L. Stanfield, H. Katinger, D. R. Burton, C. H. Wong, I. A. Wilson, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13372–13377.
[11] O. Renaudet, P. Dumy, *Org. Lett.* **2003**, *5*, 243–246.
[12] I. J. Krauss, J. G. Joyce, A. C. Finnefrock, H. C. Song, V. Y. Dudkin, X. Geng, J. D. Warren, M. Chastain, J. W. Shiver, S. J. Danishefsky, *J. Am. Chem. Soc.* **2007**, *129*, 11042–11044.
[13] J. Wang, H. Li, G. Zou, L. Wang, *Org. Biomol. Chem.* **2007**, *5*, 1529–1540.
[14] J. G. Joyce, I. J. Krauss, H. C. Song, D. W. Opalka, K. M. Grimm, D. D. Nahas, M. T. Esser, R. Hrin, M. Feng, V. Y. Dudkin, M. Chastain, J. W. Shiver, S. J. Danishefsky, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15684–15689.
[15] S. K. Wang, P. H. Liang, R. D. Astronomo, T. L. Hsu, S. L. Hsieh, D. R. Burton, C. H. Wong, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3690–3695.
[16] P. E. Nielsen, *Mol. Biotechnol.* **2004**, *26*, 233–248.
[17] M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–568.
[18] Z. L. Pianowski, N. Winssinger, *Chem. Soc. Rev.* **2008**, *37*, 1330–1336.
[19] F. Debaene, J. Da Silva, Z. Pianowski, F. Duran, N. Winssinger, *Tetrahedron* **2007**, *63*, 6577–6586.
[20] J. L. Harris, N. Winssinger, *Chem. Eur. J.* **2005**, *11*, 6792–6801.
[21] J. Harris, D. E. Mason, J. Li, K. W. Burdick, B. J. Backes, T. Chen, A. Shipway, G. Van Heeke, L. Gough, A. Ghaemmaghami, F. Shakib, F. Debaene, N. Winssinger, *Chem. Biol.* **2004**, *11*, 1361–1372.
[22] N. Winssinger, R. Damoiseaux, D. C. Tully, B. H. Geierstanger, K. Burdick, J. L. Harris, *Chem. Biol.* **2004**, *11*, 1351–1360.
[23] Y. Chevolot, C. Bouillon, S. Vidal, F. Morvan, A. Meyer, J. P. Cloarec, A. Jochum, J. P. Praly, J. J. Vasseur, E. Souteyrand, *Angew. Chem.* **2007**, *119*, 2450–2454; *Angew. Chem. Int. Ed.* **2007**, *46*, 2398–2402.
[24] K. Matsuura, M. Hibino, Y. Yamada, K. Kobayashi, *J. Am Chem. Soc.* **2001**, *123*, 357–358.
[25] K. Matsuura, M. Hibino, T. Ikeda, Y. Yamada, K. Kobayashi, *Chem. Eur. J.* **2004**, *10*, 352–359.

[26] See the Supporting Information for synthetic procedures.
[27] Calculated assuming a 2.5 Å through-space distance between three atoms in an *anti* configuration and 1.5 Å for a single bond.
[28] S. Pothukanuri, Z. Pianowski, N. Winssinger, *Eur. J. Org. Chem.* **2008**, 3141 – 3148.
[29] V. Y. Dudkin, M. Orlova, X. Geng, M. Mandal, W. C. Olson, S. J. Danishefsky, *J. Am. Chem. Soc.* **2004**, *126*, 9560 – 9562.
[30] For a recent review on glycodendrimers, see Y. M. Chabre, R. Roy, *Curr. Top. Med. Chem.* **2008**, *8*, 1237 – 1285.
[31] X. Geng, V. Y. Dudkin, M. Mandal, S. J. Danishefsky, *Angew. Chem.* **2004**, *116*, 2616 – 2619; *Angew. Chem. Int. Ed.* **2004**, *43*, 2562 – 2565.
[32] J. D. Warren, X. Geng, S. J. Danishefsky, *Top. Curr. Chem.* **2007**, *267*, 109 – 141.
[33] S. Sirois, M. Touaibia, K. C. Chou, R. Roy, *Curr. Med. Chem.* **2007**, *14*, 3232 – 3242.
[34] A. Imberty, Y. M. Chabre, R. Roy, *Chem. Eur. J.* **2008**, *14*, 7490 – 7499.
[35] J. M. Lehn, *Science* **2002**, *295*, 2400 – 2403.
[36] J. M. Lehn, *Chem. Eur. J.* **1999**, *5*, 2455 – 2463.

