

## PNA–sugar conjugates as tools for the spatial screening of carbohydrate–lectin interactions\*

Christian Scheibe and Oliver Seitz‡

*Department of Chemistry, Humboldt-Universität zu Berlin, Brook-Taylor-Str. 2,  
12489 Berlin, Germany*

**Abstract:** Multivalent carbohydrate–lectin interactions are essential for a multitude of biological recognition events. Much effort has been spent in the synthesis of potent multivalent scaffolds in order to mimic or inhibit biological carbohydrate–protein interactions. However, the defined spatial presentation of carbohydrates remained a challenging task. Peptide nucleic acid (PNA)- and DNA-based double helices are useful scaffolds that enable the controlled display of carbohydrate ligands in a modular approach. The hybridization of PNA-sugar conjugates with complementary DNA strands provides multivalent complexes with defined spatial presentation of carbohydrates, which facilitates the spatial screening of carbohydrate–lectin interactions with Ångström-scale precision.

**Keywords:** carbohydrates; DNA; lectin; multivalency; peptide nucleic acid; self-assembly; spatial screening; templates.

### INTRODUCTION

Within the diversity of interactions occurring in nature, the interaction of a carbohydrate ligand with its corresponding protein is of relatively weak affinity. Typical dissociation constants of monosaccharides are in the millimolar range. Still, carbohydrate–lectin interactions play an essential role in a myriad of biological recognition events [1,2]. Nature compensates the lack of strong affinity by presenting carbohydrate ligands in a multivalent fashion. Remarkably, the affinity enhancement obtained thereby substantially exceeds the statistical effect, i.e., the effect resulting from the increased local sugar concentration. Despite the numerous observations of this so-called “cluster glycoside effect” [3], scientists only slowly unravel its physical origin [4]. One requisite for the identification of optimal molecular features for enhanced binding affinity is the use of a scaffold that allows the exact adjustment of (1) the number of ligands, (2) the distance between the ligands, and (3) the spatial orientation of the ligands [5,6]. In the past years, nucleic acid architectures have been recognized to perfectly fulfill these requirements [7]. This article outlines the development of peptide nucleic acid (PNA)-sugar conjugates becoming a valuable tool for the investigation of carbohydrate–lectin interactions. The characteristics of nucleic acid architectures, which make the PNA/DNA duplex such a powerful scaffold, are shown. Furthermore, an overview of the studies reported in the literature will be given.

---

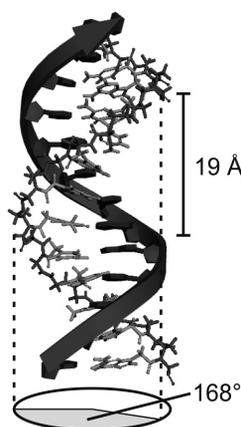
\**Pure Appl. Chem.* **84**, 1–106 (2012). A collection of invited papers based on presentations at the 16<sup>th</sup> European Carbohydrate Symposium (Eurocarb-16), Sorrento, Italy, 3–7 July 2011.

‡Corresponding author

## THE PNA/DNA DOUBLE HELIX

The hybridization of a nucleic acid molecule with its complementary counterpart follows well-known base-pairing rules and produces a double-helical structure with defined geometry. Inspired by this simple—yet powerful—assembly system, scientists did not hesitate to mimic or even optimize it. One of the most astonishing discoveries was made in 1991, when Nielsen et al. replaced the entire deoxyribose-phosphate backbone with a pseudopeptide chain consisting of *N*-(2-aminoethyl)glycine units [8]. Though this PNA mimics the behavior of DNA, it has some unique physical, chemical, and biological properties that make it a powerful biomolecular tool [9]. Unlike DNA, oligomerization of PNA relies on the simple formation of a peptide bond, making it more resistant to strong acids and bases and hence, making tagging or co-synthesis of small or macromolecules more permissive. Since the PNA backbone carries no charges, hybrids with complementary DNA show higher thermal stability, which helps to maintain the integrity of PNA/DNA duplexes at low concentrations. Furthermore, PNA is not susceptible to enzymatic degradation by nucleases which would facilitate studies in biological environments.

Structural analysis using solution-NMR revealed that the conformation of the PNA/DNA duplex is a mixture of A- and B-form DNA [10]. With an average helical twist per base pair of  $28^\circ$ , one helical turn comprises 13 base pairs. The diameter of the helix is  $23 \text{ \AA}$  with a pitch of  $42 \text{ \AA}$ . Thus, the rise per base pair is approximately  $3.2 \text{ \AA}$ . Even though the structure of the PNA/DNA duplex is specified, different approaches are used in the literature to determine the distance between two ligands presented along the double helix. For the sake of comparison we will calculate all distances based on the structure in the protein database (1pdt.pdb, see Fig. 1). The rise per base pair solely determines the distance, i.e., distances arising from the helical torsion are disregarded. Distances from linkers are also not taken into account. For example, the distance between two ligands that are separated by five nucleotides is ca.  $19 \text{ \AA}$  ( $6 \times 3.2 \text{ \AA}$ ). The helical torsion between the two ligands is ca.  $168^\circ$  ( $6 \times 28^\circ$ ).



**Fig. 1** Calculation of the distance and torsion angle in the PNA/DNA duplex for two nucleotides that are separated by five nucleotides. Structure based on 1pdt.pdb [10] (black = DNA strand, gray = PNA strand).

It should be noted that this estimation is only true for straight double helices. For example, nick sites within the helix can serve as joints that may facilitate kinking and rotating [11]. This affects in particular the torsion angle but also the distance (see below for further details). Notwithstanding, several other factors have an influence on the distance and torsion angle. Obviously, the linker connecting the carbohydrate ligand and the double helix plays a major role, i.e., the accuracy of the ligand's position decreases with increasing length and flexibility of the linker. But also the position of the carbohydrate ligand within the oligonucleotide has an influence. Due to end-fraying, the position of terminally

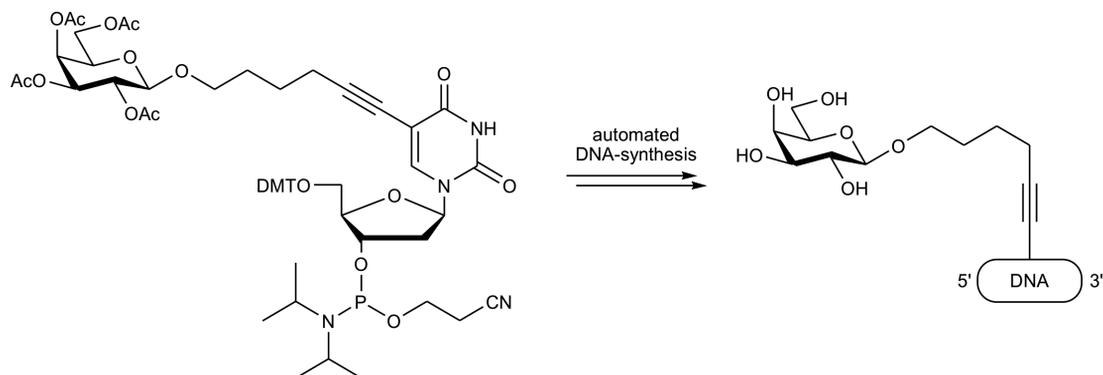
attached ligands is less defined. Hence, more accurate distances and angles can be determined for internally attached ligands.

In principle, the DNA/DNA duplex is also well suited for the exact presentation of carbohydrates. The rise and torsion per base pair in the DNA/DNA duplex is different compared to the PNA/DNA duplex. In the B-form DNA, one helical turn comprises 10.5 base pairs with a pitch of 33 Å, i.e., the distance and helical torsion between each base pair is 3.3 Å and 36°, respectively.

## DNA-PROGRAMMED PRESENTATION OF CARBOHYDRATES

### DNA-sugar conjugates

Kobayashi was the first who used self-organizing oligonucleotides for the construction of supramolecular glycoclusters [12–14]. This pioneering work was performed by using DNA/DNA duplexes as scaffold. Site-specifically galactosylated oligonucleotides were obtained via automated DNA synthesis employing galactose (Gal)-modified DNA monomers (Fig. 2). Gal-DNA-oligomers with 18, 20, and 22 nucleotides in length were synthesized. In each case, the complementary DNA strand had the same length. The DNA sequences were designed in such a way that they offered sticky ends, i.e., the left half-sequence of the Gal-DNA was complementary to the right half-sequence of the corresponding unmodified DNA and vice versa. Hybridization afforded the Gal-DNA/DNA-duplexes **1–3** in which the distance and the angle between the presented Gal ligands varied (see Table 1). Complexes **2** and **3** displayed the Gal residues in 63 and 70 Å intervals, respectively, with a helical twist of 34° each. In contrast, complex **1** displayed the Gal residues in 57 Å intervals. The helical twist between two adjacent ligands was 103°.



**Fig. 2** Internally Gal-labeled DNA oligomers.

The binding affinities of the Gal clusters were estimated by fluorometry, i.e., the decrease in fluorescence intensity was determined upon addition of Gal-complex to fluorescein isothiocyanate (FITC)-labeled RCA<sub>120</sub>. RCA<sub>120</sub> is a tetrameric lectin with two binding sites separated by 120–130 Å [15]. The strongest binding was observed for complex **2** with a Hill coefficient of 2.4, indicating cooperative binding. For complex **1**, the affinity constant and the Hill coefficient were smaller. The authors explained this decrease in binding affinity by considering the dihedral angle of 103°, which was deemed unfavorable for cooperative binding to RCA<sub>120</sub>. However, it should be kept in mind that the investigated DNA/DNA duplexes possessed nick sites that make an exact determination of the helical angle rather difficult. Surprisingly, complex **3**, which also displayed the galactose ligands with a torsional angle of 34°, did not show significant binding. This supports the idea that the distance between the carbohydrate ligands is also crucial for optimal binding to RCA<sub>120</sub>.

**Table 1** Binding parameters for the recognition of Gal-DNA/DNA complexes **1–6** by RCA<sub>120</sub> (determined via fluorometry;  $K_{af}$ , apparent affinity constant;  $n$ , Hill coefficient, , ,  DNA strand,  DNA strand, **0** Gal-ligand) [13,14].

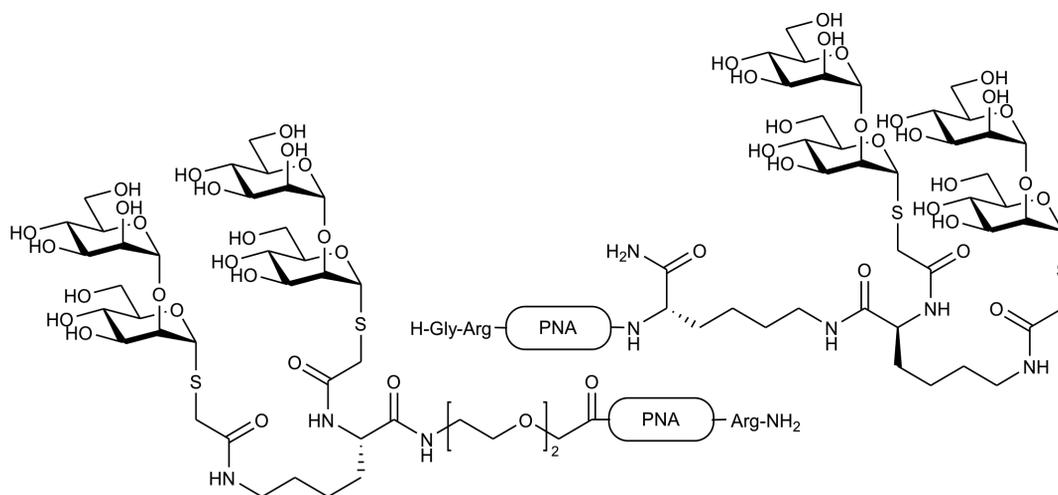
Complex	Distance/Å	Torsion/°	$K_{af}/M^{-1}$	$n$
1	57	103	$1.9 \times 10^4$	1.8
2	63	34	$5.5 \times 10^4$	2.4
3	70	34	–	–
4	63	34	–	–
5	63	34	–	–
6	63	34	$1.1 \times 10^4$	1.0

In another study, the binding affinities of complexes **4–6** were determined (see Table 1) [14]. Whereas complexes **1–3**, obtained by alternate hybridization, presented 6.5 galactose ligands on average, complexes **4–6** had a fixed length and presented 1, 2, or 3 galactose ligands. Since the complexes bound generally weak to RCA<sub>120</sub>, an affinity constant could only be determined for complex **6**. Though this was not proven, the authors assumed that the fully hybridized complexes are less flexible than the alternately hybridized complexes and, therefore, are less capable of cooperative binding to RCA<sub>120</sub>.

### PNA-sugar conjugates

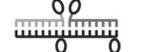
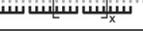
Wingsinger reported the synthesis of terminally mannosylated PNA-oligomers, which were used in hybridization experiments involving DNA template strands for the generation of oligomeric carbohydrates with controlled topology [16]. In contrast to Kobayashi, the carbohydrate ligands were post-synthetically coupled to the PNA strand, i.e., after automated solid-phase PNA synthesis. For that purpose, different thio sugars were coupled to the N- or C-terminus of chloroacetamide-substituted PNA. The obtained mannose (Man)-PNA strands were hybridized with different DNA templates yielding bivalent Man-PNA/DNA duplexes. The distance between the Man ligands was altered by variation of the DNA template sequence. If this resulted in single-strand regions in the duplex, a spacer PNA strand was added to maintain the helical structure. In addition to this sequence-controlled distance variation, the distance between different Man units within one carbohydrate ligand was also varied, e.g., by introduction of short poly(ethylene glycol) chains. This methodology allowed both short- and long-distance variation between the different Man ligands.

With the aid of surface plasmon resonance (SPR), the affinity of the different Man complexes toward surface-bound 2G12, an antibody that shows broad-spectrum activity against HIV, was determined. The crystal structure of 2G12 with the oligosaccharide Man<sub>9</sub>GlcNAc<sub>2</sub> showed that 2G12 forms an interlocked dimer resulting in three binding sites for high-Man carbohydrates [17]. Among the multitude of multivalent Man complexes, only complexes carrying tetramannose ligands in which the two dimannose units were separated by an 11-atom spacer showed significant binding (see Fig. 3). Referring to the crystal structure of 2G12, the authors stated that only this Man ligand adequately replicated the geometry of the binding site of 2G12. On closer examination, the best binding was observed for the complex with the shortest distance between the tetramannose ligands, i.e., the complex in which the distance almost exclusively arose from the linkers connecting them to the PNA backbone (see Table 2). In fact, a sequence-controlled distance increase resulted in decreased binding affinity. Since no monovalent complexes were tested for their binding affinity, an unambiguous conclusion whether the high binding affinity of complex **7** is due to optimal spatial alignment of the ligands or due to the high local



**Fig. 3** N- and C-terminally Man-labeled PNA oligomers.

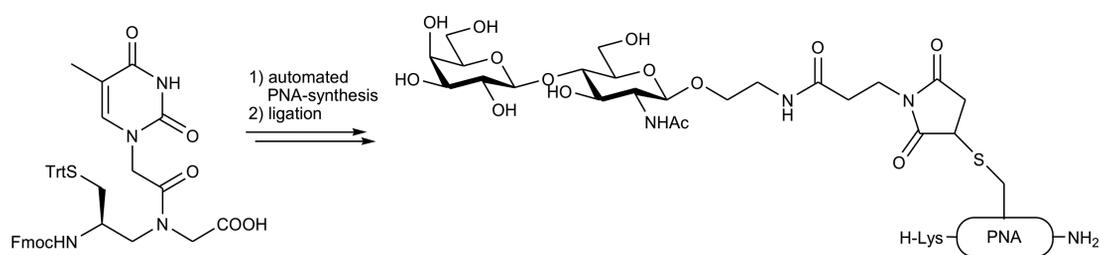
**Table 2** Binding parameters of the Man<sub>4</sub>-PNA/DNA complexes that bound to 2G12 (determined via SPR, , ,  PNA strand,  DNA strand, **0** Man<sub>4</sub>-ligand) [16].

Complex	Distance/Å	$k_d/M^{-1} s^{-1}$	$k_d/s^{-1}$	$K_D/\mu M$
7 	3	$3.2 \times 10^2$	$1.4 \times 10^{-3}$	4.2
8 	35	$2.0 \times 10^2$	$1.6 \times 10^{-3}$	8.1
9 	67	$1.8 \times 10^2$	$3.2 \times 10^{-3}$	17.3
10 	99	$9.7 \times 10^1$	$8.7 \times 10^{-3}$	89.4
11 	32	$5.1 \times 10^2$	$3.2 \times 10^{-3}$	6.4

sugar concentration (statistical effect) cannot be drawn [18]. Analogous to Kobayashi, the polymeric complex **11** was also assembled and tested for its binding affinity. Though it is anticipated to polymerize extensively, complex **11** had a higher dissociation constant than dimeric complex **7**. However, since this value is calculated for the DNA template concentration, it underestimates the affinity of the oligomeric complex **11**.

Recently, Winssinger described the hybridization of PNA-sugar conjugates with immobilized oligonucleotides [19]. This method facilitates the construction of glyco arrays and enables the combinatorial assembly of glycan structures in a high-throughput format.

We embarked on a study, in which the influence of (1) the valency of the ligands, (2) the distance between the ligands, and (3) the flexibility of the nucleic acid scaffold on the binding affinity can be assessed [20]. For the construction of the carbohydrate-labeled PNA oligomers, automated solid-phase PNA synthesis employing a thiol-modified PNA monomer and postsynthetic coupling using maleimide-activated *N*-acetylglucosamine (LacNAc) (Fig. 4) was used. The obtained PNA oligomers served as anticodons, which were hybridized with a DNA template providing the corresponding codons. Five different PNA oligomers, two of which contained LacNAc residues, provided three independent anticodon sequences. The corresponding DNA template sequence encoded four anticodons. The assembled complexes were then tested for their binding affinity toward *Erythrina cristagalli* lectin (ECL) by means of SPR. ECL consists of two protomers, each of which has a binding site for galactose-containing carbohydrates, including *N*-acetylglucosamine [21]. Analysis of the crystal structure revealed the distance



**Fig. 4** Internally LacNAc-labeled PNA oligomers.

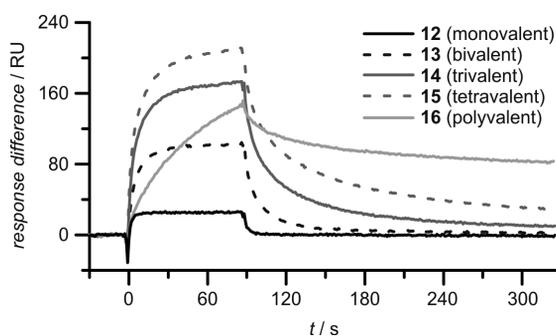
required for a multivalent substrate to simultaneously bind to both binding sites, i.e., approximately 100 Å. The discrepancy between the through-space distance of only 65 Å and the required distance of 100 Å arises from the fact that the binding sites are located on opposite sides of the lectin and, hence, a multivalent substrate has to bend around the convex surface.

#### *Influence of valency*

In an initial study, the influence of the valency of the multivalent complexes on the binding affinity was investigated. Following the assembly model described above, complexes carrying 1–4 LacNAc ligands (**12–16**) were constructed. As expected, the relative binding affinity per LacNAc ligand increased in the progression from mono- to di-, tri-, and tetravalent presentation, resulting in a 182-fold greater affinity in the tetravalent complex **15** than in the monovalent complex **12** (Table 3). The polyvalent cluster **16** was also assembled by hybridization of sticky ends. Interestingly, the SPR sensorgram revealed a dramatically reduced dissociation rate constant, a fact that is also typical for high-valent glycoclusters (Fig. 5) [22]. The reduced binding affinity, compared to complex **15**, is very likely due to the fact that the dissociation constant is for the DNA strand (compare with complex **11**).

**Table 3** Binding parameters for the recognition of the mono-, bi-, tri-, tetra-, and polyvalent LacNAc-PNA/DNA complex by ECL (determined via SPR, PNA strand, DNA strand, LacNAc-ligand) [20].

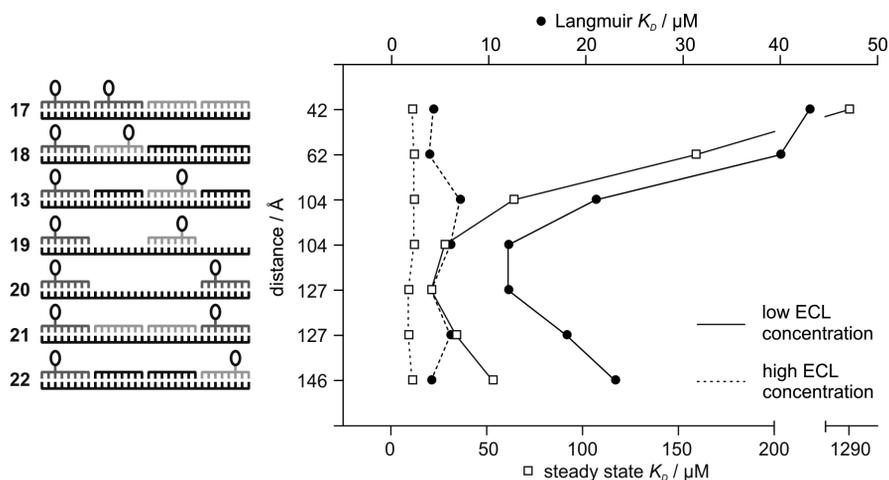
Complex	$k_d/M^{-1} s^{-1}$	$k_d/s^{-1}$	$K_D/\mu M$	Relative potency
<b>12</b>	$4.0 \times 10^2$	$3.2 \times 10^{-1}$	800	1
<b>13</b>	$4.1 \times 10^3$	$5.0 \times 10^{-2}$	12	33
<b>14</b>	$6.0 \times 10^3$	$1.5 \times 10^{-2}$	2.6	102
<b>15</b>	$6.8 \times 10^3$	$7.7 \times 10^{-3}$	1.1	182
<b>16</b>	$9.4 \times 10^2$	$2.3 \times 10^{-3}$	2.5	–



**Fig. 5** SPR sensorgrams showing the interactions between complexes **12–16** and immobilized ECL [20].

### Influence of distance

The influence of the distance between two LacNAc ligands on the binding affinity was also determined. Hence, complexes spanning distances from 42 to 146 Å (**13, 17–22**) were assembled and tested for their binding affinity. Initially, only little to no differences between the binding affinities were observed (Fig. 6, dashed lines). Considering the set-up of the SPR experiments, i.e., a solution of LacNAc-PNA/DNA complex which was passed over immobilized ECL, it was assumed that the density of immobilized lectin was so high that the bivalent substrates not only simultaneously bound to the two binding sites of a single ECL molecule but also to the binding sites of two adjacent ECL molecules. This binding event, often referred to as aggregation, leads to cross-linking of lectin molecules and can govern binding arrangements at high concentrations [23]. Though cross-linking cannot be ruled out completely, it was minimized by reducing the amount of immobilized ECL. Accordingly, the new determined binding affinities showed a clear distance dependency: the dissociation constant decreased as the LacNAc–LacNAc distance increased from 42 to 104 Å, reached a minimum before a further distance increase to 146 Å led again to increased dissociation constants (Fig. 6, solid lines). Whereas it is clear that the minimum can be attributed to a distance that is close to the spatial arrangement of the two binding sites of ECL, the decrease in binding affinity for both, shorter and longer distances, needs further



**Fig. 6** Influence of distance on the dissociation constants of bivalent complexes **13, 17–22**. Two different fitting models (Langmuir and the steady-state affinity model) were applied to the data (determined via SPR, PNA strand, DNA strand, LacNAc-ligand) [20].

explanation. The decreased binding affinity for shorter distances is due to the fact that these complexes are not able to simultaneously recognize the two binding sites of a single ECL molecule and, thus, are limited to cross-site binding. In contrast, assemblies with longer distances may still be capable of homo-site binding, albeit at the cost of strain since these complexes need to adopt energetically unfavorable conformations. But since the increased distance also increases the probability for cross-site binding, this might explain why the dissociation constants increase only moderately upon increase of the LacNAc–LacNAc spacing.

#### *Influence of flexibility*

The fact that simultaneous binding to the two binding sites of ECL can only be achieved for bivalent substrates capable of bending around the lectin prompted us to investigate the influence of substrate flexibility on the binding behavior in more detail. Though the implementation of nick sites in the double strand already provided a certain level of flexibility, the rigidity of the complexes was further decreased by the introduction of single-strand regions since these should bend more readily than completely base-paired complexes with nick sites only. In fact, complexes **19** and **20** showed the smallest dissociation constants of all tested bivalent substrates (Fig. 6, solid lines).

## SUMMARY

The studies mentioned above demonstrate that carbohydrate-PNA/DNA duplexes are well-suited scaffolds for the multivalent presentation of carbohydrates. The periodicity and rigidity of the nucleic acid helix allows the exact definition of number and spatial alignment of presented carbohydrate ligands. If necessary, the flexibility of the nucleic acid scaffold can be altered by introduction of nick sites and/or partially unpaired regions with varying length. In principle, a fluent transition from a rigid rod to an elastic string is possible. Moreover, the modularity of this approach keeps the synthetic effort relatively low.

It was shown that carbohydrate-PNA/DNA duplexes can be used as a “molecular ruler” to measure the distance between various carbohydrate binding sites. From a practical point of view, this tool could be used for the structure determination of lectins, which would facilitate the development of drugs combating infections or tumors. Moreover, the ability to precisely determine the binding affinity as a function of distance could provide insights into the molecular basis of the cluster glycoside effect.

## ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 765).

## REFERENCES

1. H. Lis, N. Sharon. *Chem. Rev.* **98**, 637 (1998).
2. H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, H. Rüdiger. *ChemBioChem* **5**, 740 (2004).
3. Y. C. Lee, R. T. Lee. *Acc. Chem. Res.* **28**, 321 (1995).
4. J. J. Lundquist, E. J. Toone. *Chem. Rev.* **102**, 555 (2002).
5. R. J. Pieters. *Org. Biomol. Chem.* **7**, 2013 (2009).
6. N. Jayaraman. *Chem. Soc. Rev.* **38**, 3463 (2009).
7. F. Diezmann, O. Seitz. *Chem. Soc. Rev.* (2011). <<http://dx.doi.org/10.1039/C1CS15054E>>
8. P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt. *Science* **254**, 1497 (1991).
9. E. Uhlmann, A. Peyman, G. Breipohl, D. W. Will. *Angew. Chem., Int. Ed.* **37**, 2796 (1998).
10. M. Eriksson, P. E. Nielsen. *Nat. Struct. Mol. Biol.* **3**, 410 (1996).
11. E. Protozanova, P. Yakovchuk, M. D. Frank-Kamenetskii. *J. Mol. Biol.* **342**, 775 (2004).

12. K. Matsuura, M. Hibino, Y. Yamada, K. Kobayashi. *J. Am. Chem. Soc.* **123**, 357 (2001).
13. K. Matsuura, M. Hibino, T. Ikeda, Y. Yamada, K. Kobayashi. *Chem.—Eur. J.* **10**, 352 (2004).
14. Y. Yamada, K. Matsuura, K. Kobayashi. *Bioorg. Med. Chem.* **13**, 1913 (2005).
15. E. C. Sweeney, A. G. Tonevitsky, D. E. Temiakov, I. I. Agapov, S. Saward, R. A. Palmer. *Proteins: Struct., Funct., Bioinf.* **28**, 586 (1997).
16. K. Gorska, K.-T. Huang, O. Chaloin, N. Winssinger. *Angew. Chem., Int. Ed.* **48**, 7695 (2009).
17. 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* **300**, 2065 (2003).
18. C. R. Cantor, P. R. Schimmel. *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*, pp. 850–852, W.H. Freeman, San Francisco (1980).
19. K.-T. Huang, K. Gorska, S. Alvarez, S. Barluenga, N. Winssinger. *ChemBioChem* **12**, 56 (2011).
20. C. Scheibe, A. Bujotzek, J. Dervede, M. Weber, O. Seitz. *Chem. Sci.* **2**, 770 (2011).
21. K. Turton, R. Natesh, N. Thiyagarajan, J. A. Chaddock, K. R. Acharya. *Glycobiology* **14**, 923 (2004).
22. X. Zeng, T. Murata, H. Kawagishi, T. Usui, K. Kobayashi. *Carbohydr. Res.* **312**, 209 (1998).
23. J. C. Sacchettini, L. G. Baum, C. F. Brewer. *Biochemistry* **40**, 3009 (2001).