Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202203591
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Biomimetic Tweezers for N-Glycans: Selective Recognition of the Core $GlcNAc_2$ Disaccharide of the Sialylglycopeptide SGP

Francesco Milanesi,[a, b] Luca Unione,[c, d] Ana Ardá,[c, d] Cristina Nativi,[a] Jesús Jiménez-Barbero,[c, d, e, f] Stefano Roelens,[a] and Oscar Francesconi*[a]

Abstract: In recent years, glycomics have shown how pervasive the role of carbohydrates in biological systems is and how chemical tools are essential to investigate glycan function and modulate carbohydrate-mediated processes. Biomimetic receptors for carbohydrates can carry out this task but, although significant affinities and selectivities toward simple saccharides have been achieved, targeting complex glycoconjugates remains a goal yet unattained. In this work we report the unprecedented recognition of a complex biantennary sialylglycopeptide (SGP) by a tweezers-shaped biomimetic receptor, which selectively binds to the core $GlcNAc_2$ disaccharide of the N-glycan with an affinity of 170 μM. Because of the simple structure and the remarkable binding ability, this biomimetic receptor can represent a versatile tool for glycoscience, opening the way to useful applications.

Introduction

Oligosaccharides (glycans, carbohydrates) are widely expressed on the surface of eukaryotic cells as glycoconjugates of proteins and lipids.[1,2] Their selective recognition by different classes of proteins, such as lectins, triggers essential physio- and pathological processes, like cell adhesion and infection, cell differentiation, inflammation, and immune response.[3–5] Glycosylation is an ubiquitous post-translational modification of proteins that leads to conjugation with O- and N-linked oligosaccharides. In particular, N-glycans, that is, N-linked oligosaccharides, all share the common pentasaccharide structure $Man_3GlcNAc_2$ [Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ] N-linked to the asparagine residue of the Asn-X-Thr/Ser (X is any amino acid except Pro) sequence in the protein backbone.[6] Depending on the structural diversity of the oligosaccharides extending beyond the common structure, N-glycans are classified as high-mannose type, complex-type, and hybrid type. N-glycans play essential roles in the chemical and biological behaviour of proteins, including protein folding and stability,[7,8] solubility,[9] resistance to proteolysis,[10] intracellular traffic,[11] antigenicity,[12] activity, and recognition by carbohydrate-binding proteins.[13] Because of glycan ubiquity, molecular agents targeting N-glycans can lead to a deeper understanding of their function and to the development of new diagnostic and prognostic strategies.[14] Typically, recognition of specific saccharides to elicit function is achieved in nature by lectins, which are glycan binding proteins specialized in recognizing specific sugar epitopes.[15] However, their pharmaceutical development has remained largely unfulfilled, due to their susceptibility to proteolysis, their ability to elicit immune response, and other features related to their protein nature.

Synthetic biomimetic receptors for carbohydrates are a heterogeneous class of molecules developed for selective recognition of sugars of biological interest.[16–21] To mimic the recognition function of lectins, the binding ability of biomimetic receptors relies exclusively on non-covalent interactions. However, achieving effective recognition of polar species like saccharides in water leveraging a biomimetic approach is a

[a] Dr. F. Milanesi, Prof. C. Nativi, Dr. S. Roelens, Dr. O. Francesconi
Department of Chemistry "Ugo Schiff" DICUS and INSTM
University of Florence
Polo Scientifico e Tecnologico, I–50019 Sesto Fiorentino, Firenze (Italy)
E-mail: oscar.francesconi@unifi.it
[b] Dr. F. Milanesi
Magnetic Resonance Center CERM
University of Florence
Via L. Sacconi 6, I-50019
Sesto Fiorentino, Firenze (Italy)
[c] Dr. L. Unione, Dr. A. Ardá, Prof. J. Jiménez-Barbero
CICbioGUNE,
Basque Research & Technology Alliance (BRTA)
Bizkaia Technology Park, Building 800,
48160 Derio, Bizkaia (Spain)
[d] Dr. L. Unione, Dr. A. Ardá, Prof. J. Jiménez-Barbero
Ikerbasque, Basque Foundation for Science
Maria Diaz de Haro 3, 48013 Bilbao, Bizkaia (Spain)
[e] Prof. J. Jiménez-Barbero
Department of Organic Chemistry,
II Faculty of Science and Technology
University of the Basque Country, EHU-UPV
48940 Leioa (Spain)
[f] Prof. J. Jiménez-Barbero
Centro de Investigación Biomédica En Red de Enfermedades Respiratorias
Madrid (Spain)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202203591

non-trivial task.[22,23] Indeed, milestones in this field have been achieved only recently, mainly by means of sophisticated macrocyclic architectures.[24–30] In contrast, acyclic receptors have been largely unexplored,[31–34] although they may present distinct advantages over macrocyclic structures in terms of simpler synthesis, more easily modifiable structure, easy adaptability to different guests and, most of all, the possibility of binding non-terminal saccharide entities.

We have recently reported design, synthesis, and applications of the acyclic, tweezers-shaped receptor **1** (Figure 1),[35,36] which effectively recognizes standard disaccharides in water at pH 7.4. Receptor **1** is based on a very simple structure, consisting of a diaminocarbazole hydrogen-bonding unit[25,37] and two anthracene rings, establishing CH-π interactions[38] with the non-polar region of carbohydrates. Receptor **1** selectively recognized the methyl β glycoside of *N,N'*-diacetylchitobiose (MeβGlcNAc$_2$), the core disaccharide of the conserved pentasaccharide structure of N-glycans, over a wide range of mono- and disaccharides, showing a remarkable affinity of 160 μM. This affinity value not only exceeds that exhibited by more complex macrocyclic receptors, but also that shown by the lectin hevein (isolated from *Hevea brasiliensis*) by one order of magnitude ($K_d = 1.61$ mM, $\log\beta = 2.79$).[39] Therefore, receptor **1** stands as the ideal scaffold to recognize complex-type N-glycans of biological interest by binding to the core GlcNAc$_2$ disaccharide.

Among N-glycans, sialylated N-glycans are of particular interest, because are the target of a class of lectins named Siglecs (sialic acid binding immunoglobulin type lectins).[40–42] Siglecs are cell surface immunomodulatory receptors, which selectively recognize specific sialylated glycans exposed on all mammalian cells and help immune cells to distinguish between self and non-self epitopes. One of the most ubiquitous sialylated N-glycans in mammals is the complex-type biantennary sialylundecasaccharide SGP, which can be obtained from egg yolk as the glycopeptide[4,43] (Figure 1). Recognition of glycans by lectins, including siglecs, usually occurs by binding to the epitope located at the end of the oligosaccharide chain; tweezers-shaped receptors may offer the opportunity to leverage a different binding site, located in the conserved unit of the glycan core, thus providing a new and potentially very useful recognition tool.

Following this alternative recognition approach, in this work we report on the unprecedented features of a tweezers-shaped receptor effectively recognizing the natural N-glycan SGP by selectively binding to the GlcNAc$_2$ core disaccharide.

Results and Discussion

Evaluation of the recognition properties of receptor **1** toward complex-type N-glycans is, unfortunately, prevented by strong self-association in water ($\log\beta_{dim} = 2.65 \pm 0.07$),[36] due to the presence of the extended aromatic surfaces of the anthracene moieties, stacking intermolecularly under the pressure of hydrophobic effects.[44] In addition, such a strong dimerization phenomenon prevents not only a quantitative determination of binding affinities, but also a structural definition of the binding modes when binding to complex glycans. To overcome the self-association issue, we have functionalized the *para-* positions of the anthracene rings of **1** with two additional phosphonate groups, giving rise to the tetraphosphonate receptor **2** (Figure 1).

Receptor **2** was prepared from commercially available 9,10-dibromoanthracene (**3**) in six steps (Figure 2a). After formylation of **3** using n-BuLi followed by DMF, the carbaldehyde **4** was protected as the dimethylpropylene acetal **5** by the pTsOH catalyzed condensation with neopentyl glycol. The bromo-substituent of the acetal intermediate was then replaced by a diethylphosphonate group using n-BuLi followed by PO(OEt)$_2$Cl. The acetal **6** was then hydrolyzed by HCl in acetone, to give the aldehyde **7**, which was condensed with the diaminocarbazole **8**, prepared according to a previously reported method,[25] to provide the corresponding bis-imino derivative. Subsequent reduction with NaBH$_4$ gave the liposoluble receptor **9**. Hydrolysis of the phosphonate esters with TMSBr, followed by treatment with methanol gave the desired receptor **2**.

The concentration dependent oligomerization state of receptor **2** was investigated by NMR. In contrast to its progenitor, dilution experiments carried out on **2** at pH 7.4 revealed sharp NMR signals in a wide concentration range, from 24 to 1 mM, together with a modest variation of ^{1}H NMR chemical shifts, suggesting a weak self-association behaviour

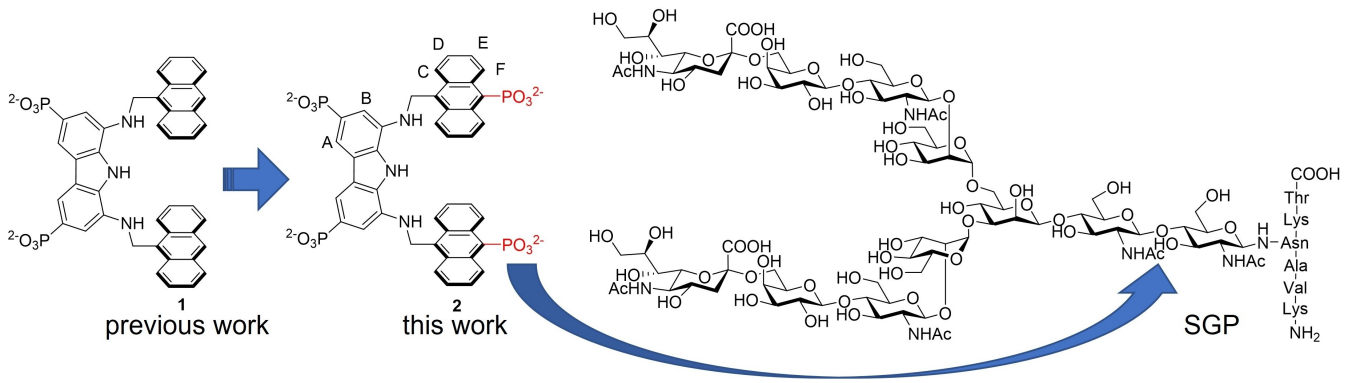

Figure 1. Concept of the work: conversion of the biomimetic receptor **1** to **2** (with proton labelling) to be tested vs. the sialylglycopeptide SGP.

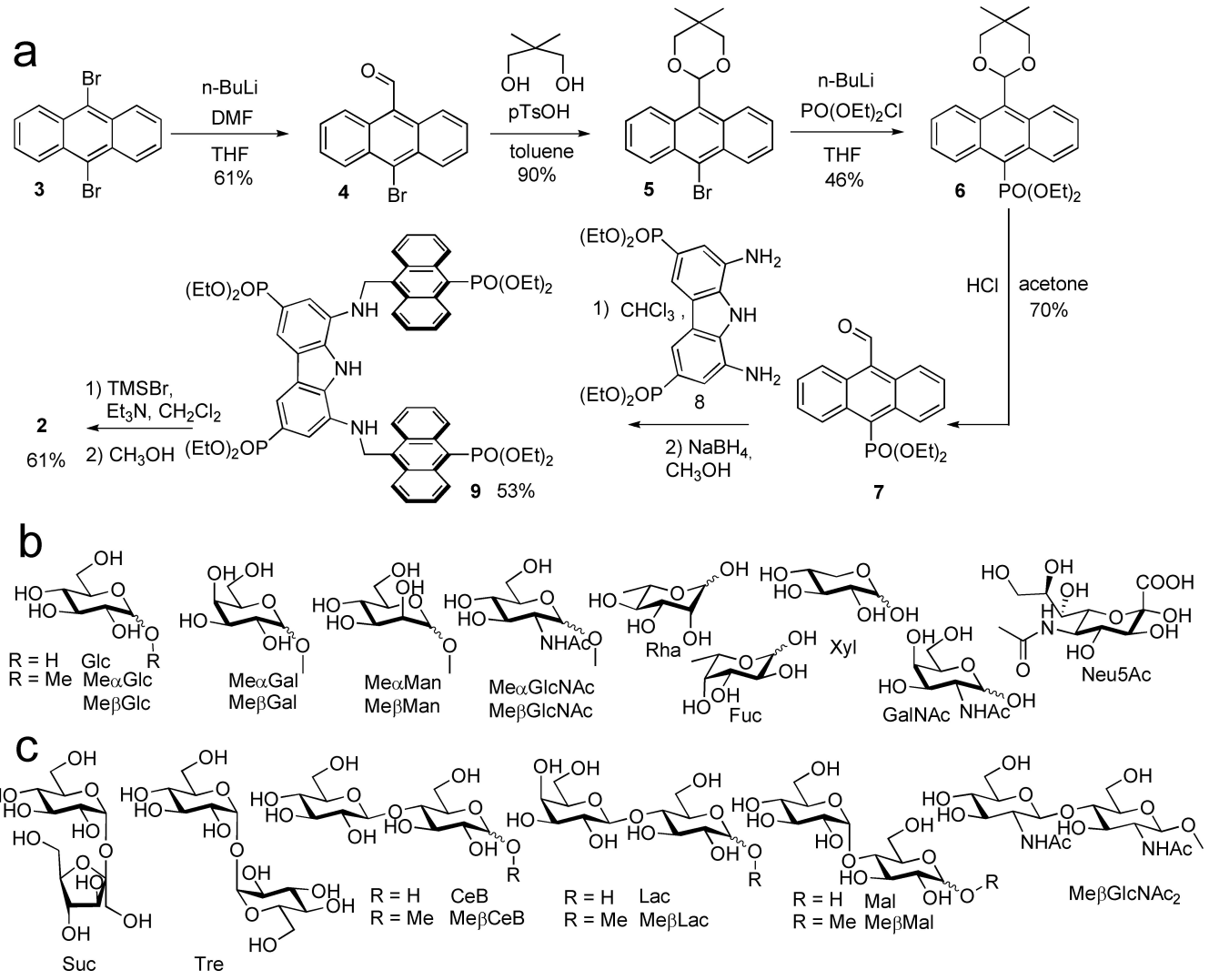

Figure 2. (a) Synthetic scheme of the biomimetic receptor **2**. Investigated monosaccharides (b) and disaccharides (c) and their abbreviations.

(Figure S17). Nonlinear regression of the corresponding data gave a dimerization constant ($\log\beta_{dim} = 0.98 \pm 0.05$) that, although still detectable, is more than one order of magnitude smaller than that observed with the parent receptor **1**. As expected, the presence of additional hydrosolubilising groups on the anthracene moieties prevented π-stacking interactions and, consequently, self-association phenomena.

To ascertain whether the introduction of phosphonate groups on anthracene rings may perturb the binding properties of receptor **1**, we tested receptor **2** toward the same set of mono- and disaccharides used for receptor **1**. A screening of the binding properties of receptor **2** was thus carried out by ^{1}H NMR spectroscopy (Figures S14-S16) toward the mono- and disaccharides shown in Figure 2b and 2c, respectively. The pool included several monosaccharides [glucose (Glc), N-acetylgalactosamine (GalNAc), rhamnose (Rha), fucose (Fuc), xylose (Xyl), sialic acid (Sia), and α and β methyl derivatives of Glc, Gal, Man, and GlcNAc], together with a few disaccharides [sucrose (Suc), trehalose (Tre), cellobiose (CeB), maltose (Mal), lactose (Lac) and the methyl glycoside of *N*,*N*′-diacetylchitobiose (MeβGlcNAc$_2$)]. In all cases, the chemical shift perturbations (CSPs) of the ^{1}H NMR signals of the sugars upon addition of equimolar amounts of **2** were monitored. Except for methyl-β-glucoside (MeβGlc), for which an appreciable variation was measured ($\Delta\delta_{H\text{-}1} = 0.05$ ppm), rather small ($\Delta\delta < 0.04$ ppm) or no variations were detected for other monosaccharides. Concerning disaccharides, MeβGlcNAc$_2$ experienced a massive perturbation of the proton signals, showing a marked line broadening that spread most signals into the spectrum baseline, which indicated chemical exchange processes between free and bound species in the medium/fast regime in the NMR time scale. While no changes were observed for Suc and Tre, a marked shift was instead observed for the β anomer of CeB and, to a smaller extent, for those of Lac and Mal, analogous to that previously reported for the parent receptor **1**. Moreover, for the CeB and Lac disaccharides, the addition of receptor **2** induced a significant line broadening of the NMR signals. This evidence strongly suggests a similar binding profile for both receptors.

Table 1. Cumulative formation constants ($\log\beta_n$)[a] and intrinsic median binding concentration (BC_{50}^0, mM)[b] for receptor to glycoside (R:G) complexes of **2** with methyl glycosides and SGP, measured at 298 K from NMR data in D_2O at pD 7.4 and from ITC data in H_2O at pH 7.4.[c] BC_{50}^0 (mM) for receptor **1** are reported for comparison.

Receptor Glycoside	R:G	**2** (NMR) log β	BC_{50}^0	**2** (ITC) log β	BC_{50}^0	**1** (NMR) BC_{50}^0
MeβGlc	1:1	1.59 ± 0.01	38.4 ± 1.7			
MeβCeB	1:1	2.40 ± 0.01	1.04 ± 0.06	2.49 ± 0.07	1.32 ± 0.43	0.94 ± 0.10
	2:1	5.79 ± 0.07		5.53 ± 0.38		
	2:2	8.00 ± 0.09				
MeβMal	1:1	2.00 ± 0.01	11.9 ± 0.3	1.78 ± 0.02	16.6 ± 0.7	31.0 ± 4.4
MeβLac	1:1	2.11 ± 0.01	9.01 ± 016	1.97 ± 0.01	10.5 ± 0.3	30.7 ± 4.7
$Me\beta GlcNAc_2$	1:1	3.47 ± 0.02	0.27 ± 0.01	3.45 ± 0.10	0.18 ± 0.04	0.16 ± 0.01
	2:1	6.48 ± 0.05		7.16 ± 0.23		
SGP	1:1			3.44 ± 0.26	0.17 ± 0.05	
	2:1			7.30 ± 0.26		

[a] Formation constants were obtained by nonlinear least-square regression analysis of NMR data. [b] Calculated from the log β values using the "BC50 Calculator" program.[45] [c] The receptor dimerization constants at pH 7.4 (**2**: log $\beta_{dim} = 0.98 \pm 0.05$) was set invariant in the nonlinear regression analysis of NMR.

A quantitative investigation was thus carried out by NMR spectroscopy on the interaction of **2** with MeβGlc, MeβCeB, MeβMal, MeβLac, and $Me\beta GlcNAc_2$. Methyl glycosides were used to avoid interconversion equilibria between anomers. The cumulative association constants reported in Table 1 were measured by ^{1}H NMR titrations in D_2O (pD 7.4) at 298 K, by simultaneously fitting the complexation induced shifts of all the available signals to the appropriate association model by nonlinear regression analysis. Because multiple binding constants were detected, affinities were assessed though the intrinsic median binding concentration parameter BC_{50}^0,[45] calculated from the measured binding constants and reported in Table 1.

At first glance, data confirm a binding profile closely similar to that observed for the progenitor **1**. Indeed, receptor **2** effectively discriminates disaccharides from monosaccharides, showing a significantly lower affinity for MeβGlc than for MeβCeB. Like seen for **1**, receptor **2** shows a distinct preference for all-equatorial disaccharides, and for MeβCeB and $Me\beta GlcNAc_2$ with respect to MeβLac and MeβMal. Most importantly, the largest affinity was observed for $Me\beta GlcNAc_2$ (270 μM), very similar to that measured for **1** (160 μM), showing that additional phosphonate groups affect binding properties only marginally.

To shed light on the energetics of binding and to support the NMR-based affinities measured for the 1,4-linked disaccharides, binding was also investigated by ITC. Because the dimerization constant of the receptor **2** is negligible in H_2O at pH 7.4, dimerization was not included in the nonlinear regression analysis of ITC data. Cumulative binding constants, together with affinity values are gathered in Table 1 to provide a direct comparison with the NMR results. The good agreement between the two independent techniques confirmed the reliability of the observed affinities. In contrast to receptor **1**, for which strong dimerization prevented the determination of thermodynamic parameters, reliable data for the formation of 1:1 complexes could be obtained for **2** (Table 2), from which it can be appreciated that enthalpy provides the impetus for the formation of all 1:1 complexes, whereas an adverse entropic contribution is consistent with a prominent role of intermolecular forces over solvophobic effects.

Table 2. Thermodynamic parameters (kJ mol^{-1}) for the formation of receptor **2** to glycoside 1:1 complexes of MeβCeB. MeβMal. MeβLac, $Me\beta GlcNAc_2$, and SGP in H_2O. pH 7.4 at 298 K.

Glycoside	$-\Delta G^0$	$-\Delta H^0$	$T\Delta S^0$
MeβCeB	14.2 ± 0.4	26.1 ± 1.4	−11.9 ± 1.0
MeβMal	10.2 ± 0.1	17.9 ± 0.5	−7.7 ± 0.6
MeβLac	11.3 ± 0.1	23.4 ± 0.5	−12.1 ± 0.6
$Me\beta GlcNAc_2$	19.7 ± 0.6	47.3 ± 2.2	−27.7 ± 2.7
SGP	19.6 ± 1.5	14.5 ± 5.0	5.2 ± 6.5

Based on the selectivity shown by receptor **2** for $Me\beta GlcNAc_2$, the determination of binding affinity toward the SGP N-glycan was attempted by NMR techniques. Because of strong signal overlapping, useful chemical shift data could not be collected for a quantitative analysis (Figure S23). Thus, the binding ability of **2** toward SGP was measured by ITC titrations (Figure 3a), which were carried out in H_2O at pH 7.4 (Table 1), giving an affinity value of 170 μM, slightly larger than that measured for $Me\beta GlcNAc_2$. In contrast, the thermodynamic parameters were unexpectedly different (Table 2): while the free energy of binding is essentially the same compared to $Me\beta GlcNAc_2$, the enthalpic contribution for SGP recognition is significantly smaller, whereas a positive entropy is responsible for binding. Results suggest a drastic change in the solvation features of SGP with respect to $Me\beta GlcNAc_2$.

The affinity of **2** for SGP measured by ITC was further confirmed by Circular Dichroism (CD) as an independent technique. Because both, the UV chromophore **2** and the chiral SGP, are silent on CD, the technique is particularly convenient, because it monitors the unambiguous formation of the complex between the two partners. CD titrations were carried out on pre-neutralized reactants in H_2O at pH 7.4. Upon addition of SGP to receptor **2**, the CD intensity increased at 254.5, 260 and 265.5 nm, featuring negative sign for the first two wavelength and positive for the latter (Figure 3b). Nonlinear least-square regression analysis of CD data fitted a 1:1 stoichiometry model, giving $\log\beta = 3.737 \pm 0.003$, corresponding to $BC_{50}^0 = 183 \pm$

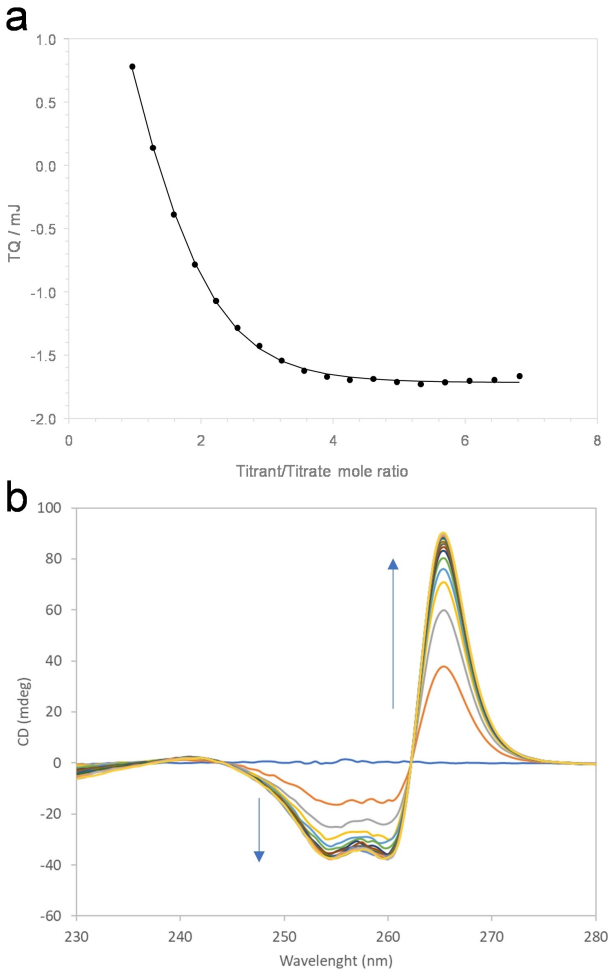

Figure 3. Titrations of **2** with SGP in H_2O (pH 7.4) at 298 K: a) ITC plot of experimental data points, [black circles - total heat (TQ) vs. **2**/SGP mole ratio] and fitted curve [solid line] calculated by non-linear regression analysis; b) CD titration spectra of **2** (63 μM) with SGP.

1 μM, closely comparable to the affinity obtained by ITC. The highly diluted conditions used for the CD titration (63 μM) account for the simpler model, in which higher stoichiometry complexes are undetectable.

It should be emphasized that, to the best of our knowledge, the molecular recognition of N-glycans by biomimetic artificial receptors has not been previously reported. Indeed, the natural products belonging to the antibiotic pradimicin family is the only known non-peptidic family binding to N-glycans .[46] Moreover, it is noteworthy that the affinity achieved by receptor **2** is fully comparable to that of natural lectins.[39]

Prompted by the above recognition properties, an investigation on the three-dimensional structure of the complex between **2** and SGP was attempted using NMR experimental data assisted by molecular modelling calculations.[47] 1D and 2D NMR spectra (Figures S24–S31) were carried out to assign all the available proton signals of SGP.

Information on the geometry of the major complex was extracted from the analysis of 2D NMR NOESY spectra recorded for a 1:1 mixture of **2** and SGP. Several intermolecular NOE contacts were unambiguously detected between the receptor and residues a, b, c, and d of the undecasaccharide, indicating that molecular recognition takes place on the N-glycan core (Figure 4a). As observed for the interaction of **2** with the simple disaccharide MeβGlcNAc$_2$, the methyl of the N-acetyl group (CH_3- a) displayed a NOE contact with the carbazole H–B proton. Moreover, H-2-b of GlcNAc$_2$ showed a NOE cross peak with the anthracene H–F, clearly indicating that the disaccharide is located between the two anthracene rings, inside the receptor cleft. Additional intermolecular contacts were inferred by the presence of NOE cross peaks involving H–D and H–E of the anthracene and H-5-c of the β-mannose residue, and with H-6-d and H-6-d′ of the α6-linked mannose residue, suggesting that the glycan extends outside the binding cleft of **2** beyond the anthracene moieties. Eventually, the carbazole H–B shows a contact with H-5-c, while H–C shows a contact with the H-6′ of residue d.

We next carried out a detailed analysis of the CSPs observed in the NMR spectra of SGP. ^{1}H-^{13}C HSQC NMR experiments (Figure 4c) showed significant changes in specific N-glycan signals upon addition of an equimolar amount of receptor **2**. A dramatic line broadening effect, spreading the signal below the limit of signal detection, was experienced by the cross-peaks corresponding to the core trisaccharide (residues a, b, and c), further confirming that recognition takes place on the N-glycan core. Concerning the two arms, the H-1 and H-2 of the α6Man (residue d) also disappeared, whereas for H-3, H-4, H-5, and H-6,6′ significant CSPs (>0.05 ppm) were detected. Perturbations observed for the α3Man (residue d′) were different, as H-1 and H-2 did not disappear, rather they showed CSPs above 0.05 ppm, while the H-3, H-4, H-5 and H-6,6′ showed modest CSPs (<0.05 ppm). Beyond the mannose residues, only marginal CSPs were detected for some of the resonances from the GlcNAc residues e and e′, whereas no variations were observed for Gal (f and f′) and Neu5Ac (g and g′). Similarly, the amino acids N, A, and K disappear, the amino acid V shows CSP, whereas the absence of CSP for N and C terminal amino acids K′ and T indicates that these residues are located far from the receptor cleft.

From the above evidence, we conclude that the recognition selectively occurs on the core GlcNAc$_2$ disaccharide, leaving terminal saccharide residues and distant amino acids unperturbed. However, the above differences in CSPs and line broadening for the two Man residues suggest that the two arms of the SGP span different conformational spaces with respect to the aromatic scaffold of the biomimetic receptor. Interestingly, conformational studies on biantennary N-glycans structurally related to SGP have shown that, beyond the "extended" conformation, other geometries of the α6 arm are also possible, including the so-called "backfolded" orientation,[48] in which the α6-linked Man is rotated toward the GlcNAc$_2$ core, so that contacts with the receptor are also expected. Thus, the observed differences in CSPs for the two Man residues point to the existence of an equilibrium between the extended and the backfolded conformations.

To provide a three-dimensional description of the major binding modes, the experimental data were supported by

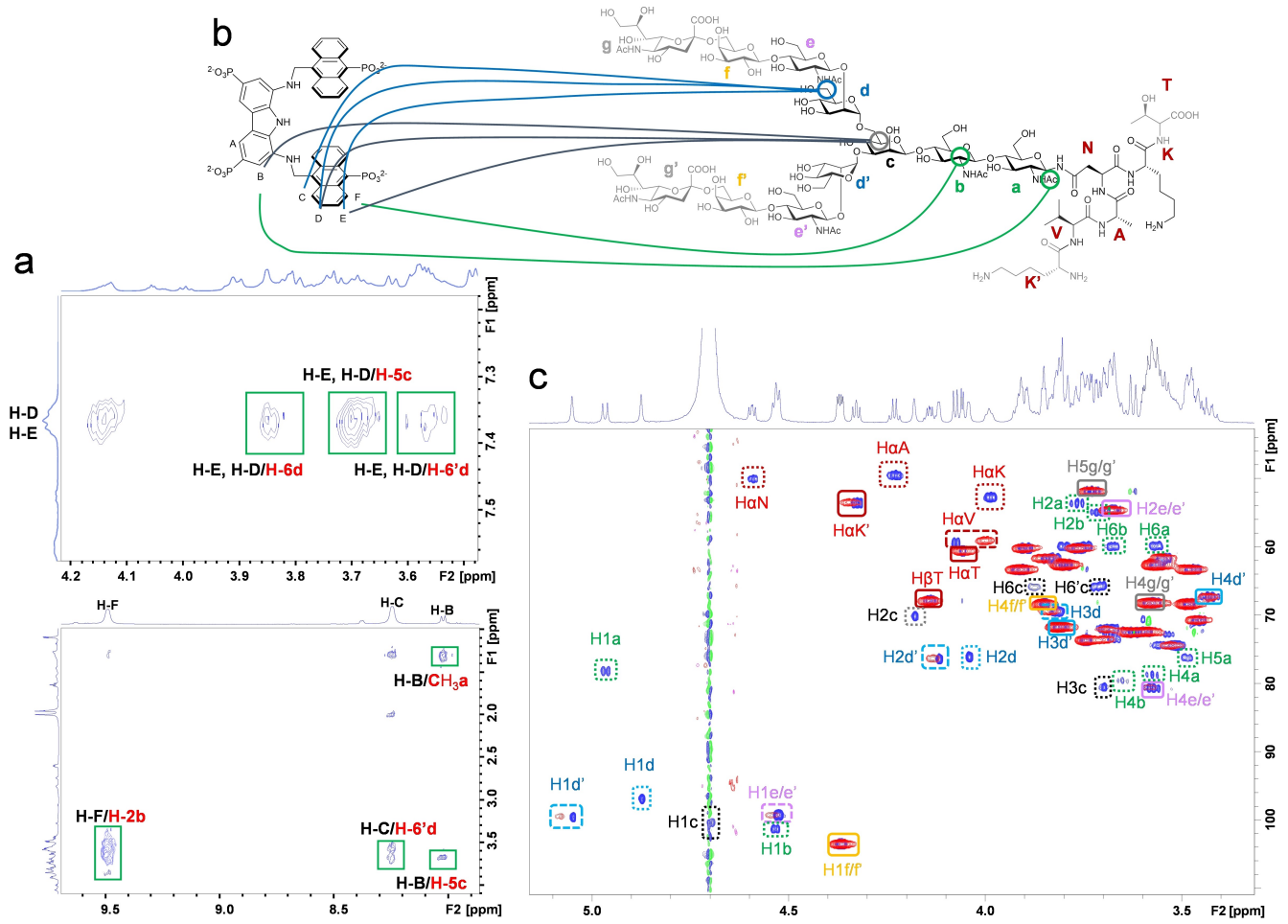

Figure 4. Studies on the binding mode between **2** and SGP: a) sections of the 800 MHz NOESY (500 ms) spectrum recorded for an equimolar mixture of SGP and **2** (2 mM each) in D_2O at 298 K. Unambiguous intermolecular NOE cross peaks are indicated by squares; b) A schematic representation of the intermolecular NOEs found between **2** and SGP is depicted at the right hand side; c) Superimposition of the ^{1}H-^{13}C HSQC spectra recorded for a 2 mM solution of SGP before (blue peaks) and after (red peaks) the addition of an equimolar amount of **2**. Selected peaks are squared and labelled: peaks that present chemical shift variation are squared using a dashed line; peaks that disappear after addition of **2** are squared using a dotted line.

molecular modelling calculations. A conformational search was carried out on the 1:1 complex formed between **2** and SGP. A family of minimum energy conformers was found that agrees with all the NOE contacts experimentally observed. A representative geometry within this minimum energy family is depicted in Figure 5a. It can be appreciated that the tweezers-shaped architecture of **2** is perfectly suited to bind the core $GlcNAc_2$ disaccharide in the N-glycan stem. Interestingly, all the NOE contacts found, including those between H-6-d and the H–D and H–E protons of anthracene, agree with an extended conformation, in which the α6-linked Man (d) is close to one of the anthracene moieties. The different chemical shift perturbation and chemical exchange behaviour observed for the two arms, which are more significant for the α6- than for the α3-linked Man, may be due to a backfolded conformation, which was indeed found in the conformational search (Figure 5b), and that may be present to a non-negligible extent in solution. Finally, several hydrogen-bonding, with O···H interatomic distances shorter than the sum of the van der Waals radii, and CH-π interactions have been detected in the calculated models involving the core $GlcNAc_2$ disaccharide (Figure 5c). As previously observed in the complex of **1** with $Me\beta GlcNAc_2$, while the tridentate diaminocarbazole of **2** binds the OH-3a, the N-acetyl group establishes CH-π interactions with the carbazole unit. Additional hydrogen bonds involving the phosphonate group and several CH-π interactions with both the anthracenes contribute to the complex stabilization.

Conclusion

In conclusion, a second-generation biomimetic receptor **2** is presented, for which the self-association issues of the progenitor have been addressed by introducing additional phosphonate substituents, leaving the binding properties toward $Me\beta GlcNAc_2$ essentially unaffected. The tweezers-shaped biomimetic acyclic architecture of **2** shows the unprecedented feature of recognizing with high selectivity the $GlcNAc_2$

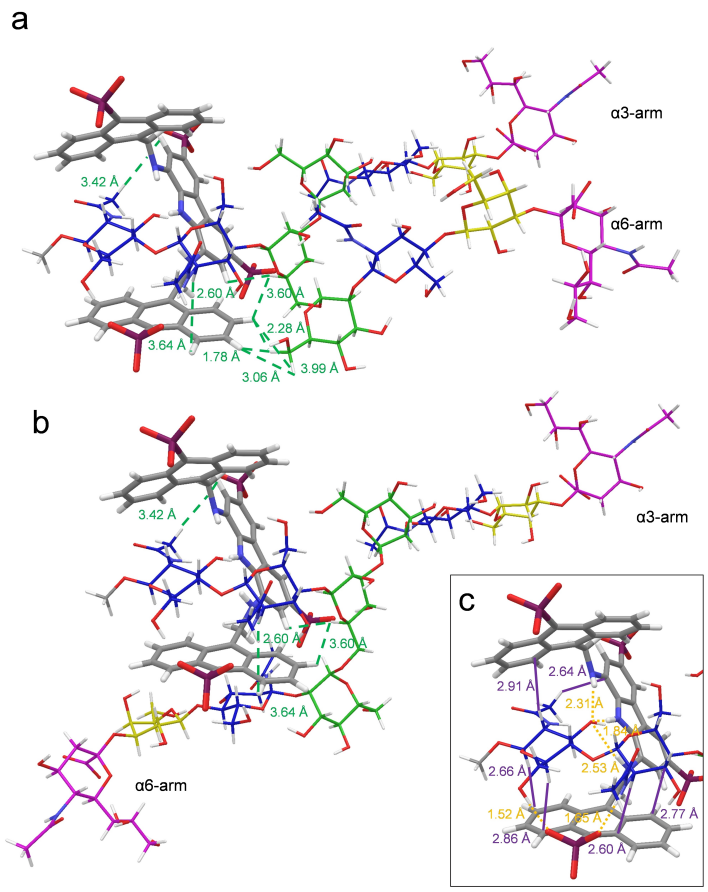

Figure 5. Representative global energy minimum structures of the complex formed between **2** and SGP, in which the peptide portion has been replaced with a methyl group: a) extended conformation; b) back-folded conformation. All the strongest intermolecular NOEs experimentally found between **2** and SGP are indicated as dashed lines, along with the corresponding distances [Å] calculated from this geometry. All observed key NOE contacts are accounted for by these 3D model, which also explains the experimental chemical shift perturbations and the distinct line broadening phenomena observed at the different N-glycan residues; c) intermolecular hydrogen-bonding and CH-π interactions found in the calculated structure are indicated as dotted (hydrogen-bonds) and solid (CH-π) lines, with corresponding oxygen/hydrogen and π-system/hydrogen distances [Å].

disaccharide unit in the core of a complex-type N-glycan. Due to the simple and easily accessible/modifiable structure, receptor **2** stands as a new and effective tool for glycoscience research. Indeed, because recognition occurs at a site different from the sialic terminal usually recognized by biological receptors, receptor **2** suggests potential applications in analysis, separation, and imaging of N-glycans, as well as modulator of N-glycan recognition processes.

Acknowledgements

We thank MIUR-Italy "Progetto Dipartimenti di Eccellenza 2018–2022" allocated to Department of Chemistry Ugo Schiff, COST Action (CA18132), MIUR-Italy PRIN2017 (2017XZ2ZBK) for granting a fellowship to F.M. and Ente Cassa di Risparmio di Firenze (Italy) is acknowledged for granting an ITC nanocalorimeter and a high-field NMR spectrometer. Open Access funding provided by Università degli Studi di Firenze within the CRUI-CARE Agreement.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: biomimetic artificial receptors · carbohydrates · chitobiose · glycan · molecular recognition

[1] B. Ernst, G. W. Hart, P. Sinaý, *Carbohydrates in Chemistry and Biology*, Wiley, **2000**.
[2] A. Varki, *Glycobiology* **2017**, *27*, 3–49.
[3] B. A. H. Smith, C. R. Bertozzi, *Nat. Rev. Drug Discovery* **2021**, *20*, 217–243.
[4] C. Reily, T. J. Stewart, M. B. Renfrow, J. Novak, *Nat. Rev. Nephrol.* **2019**, *15*, 346–366.
[5] N. Taniguchi, Y. Kizuka, *Adv. Cancer Res.* **2015**, 11–51.

[6] P. Stanley, K. W. Moremen, N. E. Lewis, N. Taniguchi, M. Aebi, in *Essential of Glycobiology*, 4th Edition (Eds.: A. Varki, R. D. Cummings, J. D. Esko, et al.), Cold Spring Harbor (NY), **2022**.
[7] D. Shental-Bechor, Y. Levy, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 8256–8261.
[8] N. G. Jayaprakash, A. Surolia, *Biochem. J.* **2017**, *474*, 2333–2347.
[9] A. Paul, D. Segal, E. Zacco, *Neural. Regen. Res.* **2021**, *16*, 2215.
[10] A. Amore, B. C. Knott, N. T. Supekar, A. Shajahan, P. Azadi, P. Zhao, L. Wells, J. G. Linger, S. E. Hobdey, T. A. Vander Wall, T. Shollenberger, J. M. Yarbrough, Z. Tan, M. F. Crowley, M. E. Himmel, S. R. Decker, G. T. Beckham, L. E. Taylor, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13667–13672.
[11] S. Hirayama, Y. Hori, Z. Benedek, T. Suzuki, K. Kikuchi, *Nat. Chem. Biol.* **2016**, *12*, 853–859.
[12] W. Dowling, E. Thompson, C. Badger, J. L. Mellquist, A. R. Garrison, J. M. Smith, J. Paragas, R. J. Hogan, C. Schmaljohn, *J. Virol.* **2007**, *81*, 1821–1837.
[13] R. L. Schnaar, *J. Allergy Clin. Immun.* **2015**, *135*, 609–615.
[14] J. Balzarini, *Nat. Rev. Microbiol.* **2007**, *5*, 583–597.
[15] N. Sharon, *Glycobiology* **2004**, *14*, 53R-62R.
[16] A. P. Davis, *Chem. Soc. Rev.* **2020**, *49*, 2531–2545.
[17] O. Francesconi, S. Roelens, *ChemBioChem* **2019**, *20*, 1329–1346.
[18] O. Francesconi, L. Donnici, M. Fragai, E. Pesce, M. Bombaci, A. Fasciani, L. Manganaro, M. Conti, R. Grifantini, R. de Francesco, C. Nativi, S. Roelens, *iScience* **2022**, *25*, 104239.
[19] M. F. Bravo, M. A. Lema, M. Marianski, A. B. Braunschweig, *Biochemistry* **2021**, *60*, 999–1018.
[20] O. Francesconi, C. Nativi, G. Gabrielli, I. De Simone, S. Noppen, J. Balzarini, S. Liekens, S. Roelens, *Chem. Eur. J.* **2015**, *21*, 10089–10093.
[21] S.-H. Park, Y. P. Choi, J. Park, A. Share, O. Francesconi, C. Nativi, W. Namkung, J. L. Sessler, S. Roelens, I. Shin, *Chem. Sci.* **2015**, *6*, 7284–7292.
[22] S. Kubik, *Supramolecular Chemistry in Water*, Wiley, **2019**.
[23] H. Yao, H. Ke, X. Zhang, S. J. Pan, M. S. Li, L. P. Yang, G. Schreckenbach, W. Jiang, *J. Am. Chem. Soc.* **2018**, *140*, 13466–13477.
[24] M. Yamashina, M. Akita, T. Hasegawa, S. Hayashi, M. Yoshizawa, *Sci. Adv.* **2017**, *3*, 2–8.
[25] O. Francesconi, M. Martinucci, L. Badii, C. Nativi, S. Roelens, *Chem. Eur. J.* **2018**, *24*, 6828–6836.
[26] D. Zhang, A. Martinez, J.-P. Dutasta, *Chem. Rev.* **2017**, *117*, 4900–4942.
[27] T. Hayashi, Y. Ohishi, H. Abe, M. Inouye, *J. Org. Chem.* **2020**, *85*, 1927–1934.
[28] T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump, A. P. Davis, *Nat. Chem.* **2016**, *8*, 69–74.
[29] R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. v. Matlock, M. G. Orchard, A. P. Davis, *Nat. Chem.* **2019**, *11*, 52–56.
[30] B. J. J. Timmer, A. Kooijman, X. Schaapkens, T. J. Mooibroek, *Angew. Chem. Int. Ed.* **2021**, *60*, 16178–16183; *Angew. Chem.* **2021**, *133*, 16314–16319.
[31] Y. Ohishi, K. Masuda, K. Kudo, H. Abe, M. Inouye, *Chem. Eur. J.* **2021**, *27*, 785–793.
[32] P. Mateus, B. Wicher, Y. Ferrand, I. Huc, *Chem. Commun.* **2018**, *54*, 5078–5081.
[33] T. S. Carter, T. J. Mooibroek, P. F. N. Stewart, M. P. Crump, M. C. Galan, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 9311–9315; *Angew. Chem.* **2016**, *128*, 9457–9461.
[34] C. Nativi, O. Francesconi, G. Gabrielli, I. de Simone, B. Turchetti, T. Mello, L. D. C. Mannelli, C. Ghelardini, P. Buzzini, S. Roelens, *Chem. Eur. J.* **2012**, *18*, 5064–5072.
[35] O. Francesconi, F. Milanesi, C. Nativi, S. Roelens, *Chem. Eur. J.* **2021**, *27*, 10456–10460.
[36] O. Francesconi, F. Milanesi, C. Nativi, S. Roelens, *Angew. Chem. Int. Ed.* **2021**, *60*, 11168–11172; *Angew. Chem.* **2021**, *133*, 11268–11272.
[37] O. Francesconi, F. Cicero, C. Nativi, S. Roelens, *ChemPhysChem* **2020**, *21*, 257–262.
[38] J. L. Asensio, A. Ardá, F. J. Cañada, J. Jiménez-Barbero, *Acc. Chem. Res.* **2013**, *46*, 946–954.
[39] J. L. Asensio, F. J. Canada, M. Bruix, A. Rodriguez-Romero, J. Jimenez-Barbero, *Eur. J. Biochem.* **1995**, *230*, 621–633.
[40] R. L. Schnaar, *J. Leukocyte Biol.* **2016**, *99*, 825–838.
[41] S. Duan, J. C. Paulson, *Annu. Rev. Immunol.* **2020**, *38*, 365–395.
[42] M. P. Lenza, U. Atxabal, I. Oyenarte, J. Jiménez-Barbero, J. Ereño-Orbea, *Cells* **2020**, *9*, 2691.
[43] Y. Kajihara, Y. Suzuki, N. Yamamoto, K. Sasaki, T. Sakakibara, L. R. Juneja, *Chem. Eur. J.* **2004**, *10*, 971–985.
[44] O. Francesconi, A. Ienco, C. Nativi, S. Roelens, *ChemPlusChem* **2020**, *85*, 1369–1373.
[45] A. Vacca, O. Francesconi, S. Roelens, *Chem. Rec.* **2012**, *12*, 544–566.
[46] a) K. Fujikawa, Y. Tsukamoto, T. Oki, Y. C. Lee, *Glycobiology* **1998**, *8*, 407–414; b) M. Enomoto, Y. Igarashi, M. Sasaki, H. Shimizu, *Tetrahedron* **2015**, *71*, 2603–2609; c) Y. Nakagawa, T. Doi, T. Taketani, k. Takegoshi, Y. Igarashi, Y. Ito, *Chem. Eur. J.* **2013**, *19*, 10516–10525.
[47] A. Ardá, J. Jiménez-Barbero, *Chem. Commun.* **2018**, *54*, 4761–4769.
[48] A. Canales, A. Mallagaray, J. Pérez-Castells, I. Boos, C. Unverzagt, S. André, H.-J. Gabius, F. J. Cañada, J. Jiménez-Barbero, *Angew. Chem. Int. Ed.* **2013**, *52*, 13789–13793; *Angew. Chem.* **2013**, *125*, 14034–14038.

Manuscript received: November 18, 2022
Accepted manuscript online: January 4, 2023
Version of record online: February 24, 2023