Chemistry
A European
Journal

Chemistry Europe
European Chemical Societies Publishing

Accepted Article

Title: Rapid On-chip Synthesis of Complex Glycomimetics from N-glycan Scaffolds for Improved Lectin Targeting

Authors: Niels Reichardt, Anna Cioce, Michel Thépaut, and Franck Fieschi

To be cited as: *Chem. Eur. J.* 10.1002/chem.202000026

Link to VoR: https://doi.org/10.1002/chem.202000026

WILEY-VCH

10.1002/chem.202000026

Rapid On-chip Synthesis of Complex Glycomimetics from N-glycan Scaffolds for Improved Lectin Targeting

Anna Cioce[1], Michel Thépaut[2], Franck Fieschi[2], Niels-Christian Reichardt*[1,3,4]

[1] CIC biomaGUNE, Paseo Miramón 182, 20009 San Sebastian, Spain.
[2] Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, 38100 Grenoble, France.
[3] CIBER-BBN, Paseo Miramón 182, 20009 San Sebastian, Spain.
[4] Basque Research and Technology Alliance (BRTA), Paseo Miramón 182, 20009 San Sebastian, Spain.

nreichardt@cicbiomagune.es

Abstract:

C-type lectin receptor (CLR) carbohydrate binding proteins found on immune cells with important functions in pathogen recognition and self and non-self-differentiation are increasingly moving into the focus of drug developers as targets for the immune therapy of cancer autoimmune diseases and inflammation and to improve the efficacy of vaccines. The development of molecules with increased affinity and selectivity over the natural glycan binders has largely focused on the synthesis of mono and disaccharide mimetics but glycan array binding experiments have shown increased binding selectivity and affinity for selected larger oligosaccharides that are able to engage in additional favorable interactions beyond the primary binding site. Here we present a platform for the rapid preparation and screening of N-glycan mimetics on microarrays that turns a panel of complex glycan core structures into structurally diverse glycomimetics by a combination of enzymatic glycosylation with a non-natural donor and subsequent cycloaddition with a collection of alkynes. All surface based reactions were monitored by MALDI-Tof MS to assess conversion and purity of spot compositions. Screening the collection of 374 N-glycomimetics against the plant lectin WFA and the 2 human immune lectins MGL and Langerin produced a number of high affinity binders as lead structures for more selective lectin targeting probes.

INTRODUCTION

N-glycosylation is a complex posttranslational modification of asparagine residues within the consensus sequence Asn-X-Ser/Thr that affects folding, stability, half-life and often the functions of eukaryotic proteins. While the biosynthesis of the common Man9-N-glycan precursor in the endoplasmic reticulum is tightly controlled by a lectin based feedback cycle, the further glycan processing in the Golgi depends on expression levels of glycosyltransferases, glycosidases and transporter proteins, concentration of sugar nucleotide donors and the steric accessibility of the protein glycosylation sites.[1] This interplay of physiological and structural factors gives rise to a heterogeneous N-glycome covering hundreds of possible structures with subtle variations in the

number and type of antennae, terminal sugar residues and core modifications. Together with glycolipids and negatively charged matrix polysaccharides, like glycosaminoglycans, mammalian glycoproteins form a dense coat of sugars on the cell surface termed the glycocalyx. They can also be secreted into the plasma and engage with carbohydrate binding proteins (GBPs) in specific molecular recognition events to modulate biological processes like cell adhesion, differentiation of self and non-self by receptors of the innate immune system or glycoprotein quality control and turnover.[1] C-type lectin receptors (CLRs) are a class of pattern recognition receptors located on the cell surface of various mammalian immune cells that among many other functions play an important role in pathogen recognition, antigen uptake or leukocyte homing.[2,3] Langerin, expressed on Langerhans cells, is an innate viral defense receptor[4] and is involved in the recognition and response to pathogens as HIV-1[5], *Mycobacterium leprae*[6] and *Candida albicans.*[7] In addition, the interaction of aberrant carbohydrate antigens expressed on tumor cells with the CLRs MGL, DC-SIGN or dectin-1 on macrophages and dendritic cells can lead to tolerogenic immune responses, which help tumor cells evade the attack from the immune system[8–10] In summary, lectins in general, and mammalian CLRs in particular, constitute an increasingly interesting group of targets for therapeutic approaches in the treatment of cancer, autoimmune and infectious diseases.

Due to their high polarity, rapid renal clearance rates and low stability against enzymatic degradation, glycans show a poor pharmacokinetic profile that complicates the development of small molecule drugs from natural carbohydrate leads. In addition, $K_{d(s)}$ for the monovalent carbohydrate protein interactions are usually in the low millimolar to micromolar range, around 3-fold lower than the affinity usually required for a pharmaceutically drug lead. Nevertheless, potent carbohydrate-based lectin antagonists have been developed for human E and P-selectin, galectin-3 and bacterial lectins like FimH or PA-I as inhibitors of biofilm formation demonstrating the general druggability of lectins with small molecules.[11] Together with the Bernardi group we had recently shown how microarrays can be used to screen small glycomimetics in a more systematic approach, which has allowed us to identify compounds presenting selectivity for Dectin-2 over DC-SIGN C-type lectins. However in this case each glycomimetic had to be synthesized individually in solution prior to its immobilization onto the array for analysis.[12]

Larger oligosaccharides, that present multiple copies of the same recognition motif and can make additional contacts with the lectin along an extended binding site, often show higher affinities than simple monosaccharides. With mannose binding lectins, Wong et al. showed three orders of magnitude higher binding affinity for the Man9 oligosaccharide than for the monovalent methylmannoside, which is similar to the affinity gain observed for dendrimer presentation of monosaccharidic ligands[13], and the Paulson group recently demonstrated up to 1500-fold increase in affinity for selected Siglec inhibitors presented on the antenna of a triantennary N-glycan compared to the monovalent variant.[14]

The presentation of particular binding motifs in the context of larger glycans can lead to an increase in selectivity for lectins with similar binding specificity. Indeed, in previous glycan array binding experiments with CLRs LSECtin, DC-SIGN and DC-SIGNR we showed a strong preference for

binding one positional bi-antennary N-glycan isomer over the other, suggesting a functional role for asymmetrically branched glycans.[15] We therefore envisaged the use of large oligosaccharide N-glycans as scaffolds for the development of a glycomimetics library towards CLRs by selective and localized chemical modifications on terminal or core residues. As the synthesis of large multiantennary glycans can include over 100 synthetic steps we looked into miniaturizing the glycomimetics synthesis and subsequent screening methodology by parallel on-chip synthesis employing our library of synthetic N-glycans as a scaffold collection.

We have previously studied the enzymatic on-chip synthesis of glycan libraries using recombinant glycosyltransferases and pairs of lectins with complimentary binding specificities to monitor the conversion of the on-chip reactions. [16] Later, we and others developed conductive array surfaces which facilitated the analysis of on-chip reactions by MALDI-TOF MS and employed these for screening the substrate specificities of glycosyltransferases and glycosidases.[17–20] The Paulson group has prepared sialoside mimetics for array-based binding assays using non-natural sialic acid derivatives to include alkyne moieties that were further diversified via copper mediated cycloaddition.[21,22] The synthesis of sialoside mimetics was further streamlined by performing the copper catalyzed triazole formation directly on the chip and screening the sialoside library for binders against siglec-7.[23] However, the conversion for the on-chip triazole formation was not established by mass spectrometry. Instead, the authors probed the chip for unreacted alkyne functions via cycloaddition with a bulky azide functionalized dye, which is unlikely to react where other smaller azides have failed to do so, potentially providing an incorrect picture of on-chip conversion.

We had previously developed a glycan microarray with a dual read-out option for analyzing on-chip transformations by mass spectrometry and interactions by fluorescence imaging, respectively.[16] Here, we describe the preparation of an improved version of this MALDI-TOF MS compatible array platform, its application to the on-chip chemo-enzymatic synthesis of a novel class of N-glycomimetics and the identification of high affinity ligands for the plant lectin *Wisteria floribunda* and the two human immune lectins *Macrophage Galactose Lectin (MGL)* and *Langerin*.

RESULTS AND DISCUSSION

Figure 1 shows an overview of the strategy employed for the on-chip synthesis of microarrays of N-glycomimetics and the substrates employed. Twenty different glycan structures **(G1-G20)** with variations in branching pattern, terminal sugar residues and core modifications were first printed on the NHS-activated hydrophobic surface, then enzymatically glycosylated with non-natural azido-N-acetyl-D-galactosamine residue and finally further derivatized by copper-catalyzed azide alkyne cycloaddition (CuAAC) with nineteen functionalized-alkynes (a1-a19). Finally the 323 new triazole based N-glycan mimetics were screened against different C-type lectins such as WFA, Langerin extracellular domain (langerin ECD) and MGL extracellular domain (MGL ECD). All on-chip transformations were carried out under aqueous conditions that are compatible with the immobilization strategy via hydrophobic interactions and were monitored by MALDI-TOF MS.

10.1002/chem.202000026

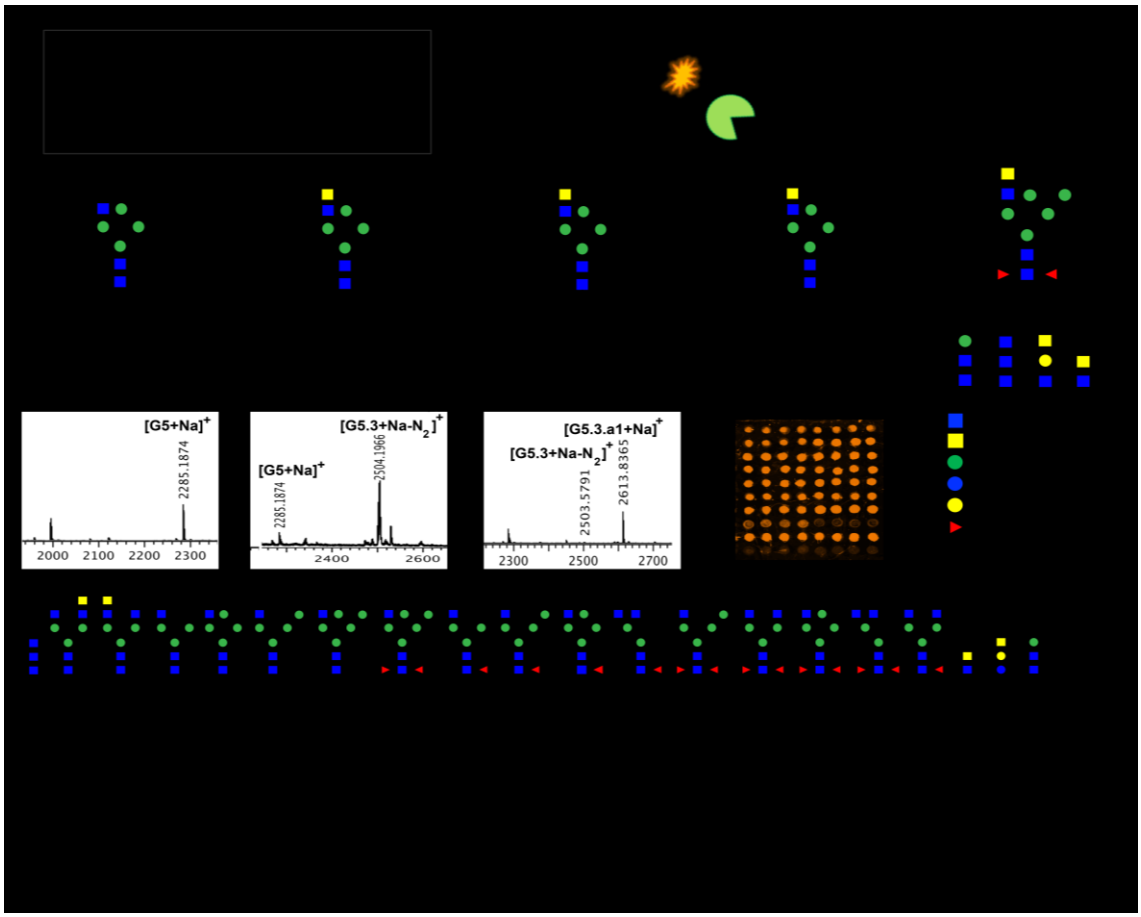

Figure 1 a) Overall strategy and linker employed for the on-chip preparation of N-glycan mimetics. **b**) selection of N-glycan scaffolds and **c)** panel of alkynes employed in the on-chip CuAAC.

Preparation of NHS-hydrophobic ITO-slides and on-chip immobilization of N-glycans.

For the preparation of the ready-to print NHS-activated hydrophobic slide we cleaned a commercial indium tin oxide glass slide with hot piranha solution and functionalized the activated surface by immersion in an octadecylphosphonic acid solution and subsequent thermal annealing to furnish a highly hydrophobic surface with a contact angle of 113° (see Figure S2, S.I.). A homogenous functionalization of the hydrophobic slide with the bidentate NHS-activated carbamate linker **1** was achieved by repeated vibrational vaporization, employing instrumentation used in sample preparation for MALDI imaging (S.I.).

From our library of synthetic glycan structures, compounds **G1** to **G17** presenting terminal N-acetylglucosamine moieties and variations in antennae structure and core modifications were chosen as scaffolds for the enzymatic and chemical diversification into glycomimetics. Higher branched glycans with 3 or more terminal GlcNAc residues were not included on the arrays as we had observed substantially lower reactivity in on-chip enzymatic transformations in preliminary assays. Since the 323 new triazole based N-glycan mimetics were synthesized on different slides, the printing of internal standards (**G18-G20**) alongside the scaffolds, for the normalization of the fluorescent read-out, was needed.

The printing conditions were carefully adjusted to achieve a sufficient ligand surface concentration for the readout with good signal strength by MALDI-TOF MS while avoiding strong background due to excessive ligand deposition in the 3 orders of magnitude more sensitive fluorescence detection.

On-chip synthesis

For the introduction of the azido N-Acetyl-D-galactosamine residues onto the printed glycan scaffolds we employed a double mutant β1,4-galactosyltransferase developed by Qasba et al. that accept UDP-GalNAc and UDP-GalNAz donors as substrates.[24] The conversion for all individual enzymatic elongations was determined by MALDI-TOF MS as the ratio of the product peak intensity to the sum of peak intensities of remaining starting material and product. We improved reaction conditions for on-chip elongation with UDP-GalNAz by performing 2 cycles of incubation with the enzyme for 8 hours by using HEPES over TRIS-buffer which reduced the amount of precipitating enzyme. Under these conditions, we observed conversion between 50-90% for the enzymatic addition of GalNAz to Glycans **G1.3-G17.3** (see figure S4, S.I.) alongside some 10% of mono-elongated for bi-antennary glycans and unreacted starting material. The incomplete elongation after two cycles is in line with previously observed surface based elongations using glycosyltransferases and might be explained by the limited access of enzymes and reagents to the glycans printed at high concentration.

In addition, incubation of subarrays with UDP-GalNAc and UDP-galactose with human GalT1 double mutant and bovine GalT, respectively, was performed to produce the series of GalNAc extended **(G1.1-G17.1)** and galactosylated derivatives **(G1.2-G17.2)** as natural benchmark substrates in the subsequent lectin binding assays. Again, the enzymatic extensions proceeded in good yield but the analysis of spot compositions by MALDi-TOF MS showed that conversion showed some variability with yields between 44-89% for the elongation with GalNAc and 26-87% for the elongation with galactose (see Figure S4, S.I.). A total of 51 compounds were added to the printed scaffolds at this stage via on-chip enzymatic synthesis.

Due to its compatibility with aqueous reaction conditions, excellent chemoselectivity and fast reaction kinetics, on-chip Cu(I) catalyzed alkyne-azide cycloaddition (CuAAC) was our method of choice for the rapid generation of a large number of structurally diverse glycomimetics. 29 structurally diverse alkynes (see Figure S5, S. I.) were coupled to the azide-functionalized N-glycans (**G1.3-G17.3**) with addition of sodium ascorbate and tris-hydroxypropyltriazolylmethylamine (THPTA) to form and stabilize intermediate Cu(I) species, respectively. For improved solubilization, alkynes were added from concentrated stock solutions in DMSO to the reaction mixture. MALDI-TOF MS analysis of the spot compositions revealed a surprisingly variable conversion for the Cu(I) catalyzed surface-based cycloaddition that depended both on alkyne and glycan structures. 10 alkynes reacted very sluggishly with yields below 20% and were removed altogether from the further synthetic strategy.

Analyzing the spot compositions by mass spectrometry was critical for revealing this reagent dependent variability and would have likely remained unnoticed by fluorescent methods which had been employed previously for measuring surface based conversion.[25]

10.1002/chem.202000026

Using the optimized protocols for both on-chip enzymatic elongation and copper catalyzed cycloadditions, we prepared various slides comprising a total of 394 ligands including glycomimetics and natural reference glycans. All reaction products were analyzed on duplicate subarrays that had been reserved for analysis by MALDI-TOF MS. In general good conversions were obtained in both enzymatic and cycloaddition steps for most compounds leading to the formation of immobilized ligands of generally good purity. Any interference in the binding assays, by unreacted starting materials or intermediates, was expected to be low and accountable as all starting scaffolds, natural references and azido intermediates had been included as control ligands in the assays.

Lectin binding assays

We chose *Wisteria floribunda Agglutinin* (WFA), a plant lectin which binds strongly to the cancer biomarker GalNAcβ1,4-GlcNAc (**G18**) and with lower affinity to galactose and α- and β-N-Acetyl-D-galactosamine residues,[26] as a first probe to optimize the incubation conditions and assess the binding properties of our array of novel N-glycan mimetics.

The fluorescence value for the interaction with glycan **G18** was arbitrarily set to 100 and used as internal standard to normalize the fluorescent intensity values across the different slides employed in the analysis. Figure 2 shows the binding pattern of WFA with all 394 ligands, organized by type of structure and alkyne, employed in the cycloaddition and normalized binding intensities displayed in the form of a heatmap. In line with the known binding specificity of WFA we observed binding close to the normalized threshold **(G18)** to structures **G2** and **G3** presenting terminal GalNAc residues with the exception of the trisaccharide GalNAcβ1-4Galβ1-4GlcNAcβ-sp (**G19**) that showed below threshold binding to WFA Galactosylation in general did not improve binding to WFA substantially with the exception of scaffold **G2**, a bi-antennary N-glycan with a GalNAc residue on the 6 arm, which bound strongly above threshold to WFA after galactosylation of the 3-arm (**G2.2**). WFA bound to structures presenting GalNAz and GalNAc residues with a similar strength but not to galactosylated structures, suggesting that a 2-acetamide moiety is required for efficient binding to WFA but slightly larger structures can be accommodated in the WFA binding site. The interaction of WFA with the panel of triazole based glycomimetics showed a differentiated pattern with a number of high affinity hits showing a 2-3 fold increase in binding over the reference compound **G18**. With respect to the parent azide structure, we observed a strong alkyne dependent affinity change after cycloaddition. For example, binding of WFA for GalNAz compound **G8.3** with a normalized RFU of 130 was reduced to an RFU of 48 after coupling with 2-ethynyl pyridine (**a2**) to afford compound **G8.3.a2**. However, when azido-derivative **G8.3** was coupled with 5-hexyn-3-ol (**a15**), to obtain compound **G8.3.a15**, binding to WFA increased nearly 3-fold to 283 RFU.

Figure 2 highlights 6 strongest WFA binders with normalized RFUs above 200; strikingly, various unique combinations of parent structures and alkyne substituents gave rise to glycomimetics with an unpredictable improvement in binding affinity. These compounds are based on 5 N-glycan scaffolds and 3 different alkynes. Notably, in both cases coupling of scaffold **G11** with cyclopropylacetylene (**a10**) and 5-hexyn-3-ol (**a15**) produced triazoles with high affinity to WFA.

10.1002/chem.202000026

Cycloaddition with 5-hexyn-3-ol also produced strong WFA binders with scaffold structures **G8** and **G12** while cyclopropylacetylene (**a10**) derived triazoles presented on either **G1** or **G11** resulted in excellent ligands for WFA.

While scaffold **G4** is structurally quite similar to **G11** or **G8** with one or two additional mannose residues, only the triazole resulting from the cycloaddition of **G4** with tert-butylalkyne **a13** produced a glycomimetic with a 2-fold increase in binding strength over the reference compound. Interestingly, a pairwise comparison of glycomimetics based on the scaffolds **G7/G8** and **G12/G16**, respectively, show the effect of a single or two core fucose residues that are far from the primary binding site on recognition by the WFA lectin. The effect of core fucose on N-glycan conformation, which favors an extended orientation of the 1,6-arm and its direct consequences for lectin recognition have been described previously by several authors.[27–30]

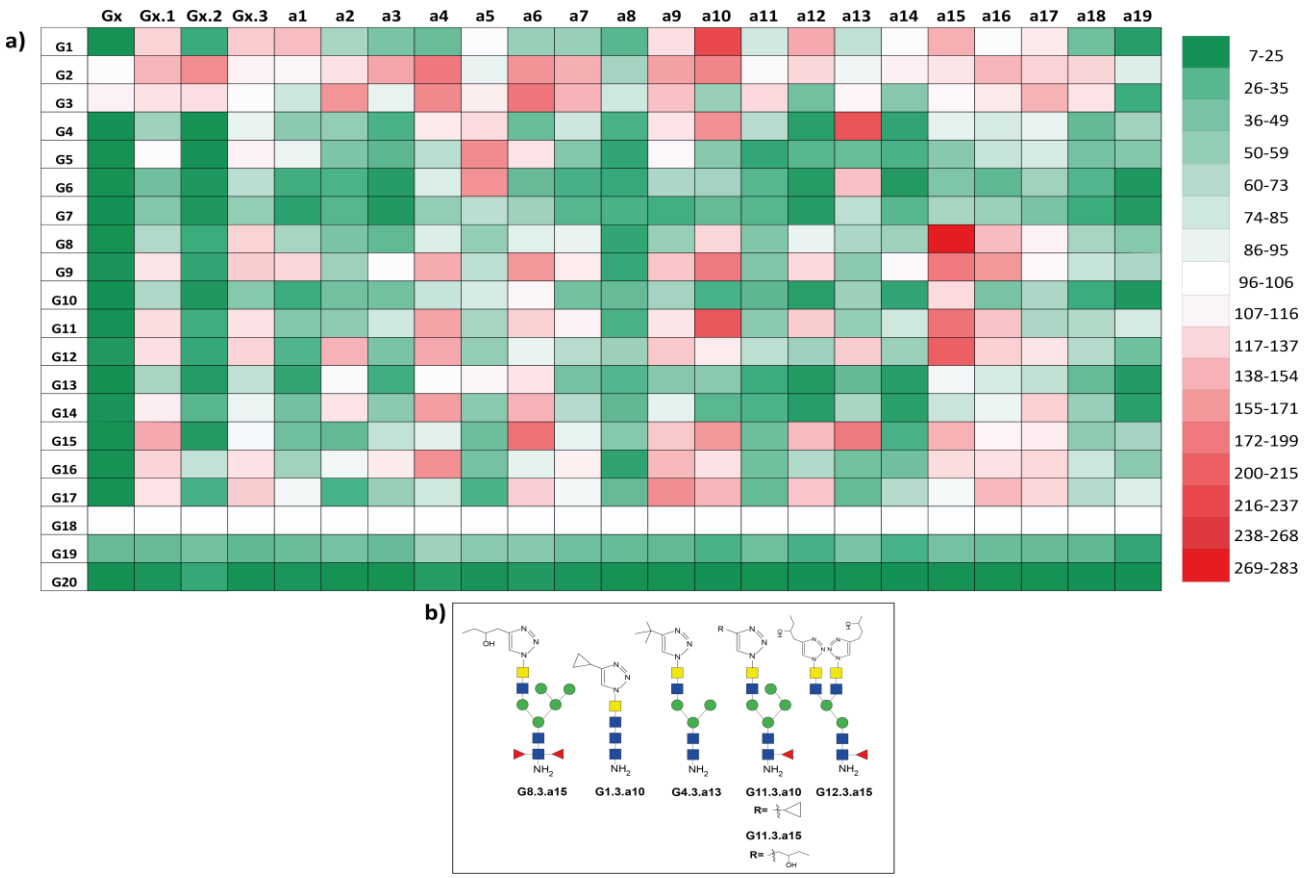

Figure 2 a) Heatmap representation of the binding data for 394 ligands with a GalNAc specific lectin from *Wisteria floribunda (WFA)*. Color-coded fluorescence value scale is represented on the right side of the table. Compound (**G18**) used as internal standard to normalize fluorescent values. Gx= starting glycans, Gx.1= GalNAc-derivatives, Gx.2= Gal-derivatives, Gx.3= GalNAz-derivatives. **b)** Selected high affinity ligands for WFA identified in the screen.

Macrophage galactose binding lectin (MGL-ECD)

We then studied the interaction of the macrophage galactose binding lectin (MGL-ECD), a human lectin with narrow binding specificity for terminal GalNAc-residues, with the 394 immobilized

10.1002/chem.202000026

structures choosing **G19** as internal reference for normalizing binding intensities over different slides.[8,31]

MGL is primarily found on macrophages and immature dendritic cells and thought to intervene in the recognition of self-gangliosides, tumor antigens, helminthic pathogens, but also Gal residue in LPS outer core of pathogenic *E.coli* as shown recently.[32,33] Its strong selectivity for the Tn antigen found on MUC1 in various solid tumors has prompted interest in understanding MGL mediated glycan antigen recognition for the development of anti-cancer immune therapies.[34]

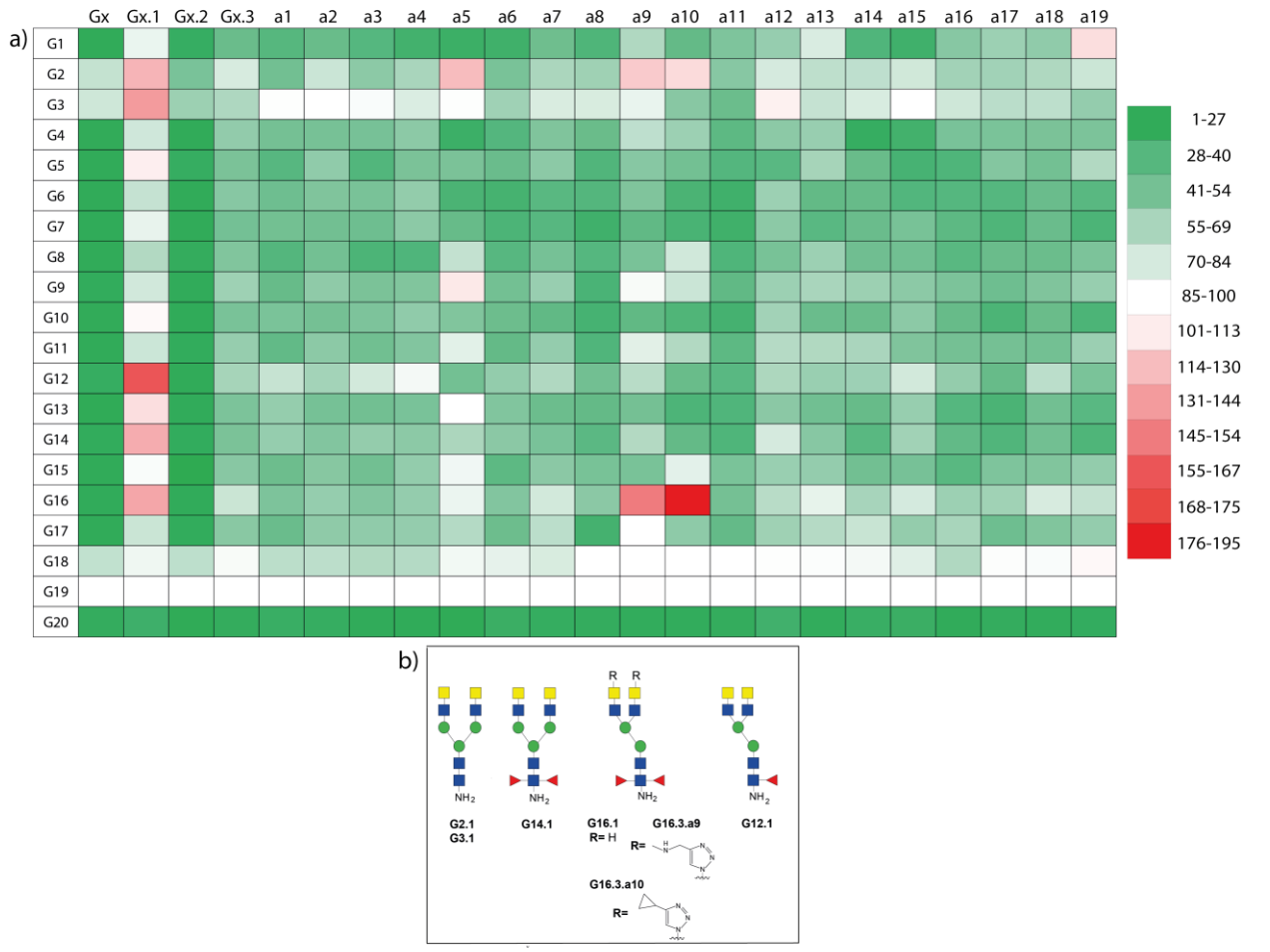

Figure 3 a) Heatmap representation of the binding data for 394 ligands with the GalNAc specific human macrophage galactose binding lectin (MGL ECD). Color-coded fluorescence value scale is represented on the right side of the table. Compound (**G19**) was used as internal standard to normalize fluorescent values. Gx= starting glycans, Gx.1= GalNAc-derivatives, Gx.2= Gal-derivatives, Gx.3= GalNAz-derivatives **b)** Selected high-affinity glycomimetics for MGL ECD identified in the screen.

In line with the carbohydrate binding specificity for MGL ECD (extracellular domain), we found moderate to strong binding only for compounds presenting terminal GalNAc residues either in printed scaffolds or after enzymatic elongation with GalNAc (**G1.1-G17.1**). **G12.1** showed the strongest binding among the natural substrates. Enzymatic elongation of scaffold structures with GalNAz (**G1.3-G17.3**) did not improve binding to MGL as frequently as observed for WFA above. The cycloaddition of GalNAz scaffolds with the panel of alkynes however, did produce a number of high affinity binders, which are summarized in **Figure 3b**.

The lower number of hits in the library screen might be explained with a more selective binding mode for MGL which recognizes predominantly terminal GalNAc residues and might allow small substitutions in C6 but not C3 and C4.[35] As the best binder we identified the bi-antennary cyclopropyl derivative **G16.3.a10** with a relative fluorescence of RFU=194; the non-natural branching pattern of **G16** lacking substitution on C6 of the β-mannose was present in 4 out of 7 strong MGL binders.

Bis-substitutions on the **G16** and **G12** scaffolds were tolerated by the lectin and improved binding while substitutions on the scaffolds **G2**, **G3** and **G14** with the natural bi-antennary branching pattern did not lead to compounds with higher affinity for MGL ECD than the native glycans.. On the basis of the high selectivity of the primary binding site of MGL, future development of MGL inhibitors should address modifications in C6 of the GalNAc residue or on other residues in the oligosaccharide that could interact favorably with amino acid residues of an extended or secondary binding site of the lectin.

Langerin ECD

Langerin is a C-type lectin receptor that mediates the pathogen recognition and antigen uptake of Langerhans cells by binding to a wide variety of D-mannose, L-fucose and D-*N*-acetylglucosamine (GlcNAc) containing carbohydrate surface glycans including bacterial, fungal and viral oligosaccharides.[7,36]

To identify novel high affinity binders of Langerin we screened our glycomimetics library against the full extracellular domain (ECD), which is a trimeric lectin.[37] Ligand **G20**, a trisaccharide that had shown very high binding affinity on our glycan array was used as internal standard for normalization of fluorescence values between different wells and slides. Fluorescence values for the glycan scaffold **G20** were set to 100. **Figure 4** summarizes the Langerin binding profile in form of a heatmap and highlights the best binders on the array.

10.1002/chem.202000026

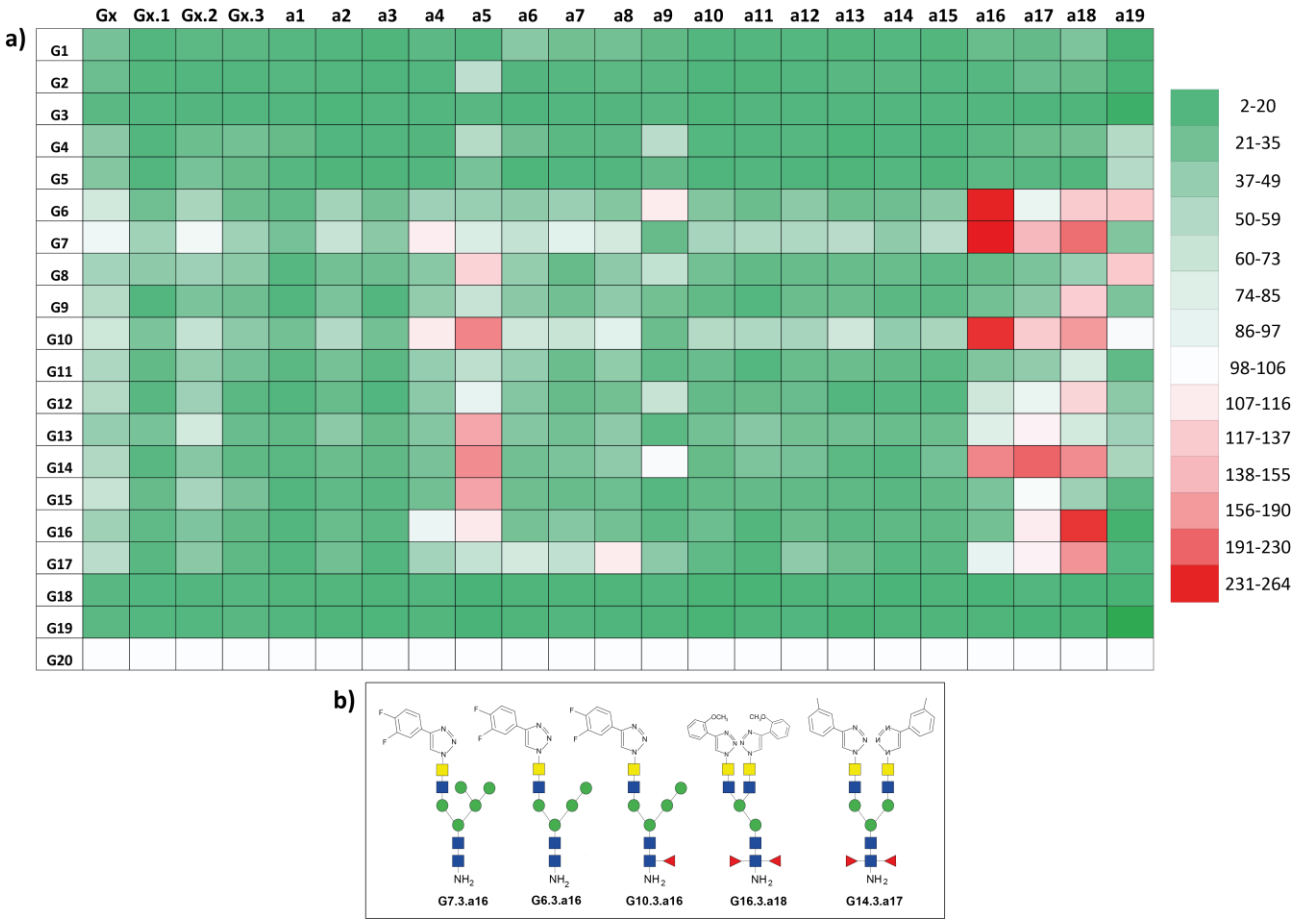

Figure 4 a) Heatmap representation of the binding data for 394 ligands with Langerin ECD, a mannose specific human C-type lectin receptor. Color-coded fluorescence value scale is represented on the right side of the table. Compound **G20** was used as internal standard to normalize fluorescent values. Gx= starting glycans, Gx.1= GalNAc-derivatives, Gx.2= Gal-derivatives, Gx.3= GalNAz-derivatives. **b)** Selected high-affinity glycomimetics for Langerin identified in the screen.

From the unmodified scaffold structures **G1-G17** only the hybrid glycan structures **G6, G7** and **G10** with accessible terminal mannose residues showed moderate, close to threshold **(G20)** binding to Langerin. As anticipated from the reported Langerin binding specificity, the enzymatic extension of scaffold structures with terminal galactose, GalNAc or GalNAz residues did not improve or even reduce binding to Langerin. On-chip derivatization of azide modified glycans **G1.3-G17.3** with the panel of 19 alkynes, however, produced a considerable number of moderate to high affinity binders. In particular, the coupling with aromatic alkynes as **a5**, **a16**, **a17** and **a18** resulted in the formation of a number of extraordinary strong binding ligands. The combinations of scaffolds and alkynes that resulted in the formation of triazole derivatives with the highest binding strength on the array (RFU>200) have been summarized in **Figure 4**. As major Langerin binders we identified compounds **G7.3.a16, G6.3.a16** and **G10.3.a16** that shared the 3,4-difluorophenyl acetylene **a16**, and **G16.3.a18** and **G14.3.a17** incorporating aromatic alkynes **a18** and **a17**, respectively (**Figure 4b**).

Unlike WFA and MGL the derivatization of the GalNAz residues with aromatic alkynes did not target interaction with the primary binding site, but the observed improved binding to Langerin

was likely due to additional favorable contacts with adjacent secondary binding sites or along an extended binding groove.

CONCLUSIONS

In conclusion, we have developed the first synthesis of a large library of N-glycan mimetics as potent leads for targeting and inhibiting C-type lectin receptors via multiple points of interaction.
By moving the chemo-enzymatic derivatization of N-glycan scaffolds to a microarray platform we were able to perform the synthesis of a large number of highly complex compounds in parallel, and employing minute amounts of valuable reagents. *In situ* analysis of reaction products by surface based MALDI-TOF MS was important for improving reaction conditions and for assessing spot composition on the array. Screening of the 394 immobilized ligands for binders against the three lectins *Wisteria Floribunda Agglutinin* (WFA), Macrophage Galactose Lectin ECD (MGL) and Langerin ECD produced high affinity leads for all three lectins.
Binding data suggested that both modifications of the recognition elements targeting the primary binding site and on the N-glycan backbone can produce an increase in binding strength over the native structures.
To study the lectin binding event with molecular precision e.g. by NMR or X-ray co-crystallization studies, selected ligands will be resynthesized in solution on a preparative scale. Our initial studies have shown that the on-chip synthesis of N-glycan mimetics is an exciting new approach for rapidly generating collections of structurally very complex glycan mimetics. These compounds show potentially far higher affinity than smaller fragments as multiple structural elements in the oligosaccharide are available for binding to secondary binding sites or extended binding grooves. Future work will move this concept further and integrate other chemistries for the preparation of mimetics and streamline the on-chip synthesis to accommodate compounds on the chip with higher density.

EXPERIMENTAL PART

Functionalization of ITO-coated glass slides with octadecylphosphonic acid (ODPA)

A commercial ITO-coated glass slide (Type I, 1.1 mm/25 ea) was cleaned in piranha solution (70 mL H_2O, 10 mL NH_3, 10 mL H_2O_2) at 70ºC for 1 hour, washed with nanopure water and dried under a stream of air. Subsequently, the slide was incubated in a 1mM solution of octadecylphosphonic acid (ODPA) in THF at r.t. for 3 hours and left at 140ºC for 20 hours. Finally, it was sonicated in methanol for 10 minutes, left in the same solvent for further 30 minutes without sonication, dried under stream of air and kept under vacuum until use.[17]

NHS-activation of hydrophobic ITO-glass slides

After washing in acetone, the hydrophobic ITO-slide was placed into a vibrational vaporization instrument (IMAGEPrep©, Bruker), sprayed with a solution of the bidentate NHS-activated carbamate linker 1 (15 mg/mL in 20 mL $CHCl_3$:MeOH 1:1) for 3 min and dried under a stream of air. The procedure was repeated twice. Finally, the slide was sonicated in nanopure water for 5 min and air-dried.

Immobilization of N-glycan structures (G1-G20) on activated ITO slides

Stock solutions of synthetic C5-amino linked N-glycans (**G1** to **G20**) (50 µM in phosphate buffer 300 mM, pH 8.7) were prepared in a 384 multi-well plate. 50 drops of each glycan solution were dispensed with the help of a Scienion S11 microarray printer, in replicates of four spots for each subarray, onto 8 subarrays of a NHS-activated hydrophobic-ITO slide and left to react overnight at 18°C under controlled humidity of 75%. The slide was dried in an oven at 40°C for 24 hours and the unreacted NHS-linker was quenched by 30 min immersion in 50 mM ethanolamine solution (in 50 mM borate buffer pH 9.3). After drying the slide under a stream of air, DHB matrix was spotted on the printed arrays with the help of a microarrayer and the immobilisation analyzed by MALDI-TOF analysis.

On-chip enzymatic elongations

A general on-chip enzymatic protocol has been adopted for the preparation of diversified N-glycan derivatives. A double mutant β1,4-galactosyltransferase (DM-GalT1) was expressed in *E. coli* and purified as previously reported by Qasba et al.[38] UDP-GalNAz or UDP-GalNAc were used as substrates for on-chip elongations with DM-GalT1 for the preparation of compound series (**G1.3-G17.3)** and (**G.1.1-G17.1**), respectively. D-galactose derivatives (**G1.2-G17.2**) were synthesized with a commercial β1,4-galactosyltransferase (GalT-C342) and UDP-galactose.

For enzymatic elongations, the slides were divided into 6 subarrays and each subarray incubated with 50 µL of enzymatic solution containing: freshly prepared HEPES 100 mM pH=7.4, 10 mM $MnCl_2$, 1 mM UDP-donor, 0.5 mg/mL enzyme, 0.2% BSA, 0.5 µL alkaline phosphatase 1 U/µL.The reaction was performed in 2 cycles of incubation at 37°C, 8 hours for each cycle; after each reaction cycle, the slide was washed with aqueous solution containing 0.1% TFA and 0.05% AcCN.

DHB matrix, containing 0.01 mM of sodium citrate, was spotted on the immobilized glycans and the conversion for each substrate was estimated by performing MALDI-TOF analysis.

On-chip Copper(I)-catalyzed Azide Alkyne Cycloaddition (CuAAC)

The slide was divided into subarrays employing a 16-well proplate® Module/6x7 mm gasket and each subarray containing azido-N-glycans **(G1.3-G17.3)** was treated with 100 µL of a solution containing: 0.058 mM $CuSO_4$, 0.29 mM THPTA, 0.58 mM alkyne in DMSO, bicarbonate buffer 300 mM, pH=9 and 0.58 mM sodium ascorbate. The reaction was left overnight at RT and the slide was then washed with water and dried under a stream of air. After spotting of DHB matrix in acetonitrile, containing 0.01 mM of sodium citrate onto selected spots, conversion to triazoles was evaluated by MALDI-TOF MS.

On-chip binding assays of N-glycan mimetics library against fluorescently labeled lectins

A total of 4 ITO slides were printed to accommodate all 394 ligands in an array layout of 80 spots/subarray, elongate scaffolds with azido-galactose and react the newly introduced azide function in separate wells with 19 different alkynes. Binding properties of novel ligands were evaluated by incubating individual sub-arrays with a labeled lectin of interest followed by

fluorescence imaging of the slide on a fluorescence scanner. The comparison of fluorescence intensities detected on different slides or different subarrays within the same slide can be challenging since reaction and washing cycles performed on the slide can introduce variation in the measurements difficult to control. To overcome this issue, fluorescent signals were normalized to an internal standard.

We used glycans **G18** (for WGA), **G19** (for MGL) and **G20** (for Langerin) as internal standards and assigned them a standard value of 100. The three compounds were not substrates for both enzymes employed in the elongations. After choosing a standard value of 100 to calculate the normalization factor, the normalization was performed for all of our slides by using the same reference sample in all the sub-arrays.

The subarrays containing immobilized N-glycan-based ligands were incubated with fluorescently labelled lectins (10 µg/mL) in buffer (Tris-HCl 100 mM, pH=7.5 with 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.2% BSA); the incubation was left 1h in the dark at R.T. for *WFA-Alexafluor647* or 1 h at 4°C for *MGL ECD-Alexafluor555* and *Langerin ECD-Cy3*. Arrays were washed with water, dried and scanned at 10 µm resolution.

ACKNOWLEDGEMENTS

For MGL-ECD and Langerin-ECD production, this work used the Multistep Protein Purification Platform (MP3) of the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI Grant ANR-10-INSB-05-02 and GRAL Grant ANR-10- LABX-49-01 within the Grenoble Partnership for Structural Biology. A.C., N.C.R. and F.F. were supported by the EU Horizon 2020 Research and Innovation Program (Marie Sklodowska-Curie Grant 642870, ETN-Immunoshape). N.C.R. additionally acknowledges funding from the Ministry of Science and Education (MINECO) Grant No. CTQ2017-90039-R and RTC-2017-6126-1 and the Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency – Grant No. MDM-2017-0720. F.F. additionally acknowledges funding from French Agence Nationale de la Recherche (ANR) PIA for Glyco@Alps (ANR-15-IDEX-02). We thank Alvaro Hernandez, Asparia Glycomics for assistance with the optimization of enzymatic glycosylations, Anna Bernardi and Giulio Goti, University of Milan for helpful discussions and help with initial cycloaddition experiments and Sonia Serna, CIC biomaGUNE for help with figures.

[1] P. Stanley, N. Taniguchi, M. Aebi, *Chapter 9. N-Glycans, Essentials of Glycobiology, 2nd Edition*, Cold Spring Harbor Laboratory Press, **2017**.

[2] W. I. Weis, M. E. Taylor, K. Drickamer, *Immunol. Rev.* **1998**, *163*, 19–34.

[3] Y. van Kooyk, *Biochem. Soc. Trans.* **2008**, DOI 10.1042/BST0361478.

[4] M. Van Der Vlist, T. B. H. Geijtenbeek, *Immunol. Cell Biol.* **2010**, *88*, 410–415.

[5] L. De Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. W. P. De Jong, T. De Gruijl, V. Piguet, Y. Van Kooyk, T. B. H. Geijtenbeek, *Nat. Med.* **2007**, *13*, 367–371.

[6] H. J. Kim, P. J. Brennan, D. Heaslip, M. C. Udey, R. L. Modlin, J. T. Belisle, *J. Bacteriol.* **2015**,

197, 615–625.

[7] H. Feinberg, M. E. Taylor, N. Razi, R. McBride, Y. A. Knirel, S. A. Graham, K. Drickamer, W. I. Weis, *J. Mol. Biol.* **2011**, *405*, 1027–1039.

[8] I. G. Zizzari, C. Napoletano, F. Battisti, H. Rahimi, S. Caponnetto, L. Pierelli, M. Nuti, A. Rughetti, *J. Immunol. Res.* **2015**, *2015*, 1–8.

[9] S. Manicassamy, B. Pulendran, *Immunol. Rev.* **2011**, *241*, 206–227.

[10] K. Upchurch, S. Oh, H. Joo, *Recept. Clin. Investig.* **2016**, *3*, 1–7.

[11] B. Ernst, J. L. Magnani, *Nat. Rev. Drug Discov.* **2009**, *8*, 661–77.

[12] L. Medve, S. Achilli, S. Serna, F. Zuccotto, N. Varga, M. Thépaut, M. Civera, C. Vivès, F. Fieschi, N. Reichardt, et al., *Chem. - A Eur. J.* **2018**, *24*, 14448–14460.

[13] P.-H. Liang, S.-K. Wang, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 11177–84.

[14] W. Peng, J. C. Paulson, *J. Am. Chem. Soc.* **2017**, *139*, 12450–12458.

[15] B. Echeverria, F. Fieschi, J. Pham, S. Achilli, C. H. Hokke, N.-C. Reichardt, M. Thépaut, C. Vivès, S. Serna, *ACS Chem. Biol.* **2018**, *13*, 2269–2279.

[16] S. Serna, J. Etxebarria, N. Ruiz, M. Martin-Lomas, N.-C. Reichardt, *Chemistry* **2010**, *16*, 13163–75.

[17] A. Beloqui, J. Calvo, S. Serna, S. Yan, I. B. H. Wilson, M. Martin-Lomas, N. C. Reichardt, *Angew. Chem. Int. Ed.* **2013**, *52*, 7477–7481.

[18] A. Beloqui, A. Sanchez-Ruiz, M. Martin-Lomas, N.-C. Reichardt, *Chem. Commun. (Camb).* **2012**, *48*, 1701–3.

[19] C. J. Gray, A. Sánchez-Ruíz, I. Šardzíková, Y. A. Ahmed, R. L. Miller, J. E. Reyes Martinez, E. Pallister, K. Huang, P. Both, M. Hartmann, et al., *Anal. Chem.* **2017**, *89*, 4444–4451.

[20] S. Y. Tseng, C.-C. Wang, C.-W. Lin, C.-L. Chen, W.-Y. Yu, C.-H. Chen, C.-Y. Wu, C.-H. Wong, *Chem. Asian J.* **2008**, *3*, 1395–405.

[21] C. D. Rillahan, E. Schwartz, R. McBride, V. V. Fokin, J. C. Paulson, *Angew. Chem. Int. Ed.* **2012**, *51*, 11014–11018.

[22] O. Blixt, S. Han, L. Liao, Y. Zeng, J. Hoffmann, S. Futakawa, J. C. Paulson, *J. Am. Chem. Soc.* **2008**, *130*, 6680–6681.

[23] C. D. Rillahan, E. Schwartz, C. Rademacher, R. McBride, J. Rangarajan, V. V. Fokin, J. C. Paulson, *ACS Chem. Biol.* **2013**, *8*, 1417–1422.

[24] N. Mercer, B. Ramakrishnan, E. Boeggeman, L. Verdi, P. K. Qasba, *Bioconjug. Chem.* **2013**, *24*, 144–152.

[25] C. D. Rillahan, E. Schwartz, C. Rademacher, R. Mcbride, J. Rangarajan, V. V Fokin, J. C. Paulson, **2013**, *9*.

[26] T. Kurokawa, M. Tsuda, Y. Sugino, *J. Biol. Chem.* **1976**, *251*, 5686-5693.

[27] C. Unverzagt, S. André, J. Seifert, S. Kojima, C. Fink, G. Srikrishna, H. Freeze, K. Kayser, H. J. Gabius, *J. Med. Chem.* **2002**, *45*, 478–491.

[28] M. Takahashi, Y. Kuroki, K. Ohtsubo, N. Taniguchi, *Carbohydr. Res.* **2009**, *344*, 1387–1390.

[29] W. Nishima, N. Miyashita, Y. Yamaguchi, Y. Sugita, S. Re, *J. Phys. Chem. B* **2012**, *116*, 8504–8512.

[30] A. Gimeno, N. C. Reichardt, F. J. Cañada, L. Perkams, C. Unverzagt, J. Jiménez-Barbero, A. Ardá, *ACS Chem. Biol.* **2017**, *12*, 1104–1112.

[31] S. J. Van Vliet, E. Saeland, Y. Van Kooyk, *Trends Immunol.* **2008**, 83–90.

[32] S. J. van Vliet, E. van Liempt, E. Saeland, C. A. Aarnoudse, B. Appelmelk, T. Irimura, T. B. H. Geijtenbeek, O. Blixt, R. Alvarez, I. van Die, et al., *Int. Immunol.* **2005**, *17*, 661–669.

[33] M. M. Maalej, R. E. Forgione, R. Marchetti, F. B. Bulteau, M. T. Thepaut, R. Lanzetta, C. Laguri, J. Simorre, F. Fieschi, A. Molinaro, et al., *ChemBioChem* **2019**, cbic.201900087.

[34] N. Higashi, K. Fujioka, K. Denda-Nagai, S. I. Hashimoto, S. Nagai, T. Sato, Y. Fujita, A. Morikawa, M. Tsuiji, M. Miyata-Takeuchi, et al., *J. Biol. Chem.* **2002**, *277*, 20686–20693.

[35] F. Marcelo, N. Supekar, F. Corzana, J. C. Van Der Horst, I. M. Vuist, D. Live, G. J. P. H. Boons, D. F. Smith, S. J. Van Vliet, *J. Biol. Chem.* **2019**, *294*, 1300–1311.

[36] V. Porkolab, E. Chabrol, N. Varga, S. Ordanini, I. Sutkevičiute, M. Thépaut, M. J. García-Jiménez, E. Girard, P. M. Nieto, A. Bernardi, et al., *ACS Chem. Biol.* **2018**, *13*, 600–608.

[37] M. Thépaut, J. Valladeau, A. Nurisso, R. Kahn, B. Arnou, C. Vivès, S. Saeland, C. Ebel, C. Monnier, C. Dezutter-Dambuyant, et al., *Biochemistry* **2009**, *48*, 2684–2698.

[38] B. Ramakrishnan, P. S. Shah, P. K. Qasba, *J. Biol. Chem.* **2001**, *276*, 37665–37671.