ChemPlusChem

Review
doi.org/10.1002/cplu.202300598

Chemistry Europe
European Chemical Societies Publishing

www.chempluschem.org

Towards Biomimetic Recognition of Glycans by Synthetic Receptors

Francesco Milanesi,[a] Stefano Roelens,[a] and Oscar Francesconi*[a]

Carbohydrates are abundant in Nature, where they are mostly assembled within glycans as free polysaccharides or conjugated to a variety of biological molecules such as proteins and lipids. Glycans exert several functions, including protein folding, stability, solubility, resistance to proteolysis, intracellular traffic, antigenicity, and recognition by carbohydrate-binding proteins. Interestingly, misregulation of their biosynthesis that leads to changes in glycan structures is frequently recognized as a mark of a disease state. Because of glycan ubiquity, carbohydrate binding agents (CBAs) targeting glycans can lead to a deeper understanding of their function and to the development of new diagnostic and prognostic strategies. Synthetic receptors selectively recognizing specific carbohydrates of biological interest have been developed over the past three decades. In addition to the success obtained in the effective recognition of monosaccharides, synthetic receptors recognizing more complex guests have also been developed, including di- and oligosaccharide fragments of glycans, shedding light on the structural and functional requirements necessary for an effective receptor. In this review, the most relevant achievements in molecular recognition of glycans and their fragments will be summarized, highlighting potentials and future perspectives of glycan-targeting synthetic receptors.

1. Introduction

1.1. Roles of Glycans

Glycans are ubiquitous in nature where their roles range from structural functions to energy resources, up to information carriers as ligands for several classes of proteins.[1,2] Glycans can be structured as free polysaccharides, such as in cellulose, chitin, starch, and glycogen, or as glycoconjugates with lipids and the linkage to the aglycone, as in N- and O-glycans. Protein glycosylation is a main post-translational modification, with significant effects on protein folding, stability, solubility, resistance to proteolysis, intracellular trafficking, antigenicity, and recognition by carbohydrate-binding proteins.[3,4] Widely expressed on the surface of eukaryotic cells as constituents of the glycocalyx, glycans mediate a variety of events in cell–cell, cell–matrix and cell–molecule interactions, including the adhesion and infection by pathogenic microorganisms such as viruses and bacteria.[5,6] Di-, tri- and poly-branched glycans can show high structural complexity and, although generally only the terminal residues are the elective epitopes recognised by proteins,[7,8] the entire polysaccharide is responsible for the correct folding and orientation of these complex structures.[9] Carbohydrate-protein recognition is involved in an enormous number of physiological and pathological processes including cell–growth,[10] neural development,[11] immune system regulation,[8] tumour growth and metastasis,[12] and inflammation.[13] Moreover, carbohydrates are important diagnostic biomarkers in cancer.[14] Indeed, in tumour cells the glycosylation patterns of proteins changes with the expression of tumour-associated carbohydrate antigens (TACAs) such as Thomsen antigen (Tn), Thomsen-Friedenreich antigen (TF), Sialyl-Tn antigen (STn), sialyl-Lewis a and sialyl-Lewis x antigens.[7]

[a] Dr. F. Milanesi, Dr. S. Roelens, Prof. O. Francesconi
Department of Chemistry "Ugo Schiff", DICUS and INSTM
Università degli Studi di Firenze
Campus Sesto, 50019 Sesto Fiorentino, Firenze (Italy)
E-mail: oscar.francesconi@unifi.it

1.2. Carbohydrates recognition in nature

Carbohydrates can be rather tricky ligands to recognize. Because of their polar hydroxyl groups, carbohydrates appear at first glance to be a cluster of water molecules, and when solvated in an aqueous medium they establish strong hydrogen bonding interactions with the solvent.[15] Consequentially, water hampers the recognition by proteins not only by competing for the binding site, but also by strongly solvating the ligands. In addition, due to their high hydrophilicity, most carbohydrate-protein interactions occur at the interphase between the solvent and the protein surface.[16] All these phenomena contribute to the low binding affinities between carbohydrates and proteins.[17,18] In addition, selectivity is also a difficult task, since the most common monosaccharides differ only in the stereochemistry of a single stereogenic centre.

Due to the plethora of carbohydrate-mediated processes, the need for new tools recognizing carbohydrates has grown in recent years with the aim of studying and understanding the molecular basis of the recognition process, but also to develop new diagnostic and prognostic strategies.[19] Initially, lectins and antiglycan antibodies were the first choice used as carbohydrate-binding agents (CBAs) to tackle this task.[20,21] However, although they are currently extensively used in glycobiology, their therapeutic development is unfortunately hampered by their protein nature that gives poor stability under physiological conditions. In addition, the lack of specificity together with the apparent immunogenicity of the lectins and the difficulties in purifying the antibodies has further limited their use in therapy and imaging.[22]

1.3. Artificial carbohydrate-binding agents (CBAs)

To overcome the problems given by protein CBAs, in the last decades, artificial alternatives have been developed. To properly recognize a carbohydrate, an artificial CBA must address both polar interactions and hydrophobic effects and satisfy the correct orientations of binding groups.[19] Among the classes of artificial CBAs is possible to distinguish between those that use boronic acids to bind sugars, which forms strong and reversible covalent bonds with 1,2- or 1,3-diol groups,[23] and those that, following a biomimetic approach, mimics the recognition

function of proteins exclusively relying the recognition on non-covalent interactions.[24,25] Recent advances in boronic-acid-based receptors have been extensively reviewed over the last few years[26–29] and include the interesting results obtained with molecular imprinted polymers (MIPs) not only in the recognition of biologically relevant glycan,[30] such as the blood antigens H, B and A,[31] but also of both the glycan and peptide structure of glycoproteins such as in the case of the highly mannosylated glycans of ovalbumin (OVA) and horseradish peroxidase (HRP).[32]

Another class of artificial CBAs that has shown to be promising for carbohydrate recognition is that of the aptamers.[33] Aptamers are functional single-stranded oligonucleotides capable of binding specific small molecules or proteins which are usually prepared through an in vitro evolution process called systematic evolution of ligands by exponential enrichment (SELEX).[34] In the field of carbohydrate recognition, aptamers have only recently been introduced. Although the first results were obtained with aptamers functionalised with boronic acid binding groups, more recently aptamers following a biomimetic approach have proven to be able to effectively recognise glycans. Interesting results have been obtained in the discrimination of different protein glycoforms[35] as well as in the binding of the glycosylated form of human prostate specific antigen (hPSA)[36] or in the targeting of the highly mannosylated glycoprotein gp120 of HIV.[37]

The most studied class of biomimetic CBAs is that of synthetic receptors for carbohydrates. Since the first example developed by Aoyama in 1988,[38] this heterogeneous group of small molecules, tailored-made for carbohydrate recognition, has grown extensively to achieve effective binding properties and selectivity in the recognition of biologically relevant saccharides.[39,40] The successful use of biomimetic CBAs in a biological context has further raised interest in the topic, which is still a hot field of current research.[25,41] However, development of effective biomimetic receptors that have to compete with a polar solvent such as water is indeed a non-trivial task. For this reason, a common approach in the design of new structures consists in beginning the investigation of binding properties in a low competitive medium such as an organic solvent, where polar interactions are enhanced, with the aim to move forward towards aqueous media once effective recognition properties have been achieved.[42]

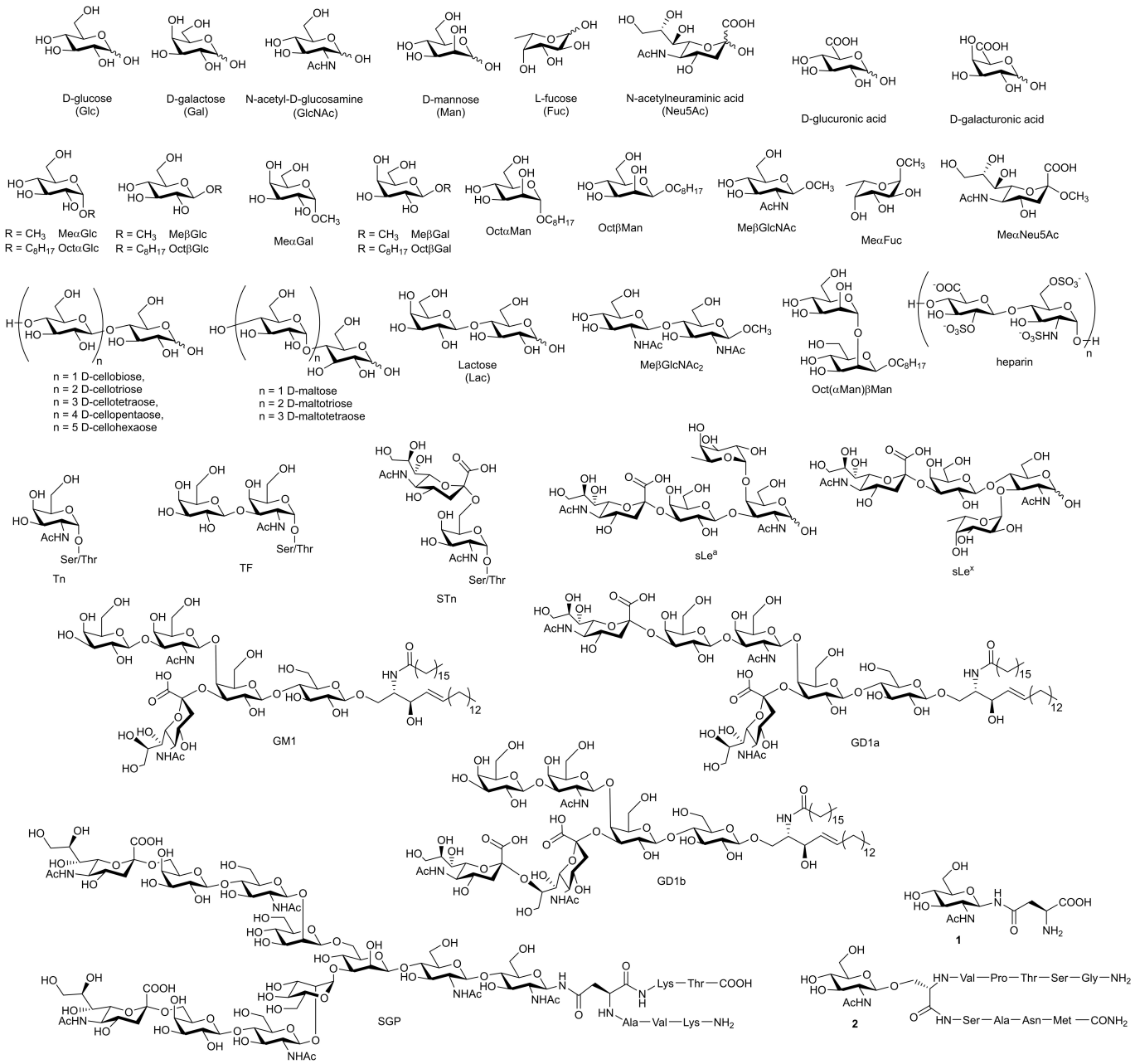
To address the binding requirements for an effective recognition and to find the optimal balance between rigid building blocks that pre-organize the binding groups for interaction and flexible structures that enhance adaptability to the guest, several architectures have been adopted including covalent macrocycles, coordination-driven macrocycles, acyclic structures, cages and foldamers.

In the last years, much effort has been devoted to the development of synthetic receptors for glucose, often driven by an interest in developing new potential glucose sensors to monitor its levels in the bloodstream.[43–45] However, with the exception of glucans (a class of glycans that are polymers of D-glucose), it is most common to find monosaccharides other than D-glucose as constituents of glycans – e.g. D-galactose (Gal), L-fucose (Fuc), D-mannose (Man), N-acetyl-D-glucosamine (GlcNAc), N-acetyl-D-galactosamine (GalNAc) and N-acetylneuraminic acids (Neu5Ac) (see Scheme 1). Many step forwards have been made in the recognition of different type of monosaccharides by several research groups and their results have been collected in valuable reviews.[15,25,39,46,47] However, the saccharides constituting glycans are part of complex structures, and there are relatively few examples in the literature of synthetic receptors for di- and oligosaccharides, and even fewer for glycoconjugates. Indeed, to recognize more complex ligands,

Francesco Milanesi received his PhD in Structural Biology in 2022 at the University of Firenze. During his PhD he worked on the development of both synthetic receptors for carbohydrates and glycomimetics and he spent a period as visiting student at CIC bioGUNE (Derio, Spain). Currently he is a postdoctoral researcher at the Department of Chemistry of the University of Firenze. His research interests are in supramolecular chemistry and carbohydrates chemistry.

Stefano Roelens graduated in chemistry at the University of Roma in 1976. After five years at Montedison Chemical Company as a research scientist, in 1982 he joined the National Research Council of Italy (CNR), holding a permanent research position at the University of Firenze, where he was Project Leader from 1995. He has been research fellow at the University of Montréal, and appointed Professor at the Universities of Messina and Firenze. He has been Senior Scientist of the CNR (1999–2015), leading research projects in Supramolecular Chemistry at the Chemistry department of the University of Firenze, where he is currently research associate of the INSTM. His scientific interests are focused on molecular recognition and self-assembly by hydrogen bonding, design and synthesis of artificial receptors, host-guest interactions, molecular recognition of carbohydrates.

Oscar Francesconi received his PhD in 2010 from University of Firenze where he worked on synthetic receptor for carbohydrates. After a period as visiting researcher at Centro de Investigaciones Biológicas in Madrid, he worked as a postdoctoral researcher and then as Assistant Professor at the Department of Chemistry of the University of Firenze, where he is currently an Associate Professor in organic chemistry. His interests are in supramolecular chemistry and physical-organic chemistry, with particular focus on the biomimetic recognition of carbohydrates. He received a Junior Research Award 2016 from the Italian Chemical Society and the Research Award 2018 from CINMPIS.



Scheme 1. Saccharides, alkyl glycosides and glycoconjugates with corresponding abbreviations.

receptors have to fulfil structural and functional requirements that are often different from those for a single monosaccharide.

The focus of this review is to summarize and compare the work that has been done in the field of biomimetic recognition of carbohydrates targeting glycans and their constituting saccharides. The review will be organized in sections treating different saccharides including mannosides and highly mannosylated proteins in Section 2, sialic acid and related glycans in Section 3, galactosides and the structurally related fucosides in Section 4, GlcNAc and related O- and N-glycans in Section 5, and glycan polysaccharides in Section 6.

2. D-Mannose and mannosylated glycans

D-Mannose is a monosaccharide found both as part of the core structure of N-glycans ($Man_3GlcNAc_2$) and as the terminal residue of high-mannose type N-glycans ($Man_{5\text{-}9}GlcNAc_2$, see Figure 1), which are commonly found on the surface of several pathogens.[48,49] Indeed, high-mannose type N-glycans exposed.

On enveloped viruses such as HIV,[50,51] Zika virus,[52] Ebola virus[50] and SARS-CoV-2,[53] have recently been shown to be central to the infection mechanism and have emerged as key targets for novel therapeutic approaches. For instance, oligomannose N-glycans on glycoprotein gp120 of HIV facilitate the viral entry by binding to the receptor domain of the host cell

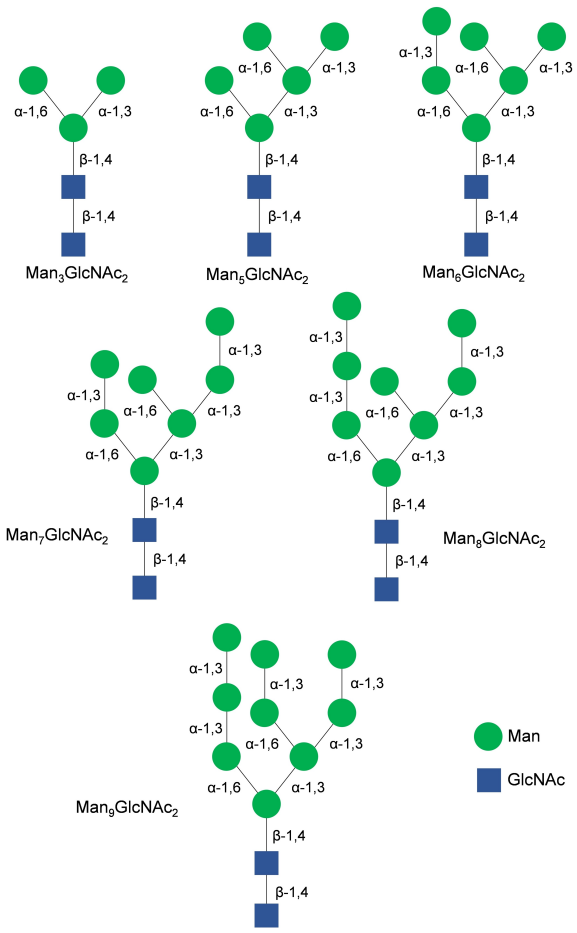
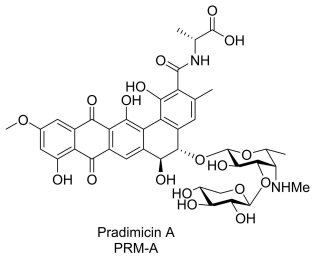


Figure 1. Principal high-mannose type glycans structures ($Man_{5\text{-}9}GlcNAc_2$). Man is represented as green dots, GlcNAc as blue squares.

protein DC-SIGN, while oligomannosides on the spike protein of SARS-CoV-2 are known to stabilize the "up" state of the protein,[54] which is required for binding to the host cell angiotensin converting enzyme 2 (ACE2). High-mannose type N-glycans are also overexpressed in aberrant glycans found on certain cancer cells including those of human liver,[55] lung,[56] breast,[57,58] and prostate,[59] so becoming an appealing target for cancer therapy and imaging.

Because of the importance of oligomannosides, researchers have focused on CBAs that target D-mannose. These CBAs include the only known non-peptidic CBA of natural origin, the antibiotic pradimicin A (PRM–A) (Scheme 2).[60] PRM–A has shown potent antimicrobial activity against fungi, such as *Candida albicans*, by altering membrane permeability upon mannoside recognition in a Ca^{2+} dependent manner.[61,62] Recently, a PRM–A derivate functionalised with a fluorescent probe was successfully used to stain the cell wall of *Candida rugosa*, demonstrating the versatility of this tool in visualizing cell surface glycans.[62b] Furthermore, PRM–A has also been shown to prevent HIV entry by recognizing mannosides exposed on the gp120 glycoprotein of the viral envelope, showing an $EC_{50} = 3.33$ μM against HIV-1.[46]

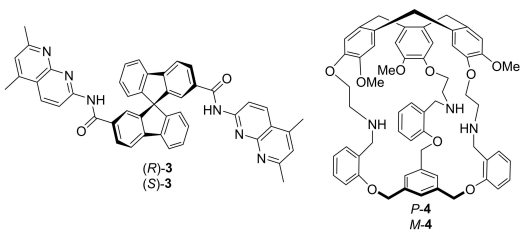


Scheme 2. Chemical structure of pradimicin-A.

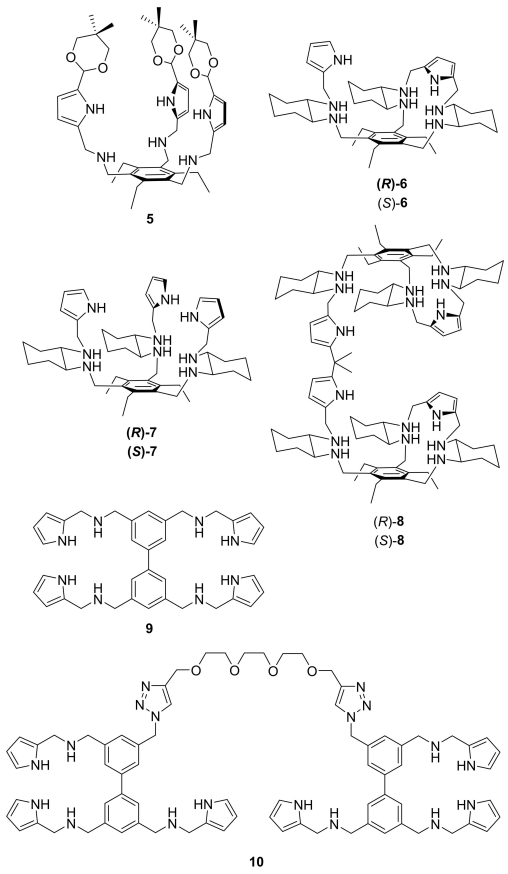
In the field of supramolecular chemistry, many efforts have been made in the development of biomimetic synthetic receptors for mannosides. In a pioneering stage Diederich and coworkers have reported a family of spirobifluorene cleft structures showing selective mannose recognition (see **3** in Scheme 3).[63] These chiral receptors, developed to study the role of chirality in the recognition of chiral species like carbohydrates, showed good affinity and appreciable enantioselectivity towards OctαMan in chloroform, as (*S*)-**3** binds with a $K_a = 1.27\times10^3$ M^{-1} while (*R*)-**3** shows a $K_a = 8.7\times10^2$ M^{-1}.

Instead, using receptor chirality to tune selectivity between different carbohydrates was a strategy pursued by Martinez and collaborators with a pair of chiral hemicryptophane cages, the *P* (right-handed helix) and the *M* (left-handed helix), that showed opposite preferences between D-glucose and D-mannose.[64] Indeed, the *M*-**4** receptor (Scheme 3) was able to recognize in chloroform OctβGlc with a $K_a = 9.93\times10^2$ M^{-1} and OctβMan with a $K_a = 7$ M^{-1} while the *P*-**4** enantiomer recognized OctβGlc with a $K_a = 4.5\times10^1$ M^{-1} and OctβMan with a $K_a = 8.56\times10^2$ M^{-1}. The pre-organization provided by the cage structure helps the receptor to exploit the chirality in the enantiodiscrimination of the guest reaching a significant difference in mannose recognition between the *P* and *M* receptors of two orders of magnitude.

Taking advantage of chirality to find the best match with the saccharide was also the approach taken by Roelens and coworkers with a family of chiral receptors based on 1,2-*trans*-diaminocyclohexane as chiral building blocks. These receptors belong to the larger family of aminopyrrolic receptors previously developed by the same group,[65,66] which have shown selectivity towards mannosides, as observed for the tripodal receptor **5** (Scheme 4),[67] which recognizes OctβMan with a $BC^0_{50} = 6.8\times10^{-4}$ M[68] in a more competitive solvent than chloroform, such as acetonitrile. The macrocyclic receptor **6** and the



Scheme 3. Chemical structure of receptors **3** and **4**.



Scheme 4. Chemical structure of receptors **5**–**10**.

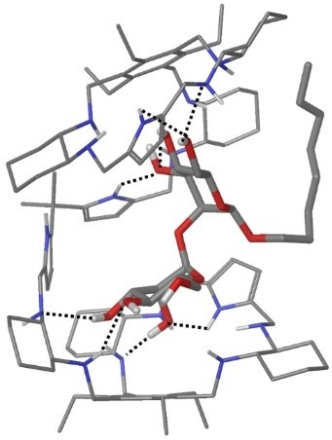


Figure 2. Global minimum structure obtained from a search of the conformational space for the complex between (*S*)-**8** and Oct(αMan)βMan. Hydrogen bonds are depicted as dashed lines. Adapted from ref. [70] with permission © 2013 John Wiley & Sons, Inc..

acyclic receptor **7** (Scheme 4) are the most important examples of these chiral receptors.[69] Indeed, receptor (*S*)-**6** showed the best affinities for OctβMan, with a $BC_{50}^{0} = 8.3\times10^{-5}$ M in acetonitrile, and a marked (1:15) enantioselectivity between the *R* and *S* receptors. Conversely, receptor (*R*)-**7** showed selectivity for OctαMan with a $BC_{50}^{0} = 1.27\times10^{-4}$ M, but with a milder enantioselectivity (1:7), most probably explained by the higher flexibility of its structure, which better adapts the opposite chirality to the same guest. Because of the importance of dimannosides as terminal unit of high-mannose type glycans, the same group has more recently developed a set of chiral ditopic receptors for the recognition of Manα(1–2)Man target.[70] All the set showed selectivity towards dimannosides over monosaccharides and in particular receptor **8** (Scheme 4) demonstrated the most interesting binding properties towards Oct(αMan)βMan in a competitive organic solvent mixture of DMF 30% in chloroform. Indeed (*S*)-**8** showed to enantioselectively recognize the dimannoside with a marked affinity with a $BC_{50}^{0} = 1.5\times10^{-5}$ M (see binding mode in Figure 2) whereas the enantiomer (*R*)-**8** showed an affinity two orders of magnitude lower with a $BC_{50}^{0} = 1.32\times10^{-3}$ M.

The family of aminopyrrolic receptors has been successfully used to target high-mannose type glycans in different biological contexts. Indeed, in analogy to PRM–A aminopyrrolic receptors showed antimycotic activity up to MIC = 2 μg mL^{-1} towards *Pichia norvegiensis* and *Prototheca wickerhamii* comparable to known antibiotics such as amphotericin B.[71] Aminopyrrolic receptors also demonstrated binding abilities towards the mannosylated gp120 glycoprotein of HIV as measured by surface plasmon resonance (SPR) analysis and showed antiviral activity inhibiting HIV-1 entry in T-lymphocyte cells with an EC_{50} up to 1.7 μM.[72] Antiviral activity of aminopyrrolic receptors was also evaluated against another enveloped viruses, the SARS-CoV-2. Aminopyrrolic receptors demonstrated a broad neutralizing activity with IC_{50} ranging from 7.9 to 16.8 μM towards several variants of the spike protein.[73] Antiviral activity appears to be due to the binding to the mannosylated spike protein, which in turn inhibits binding of the latter to ACE2 receptor, as suggested by saturation transfer difference (STD) experiments carried out with the recognition binding domain (RBD) of the wild-type spike protein. In analogy to certain mannose binding lectins and PRM–A, aminopyrrolic receptors demonstrates the capability to induce a caspase-dependent apoptosis with low micromolar IC_{50} in cancer cells expressing high level of mannose glycans.[74] Interestingly, cytotoxic activity was dependent on the expression levels of mannose glycans on the cell surfaces and on the ability of receptors to recognise mannosides. In particular, compounds (*R*)-**7** and (*S*)-**7** showed IC_{50} values of 1.3 and 2.7 μM, respectively, and their apoptotic activity was inhibited in a concentration dependent manner by addition of MeαMan. Moreover, the ability of mannose-binding concanavalin A (ConA) to bind cells overexpressing mannosides was attenuated in the presence of (*R*)-**7** and (*S*)-**7**. More recently, receptor (*R*)-**7** has been used to develop functionalized niosomes for mannose-targeted doxorubicin delivery.[75] In vitro studies towards triple-negative cancer cells (MDA-MB-231), overexpressing high-mannose type glycans, showed for the functionalized niosomes a cytotoxic activity comparable to free (*R*)-**7** demonstrating the correct presentation of the CBA in niosomes, and for functionalized niosomes loaded with doxorubicin a cytotoxic activity comparable to free doxorubicin but with an appreciable increment in apoptosis given by the CBA.

Finally, comparison studies with normal H9C2 cells showed a protective role of the formulation on cardiomyocytes.

The class of aminopyrrolic receptors was further expanded by Braunschweig and coworkers using a biphenyl scaffold.[76] They report in 2013[77] and later in 2018[78] the binding properties of the tetrapodal receptor **9** (Scheme 4) able to selectively recognize OctβMan with a $K_{a1} = 1.2\times10^{3}\ M^{-1}$ and $K_{a2} = 3.0\times10^{1}\ M^{-1}$ corresponding to 1:1 and 2:1 receptor–sugar association constants in dichloromethane. Later the same group published the antiviral activity against Zika virus of a library of receptors for mannosides.[79] Among the tested compounds, the ditopic receptor **10** (Scheme 4) resulted the most promising. The receptor is structured with two biphenyl cores, functionalized with aminopyrroles and linked through a tetraethylen glycol spacer. The authors evaluated the infection inhibition in different cell lines (HeLa and Vero) and compound **10** reported an $IC_{50} = 0.24\ \mu M$ in HeLa and $IC_{50} = 0.16\ \mu M$ in Vero, together with a negligible toxicity.

Inspired by the calcium-mediated mannose binding of PRM-A[62b] and the mannose recognition property of certain C-type lectins, that employ Ca^{2+} ions to coordinate the guest (see Figure 3a), synthetic receptors that use metal ions for mannose recognition were developed and shown to be effective. One interesting study is that of Striegler and coworkers who have presented in 2003 a binuclear copper(II) complex **11** (Figure 3) able to recognize D-mannose in water with a $pK_{app} = 4.06$.[80] Unfortunately, strong interactions are only observable at alkaline pH (pH = 12.4), quite far from physiological conditions.

More recently in 2018 Huc and coworkers proposed an aromatic oligoamide metallofoldamer (**12**, Figure 3) for mannose recognition.[81] The oligoamide is adaptable while in solution and it forms the receptor cavity upon addition of Cu^{2+} ions. The helical receptor binds D-mannose over other monosaccharides with $K_a = 4.50\times10^{2}\ M^{-1}$ in a 20% mixture of DMSO in chloroform. Crystal structure of the complex revealed an unusual second-coordination sphere interaction between Cu^{2+} hydrates and the carbohydrate that result in the moderate affinity observed.

A foldamer architecture for mannose recognition was also adopted by Abe and Inouye by mean of a pyridine-acetylene-phenol hexamer **13** (Figure 4) that fold upon binding to the saccharide ligand.[82] Although the nanomolar affinities observed in dichloromethane, low selectivity was observed among octyl α- and β-glycosides. However, a certain degree of selectivity towards D-mannose was observed in solid-liquid extraction experiment in dichloromethane where D-mannose was predominately extracted over other monosaccharides. A nonameric analogue **14** of the previous receptor having alkenyl side chains was used to form helical structure upon binding with D-mannose extracted from solid followed by metathesis reaction to staple the templated helix.[83] D-mannose extraction experiment from solid in dichloromethane showed again a selectivity towards this sugar.

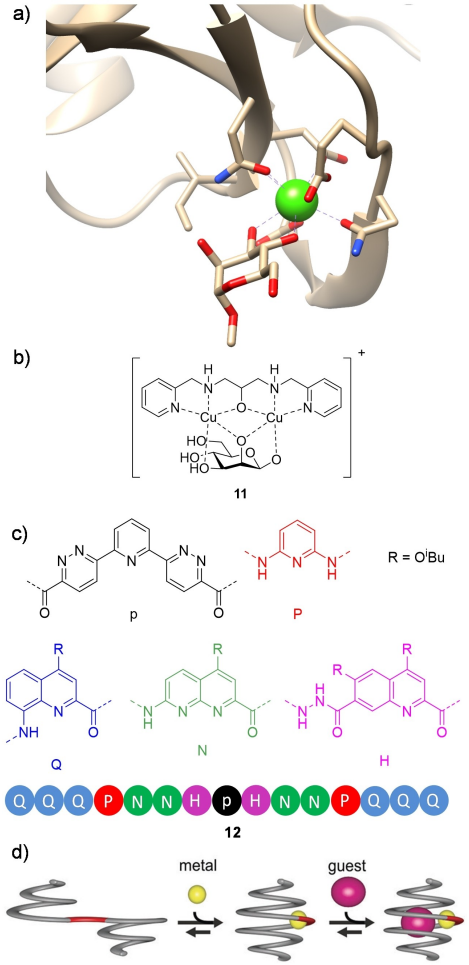


Figure 3. a) Detail of the binding site of the CRD4 of mannose receptor (CD206) of macrophages in complex with MeαMan (from PDB entry: 7JUB); b) structure of receptor **11**; c) schematic representation of receptor **12** sequence; d) representation of folding upon metal coordination followed by metal-assisted encapsulation of the ligand. Adapted from ref. [81] with permission from The Royal Society of Chemistry.

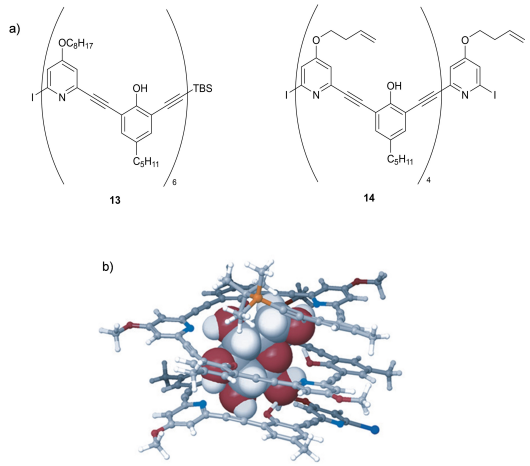


Figure 4. a) Structures of receptors **13** and **14**; b) model of the helical complex of **13** (side alkyl chains were substituted with methyl groups for simplicity) with β-D-glucose obtained by DFT calculation. Adapted from ref. [82] with permission. © 2015 John Wiley & Sons, Inc.

3. Sialic acid and sialylated glycans

Sialic acids are a family of 9-termed carbohydrates of which N-acetylneuraminic acid (Neu5Ac), often simply referred to as sialic acid, is the most representative. Sialic acid is a common terminal residue in mammalian glycans, and it is found in many TACAs. The recognition of Neu5Ac is of note one of the most crucial processes in the immune response.[84] Indeed, many protein families expressed on the surface of immune cells have the capability to recognize sialic acid-containing self-antigens. Upon recognition of sialylated glycans different proteins can up- or downregulate the immune response.[85] This mechanism is used to overcome autoimmune responses, but many pathogens mimic sialic acid-containing glycans to evade immune cells.[86] Sialic acid recognition is usually driven by a salt bridge between the carboxylic group and an arginine in the binding site of the protein. Taking advantage of the electrostatic interaction established with cationic hosts is a strategy that has been widely pursued in the development of functional materials for the separation, sensing and analysis of sialylated glycans.[87]

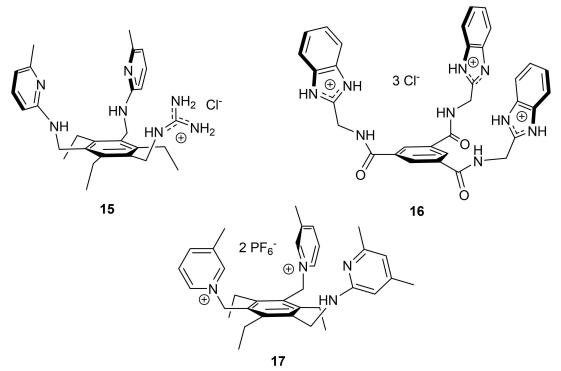
Despite the large number of processes in which sialic acid is involved, there are few examples of artificial CBAs targeting this sugar in the literature, and most of them feature the use of boronic acids.[88] However, some biomimetic synthetic receptors can also be found, such as the tripodal receptors **15** and **16** of Mazik and coworkers (Scheme 5).[89] The authors used a biomimetic approach by means of charged groups to establish a strong interaction with carboxylic group of the sialic acid. Receptor **15** is endowed with a guanidine residue and two pyridine moieties while receptor **16** is decorated with three methylbenzimidazoles, linked to the benzene scaffold through amide bonds. Both the receptors are capable to strongly recognize the tetramethylammonium salt of sialic acid in a highly polar medium such as the mixture water/DMSO 1:9. The receptors showed a different stoichiometry in the complexation of the sugar, as highlighted by experimental data, and confirmed by modelling. Indeed, receptor **15** can accommodate 1 to 2 sugar molecules in its binding site, with the $K_{a1} = 1.5\times10^5\ M^{-1}$ and a $K_{a2} = 3.2\times10^4\ M^{-1}$. Instead, receptor **16** can accommodate 1 to 3 sugar molecules in its binding site, with $K_{a1} = 2.0\times10^5\ M^{-1}$, $K_{a2} = 6.1\times10^4\ M^{-1}$ and $K_{a3} = 1.2\times10^4\ M^{-1}$. Later the same group has developed a set of tripodal receptors bearing a combination of cationic and neutral binding arms that turned out to be effective in the recognition of Neu5Ac also in aqueous media.[90] Indeed, one of the most effective receptors of the set, **17**, accommodating up to two guests in the binding site showed an affinity with $K_{a1} = 7.68\times10^3\ M^{-1}$ and $K_{a2} = 2.6\times10^3\ M^{-1}$ in a 1/5 mixture of D_2O/DMSO (Scheme 5).

In 2016, Davis and collaborators published a pyrene-based receptor **18** (Scheme 6), bearing guanidines both as hydro solubilizing and binding groups for Neu5Ac recognition.[91] The tetrapodal receptor presents two identical binding sites and can effectively recognize MeαNeu5Ac in water at physiological pH with a 1:2 (R:G) binding stoichiometry. The constants of the equilibria were reported to be $K_{a1} = 1.31\times10^3\ M^{-1}$, $K_{a2} = 5.70\times10^2\ M^{-1}$ and $K_{a3} = 30\ M^{-1}$. The third constant is weaker, but its presence is justified by the structure of the receptor that can potentially recognize till four units of Neu5Ac.

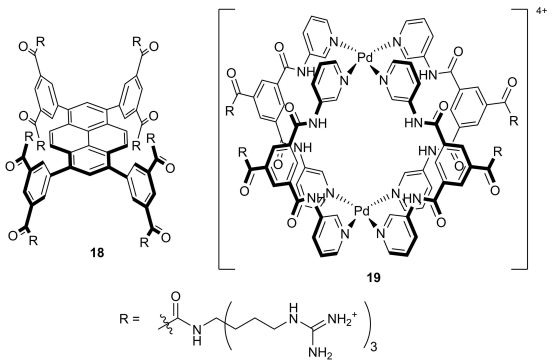
Recently Mooibroek and collaborators prepared a new synthetic CBA for sialic acid freely soluble in water and assembled by metal coordination. The metal-assembling strategy is indeed convenient to obtain polycyclic structures that generally requires long synthesis concerning macrocyclizations with low yields. Receptor **19** (Scheme 6) is a Pd_2L_4 cage, where L is a dipyridyl ligand.[92] The guanidinium terminal functions of R groups are responsible of the water solubility and the receptor shown to recognize Neu5Ac in water with an affinity of $K_a = 24.0\ M^{-1}$ and good selectivity among other anionic monosaccharides such as glucuronic acid and galacturonic acid.

Following a dynamic combinatorial approach Ravoo and coworkers have assembled a set of cyclic hexapeptides as carbohydrate receptors by means of a dynamic combinatorial library (DCL) of tripeptides using the reversible disulfide exchange reaction.[93] Composition of DCL in the presence of Neu5Ac changed yielding an amplification of two times the cyclic homodimer HisHis **20** (Scheme 7) that turned out to be able to bind Neu5Ac in water with a 1:2 (R:G) binding stoichiometry with $K_{a1} = 143\ M^{-1}$, $K_{a2} = 5.08\times10^3\ M^{-1}$ as determined by ITC experiments.[94]

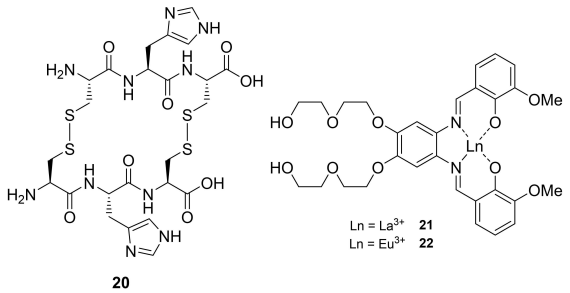
Inspired by the central role of coordination of saccharides to metal ions in C-type lectins, Strongin and coworkers have taken advantage of the similar properties of trivalent lanthanides and Ca^{2+} to develop water-soluble salophene-lanthanide complexes



Scheme 5. Chemical structure of receptors 15–17.



Scheme 6. Chemical structure of receptors **18** and **19**.



Scheme 7. Chemical structure of receptors **20**, **21** and **22**.

as fluorescent sensors for the detection of carbohydrates.[95] If using the complex with La^{3+} (**21**) yielded effective recognition at physiological pH towards D-glucose, maltose and maltotriose (500, 1666 and 2500 M^{-1} respectively), when Eu^{3+} is used instead the receptor (**22**) turned out to be able to selectively detect sialic acid-containing gangliosides, such as the monosialo GM1 and the disialo GD1a and GD1b, over asialo analogues (Scheme 7). The recognition involves multiple coordination sites belonging both to sialic acid and proximal oligosaccharides and presents higher selectivity with respect to europium(III)-tetracycline complexes.

4. Galactosides and Fucosides

D-Galactose is the epimer in position 4 of D-glucose. The first dietary resource of D-galactose is the disaccharide lactose found in milk and in dairy products. D-Galactose is one of the main constituents of glyconjugates both as simple D-galactose or N-acetyl-D-galactosamine and can also be found in many tumoral markers such as Tn, TF and STn. D-Galactose could be present both in the α configuration and the β, however, the α configuration is the most common as terminal epitope.[96]

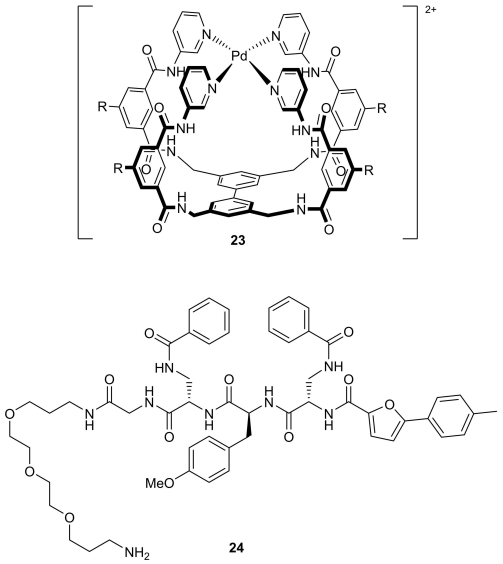
L-Fucose is structurally related to D-galactose being its 6-deoxy L-analogue. As this latter it is an essential constituent of glyconjugates, in which is usually been found as α anomer as terminal saccharide. L-Fucose can be linked to $GlcNAc_2$ disaccharide in the core structure of several N-glycans where its presence is crucial to the glycan conformation and to the orientation of the different antennae.[97] L-Fucose is also a constituent of the glycans in the ABO blood group and the Lewis antigen systems.

Since the design of a selective receptor for axial-substituted carbohydrates such as D-galactose and L-fucose proved to be more challenging than for all-equatorial carbohydrates, several different approaches have been taken to achieve an effective recognition. Inspired by previously reported polyamidic receptors for all-equatorial carbohydrates, Mooibroek and coworkers have proposed a hybrid design based on a covalent scaffold part of a coordination-driven macrocyclic structure.[98] Receptor **23** (Scheme 8) has a temple-like architecture in which the floor is represented by a biphenyl moiety covalently linked to four isophtalamide pillars and the roof is constituted by pyridine units chelating a Pd^{2+} ion. The receptor exclusively recognizes β-octyl galactopyranoside over other α and β glycosides in a competitive organic mixture of 10% DMSO in CD_2Cl_2 with a $K_a = 5.5\times10^2$ M^{-1}. Decreasing the polarity of the solvent with a mixture of 5% DMSO in CD_2Cl_2 increases the affinity for the galactoside by one order of magnitude while the affinity for mannosides and glucosides becomes measurable but remains two orders of magnitude lower than that for the β- galactoside.

Many efforts have been taken on the recognition of galactose containing oligosaccharides. Of note is the example of the peptide receptor **24** of Hall *et al.* (Scheme 8).[99] Despite the study is mostly focused on strongly effective bis(boroxoles) derivatives for the recognition of galactose oligosaccharides, receptor **24** stands out for the biomimetic approach followed. Indeed, using competitive ELISA experiments the peptide-like receptor showed an $IC_{50} = 100$ μM for the TF antigen in aqueous media, an affinity only four times weaker that the corresponding bis(borozole) derivative, highlighting the role of non-covalent interactions in the recognition process.

Achieving complex macrocyclic structures by self-assembled coordination to metal ions was a strategy followed by He and collaborators that have presented in 2012 a metal-organic octanuclear triangular prism **25** recognizing saccharides.[100] The constructive and chelating unit of the complex is H_4TRBS showed in Figure 5b). H_4TRBS contains several amide moieties that constitutes the tridentate chelating unit for metal coordination. The metal-organic polyhedral complex **25** was obtained by treatment of the chelating unit with $Ce(NO_3)_3\cdot6H_2O$. The formation of the described structure was confirmed by X-ray crystal analysis. Binding properties of the Ce-TRBS were analysed by fluorescence titrations revealing a selective recognition for lactose over other mono- and disaccharides with a $logK_a = 3.06$ in DMF/acetonitrile 1:9. Because of the optical properties of the complex, this receptor works as a selective chemosensor for lactose.

Despite the success in molecular recognition of galactosides in organic solvents, the examples of receptors effective towards



Scheme 8. Chemical structure of receptors **23** and **24**.

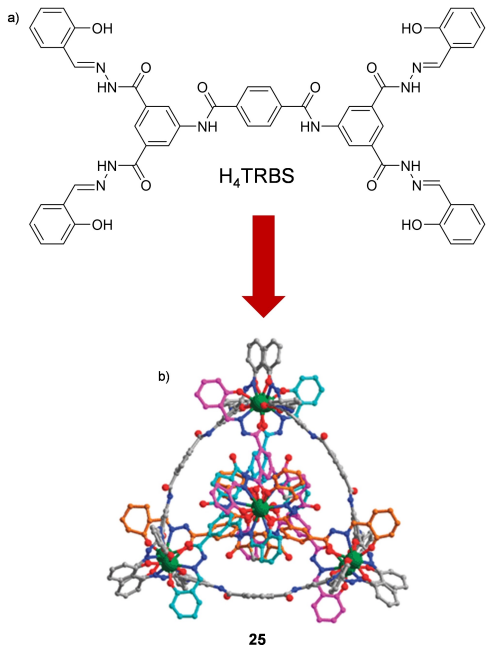


Figure 5. a) Structure of H_4TRBS; b) molecular structure of the octa-nuclear bicoronal triangular prism **25**. Adapted from ref. [100] with permission from The Royal Society of Chemistry.

this axially substituted monosaccharide in water are rare in the literature. An example comes from Francesconi and coworkers, which reported in 2018 the macrocyclic receptor **26** (Figure 6a),[101] constituted by two anthracenes for CH–π interactions with the aliphatic backbone of the sugar and two diaminocarbazole as tridentate hydrogen bonding units.[102] The receptor turned out to be able to accommodate all-equatorial saccharides, such as MeβGlc ($BC_{50}^0 = 1.3\times10^{-3}$ M), but remarkably, the molecule showed an interesting affinity for MeαGal with a $BC_{50}^0 = 1.2\times10^{-3}$ M, resulting in one of the most effective receptors for galactosides in water. However, the most relevant

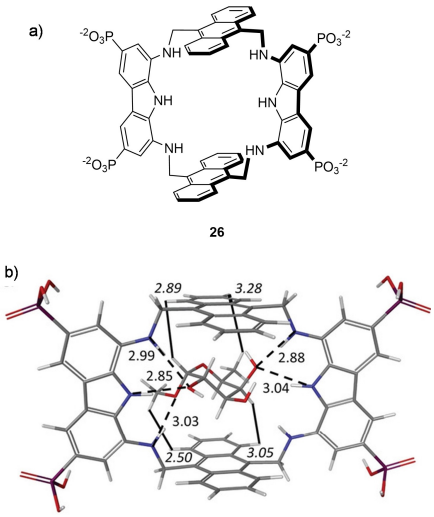


Figure 6. a) Structure of receptor **26**; b) complex between **26** and MeαFuc. Adapted from ref. [101] with permission. Copyright: Wiley-VCH, 2018.

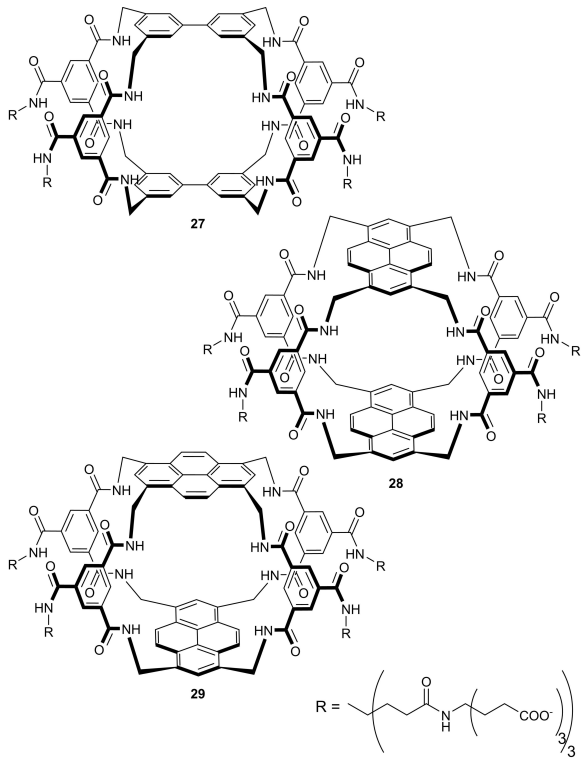
result of **26** was the affinity reported for MeαFuc for which this receptor showed selectivity. Indeed, the α-fucoside was recognized with a micromolar affinity ($BC_{50}^0 = 3.6\times10^{-4}$ M) that is closely comparable with those reported for several fucose/galactose-binding natural lectins, and an impressive α/β selectivity of 30-folds. Selectivity towards 1,4-diaxial substituted monosaccharides, such as the α anomers of D-galactose and L-fucose, appears to be governed by hydrogen bonding interactions with the diaminocarbazole binding units as suggested by thermodynamic parameters showing a strong enthalpic contribution and NMR-based minimum energy structure of the complex with MeαFuc showing hydrogen bonding interactions involving both the axial oxygens (Figure 6b).

5. N-Acetyl-D-glucosamine in O- and N-glycans

N-Acetyl-D-glucosamine is one of the most abundant sugars in Nature. Indeed, this carbohydrate represent the repeating unit of chitin, the polysaccharide that constitute the exoskeletons of arthropods and the cell wall of fungi. In mammals, GlcNAc does not have a structural role, but it is instead one of the most common saccharides of N- and O-glycans. For instance, O-GlcNAc is a dynamic post-translational glycosylation of proteins at level of serine or threonine residues that have role in critical biological processes and in the aetiology of diabetes and neurodegeneration.[103,104] GlcNAc can also be found both in the core region of N-glycans as $GlcNAc_2$ disaccharide directly linked a residue of Asn part of the protein backbone, as well as in the antennas of hybrid and complex N-glycans but generally not as a terminal residue. Although GlcNAc of the core is generally not recognized by proteins in mammals, the targeting of GlcNAc with artificial CBAs could be an effective strategy for protein glycan analysis and separation, and to study glycans as recognition sites.

In the recognition of an all-equatorial carbohydrate such as GlcNAc, temple structures turn out to be effective. Indeed, in 2009 Davis and coworkers reported a water-soluble polyamidic receptor **27** (Scheme 9) based of biphenyl roof and floor.[105] Receptor **27** selectively recognizes GlcNAc ($K_a = 5.6\times10^1$ M^{-1}) over other mono- and disaccharides especially as its methyl β glycoside (MeβGlcNAc, $K_a = 6.3\times10^2$ M^{-1}). Interestingly, no binding was detected for the disaccharide $GlcNAc_2$ and a weak binding was observed for the glycoside **1** N-linked to a residue of Asn ($K_a = 4$ M^{-1}). Contrarily, the glycoside **2** O-linked to a decapeptide based on a sequence from casein kinase II, known to be subject to O-GlcNAcylation, was strongly recognized with a $K_a = 1.04\times10^3$ M^{-1}. NMR based minimum energy structure revealed the origin of selectivity towards GlcNAc by means of hydrophobic contacts made by the methyl of the NHAc group with the aromatic rings of the receptor.

A step forward in the recognition of MeβGlcNAc and of the glycopeptide **2** was made in 2016 by the same group with a polyamidic temple structure based on pyrenyl floor and roof.[106] While in receptor **28** the pyrenyl units are "eclipsed" in receptor **29** are "stagged" (Scheme 9). The receptors are both hydrosoluble thanks to a polycarboxylic dendritic residue on each of



Scheme 9. Chemical structure of receptors **27**–**29**.

the four pillars. Compounds **28** and **29** showed high affinities for MeβGlcNAc with association constants of $K_a = 2.1\times10^3\ M^{-1}$ and $K_a = 1.82\times10^4\ M^{-1}$ respectively. The stronger binding observed with **29**, that exceeds that observed with the GlcNAc-binding lectin wheat germ agglutinin, was ascribed to a more extended pattern of interactions with the sugar as highlighted by NMR and modelling studies. Contrary to the monosaccharide, binding studies towards **2** revealed a strong binding only with receptor **28** ($K_a = 6.7\times10^4\ M^{-1}$), while in contrast binding with receptor **29** was no quantifiable. The boost in affinity moving from MeβGlcNAc to **2** was explained by additional interactions occurring between the peptide aglycon and the receptor including some with the dendritic side chains as envisaged by NMR based minimum energy structure (Figure 7).

Designing effective synthetic receptors for a relatively large guests such as an oligosaccharide can be challenging since an increment in the complexity of the guest is generally addressed with a higher complexity of the receptor. An opposite approach was instead followed by Francesconi and coworkers that in 2021 have reported the simple acyclic receptor **30** that turned out to be selective and particularly effective in the recognition of $GlcNAc_2$ in water (Figure 8).[107] The tweezer-shaped architecture showed no binding towards a broad panel of monosaccharides including MeβGlcNAc, while it presents a micromolar affinity with a $BC^0_{50} = 1.6\times10^{-4}$ M towards the disaccharide $Me\beta GlcNAc_2$. This guest was selectively recognized over other structurally related 1,4-linked disaccharides including cellobiose, maltose and lactose for which affinities one order of magnitude lower were measured. The marked affinity for $Me\beta GlcNAc_2$, which exceeds of one order of magnitude also that measured for the pseudo-lectin hevein from *Hevea brasiliensis*, was explained by the extended network of hydrogen bonds, CH-π interactions and hydrophobic contacts established with the disaccharide, while the selectivity with respect to cellobiose is ascribable to additional interactions with one of the two NHAc groups interacting with the diaminocarbazole unit of the receptor.[108]

More recently the same tweezer-shaped architecture proved to be also effective in targeting of the $GlcNAc_2$ disaccharide within the core structure of a complex N-glycan.[109] Indeed, the tetraphosphonate analogue **31** (Figure 9), prepared to circumvent the self-association phenomenon of the progenitor **30**, demonstrated the capability to intercept and strongly recognize the $GlcNAc_2$ disaccharide at the stem of sialoglycopeptide SGP with the same affinity observed for the disaccharide ($BC^0_{50} = 1.7\times10^{-4}$ M). Despite the undecasaccharidic structure and the pentapeptide chain, the recognition occurs exclusively with the $GlcNAc_2$ disaccharide as confirmed by NMR studies. Moreover, the binding seems to not alter the presentation of the glycan terminals as suggested by NOESY and modelling studies in which it was possible to establish the existence in solution of the two most common conformations of the ligand, the "extend" and the "backfolded", differing in the torsional angles around the Manα1–6Man linkage.[110]

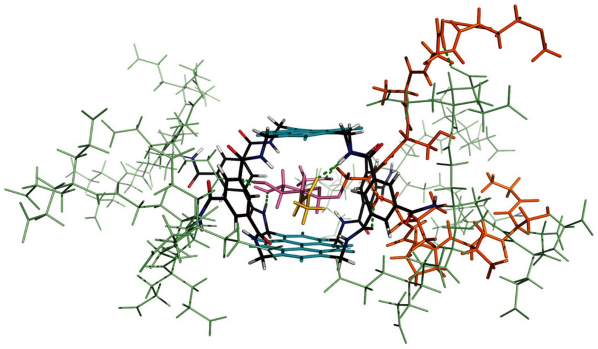


Figure 7. Complex between O-glycan **2** and receptor **29**. Adapted from ref. [106].

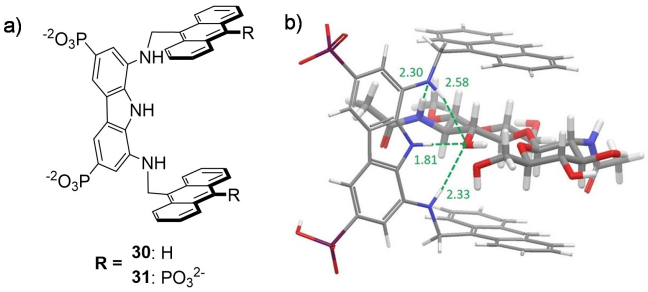


Figure 8. a) Structure of receptors **30** and **31**; b) complex between **30** and $Me\beta GlcNAc_2$. Adapted from ref. [107].

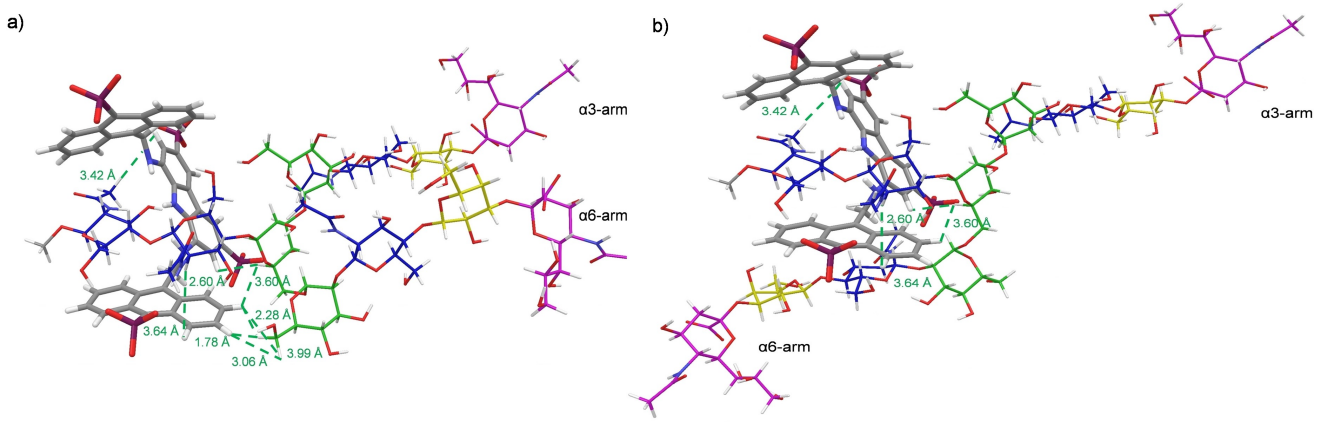
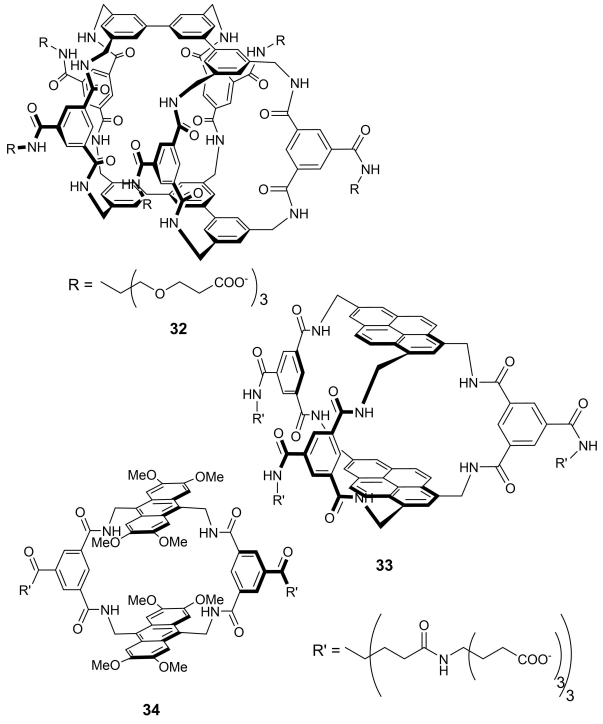


Figure 9. a) Complex between **31** and SGP with the glycopeptide in the "extended" conformation; b) Complex between **31** and SGP with the glycopeptide in the "backfolded" conformation. Adapted from ref. [109].

6. Polysaccharides

Polysaccharides are the most representative macromolecules in nature. Their roles are various, from the structural function of cellulose and hyaluronic acid to energy storage of starch and glycogen up to the anticoagulant function of heparin. The recognition of polysaccharides with small molecules is, of note, a difficult task due to the dimensions and the complexity of these structures. However, the recognition of repeating units of polysaccharides could be an effective strategy to alter the physical and chemical properties of these polymers. Cellulose is composed by repeating unit of cellobiose, an all-equatorial disaccharide the recognition of which has been largely explored with synthetic receptors. The first important result towards this sugar was obtained by Davis and coworkers in 2007. The authors described receptor **32** (Scheme 10), a temple structured receptor with two *meta*-terphenyl platforms as roof and floor that recognizes cellobiose in water with an affinity of $K_a = 6.0\times10^2\ M^{-1}$.[111]

More recently, a simplification of the temple structure led the same group to publish the bicyclic receptor **33** (Scheme 10), composed of three polyamidic hydrogen-bonding units and two pyrenes.[112] The affinity for cellobiose in water was significantly increased with respect to **32** ($K_a = 3.90\times10^3\ M^{-1}$), but more interestingly, the receptor showed an increase in affinity as the oligomer increases in length, resulting in a $K_a = 5.20\times10^3\ M^{-1}$ for cellotriose and $K_a = 1.20\times10^4\ M^{-1}$ for cellotetraose, while longer oligomers were recognized with an average affinity of $K_a = 8.80\times10^3\ M^{-1}$. Indeed, receptor **33** turned out to be able to form threaded complexes not only with these oligomers but also with cellulose as assessed by circular dichroism and atomic force microscopy experiments. If recognition of all-equatorial oligosaccharides can be challenging, designing effective receptors for non-all-equatorial oligosacchairdes may be perhaps even more so. Simplifying the design of "temple" receptors Davis and coworkers have more recently developed a monocyclic receptor **34** (Scheme 10) endowed with only two isophthalamides and two anthracenes bearing four methoxy groups.[113] Receptor **34** showed selectivity



Scheme 10. Chemical structure of receptors **32**–**34**.

towards maltose and maltodextrins over other mono- and 1,4-linked disaccharides including cellodextrins presenting an affinity of $K_a = 5.80\times10^2\ M^{-1}$ for maltose, $K_a = 1.15\times10^3\ M^{-1}$ for maltotriose and $K_a = 1.62\times10^3\ M^{-1}$ for maltotetraose. The four MeO groups play an essential role in tuning the selectivity towards the α-linked over β-linked, providing additional interaction with the glycosyl units external to the cavity thanks to the curved profile of maltotriose guest, effect that does not apply to cellotriose as can be envisaged by Figure 10.

Heparin is a natural charge dense glycosaminoglycan (GAG), commonly used as anticoagulant drug. Given its medical relevance, many efforts have been taken in the development of sensors for this polysaccharide. In nature, heparin is recognized through key interactions between anionic groups and positively

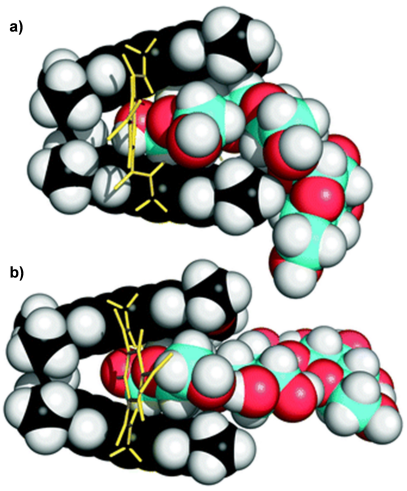
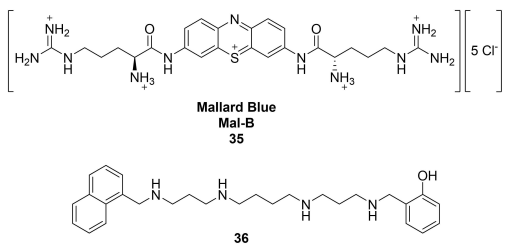


Figure 10. a) Complex between receptor **34** and maltotriose; b) complex between receptor **34** and cellotriose. Adapted from ref. [113] with permission. from The Royal Society of Chemistry.

charged amino acids in the binding sites.[114] To achieve an effective recognition following a biomimetic approach, in 2013, Smith and coworkers prepared a new dye, called by the authors Mallard Blue (MalB **35**), able to selectively recognize heparin in water (Scheme 11).[115] Binding studies performed by UV/Vis spectroscopy showed a reduction of MalB absorbance upon addition of heparin that is not significantly reverted by addition of equimolar amounts of similarly charged cationic species, suggesting that although binding is driven by electrostatic interactions the structure of MalB is determinant for selective heparin recognition. MalB turn out to be also effective in human and horse serum, in condition in which a traditional heparin sensor such as Azure A did not work.[116]

More recently, Alfonso and collaborators proposed another polycationic receptor for heparin (**36**, Scheme 11).[117] The receptor is structured with two aromatic moieties for CH-π interactions linked together through a spermine linker bearing the ammonium groups. Calorimetric measurements showed an affinity for heparin of $K_d = 1.30\times10^{-6}$ M in aqueous media with an average stoichiometry of 7.5 molecules of **36** for each molecule of heparin. A three-dimensional description of the complex (Figure 11), in which the flexible structure of **36** folds to interact with the anionic oligosaccharide, was obtained by molecular dynamic calculations.



Scheme 11. Chemical structure of receptors **35** and **36**.

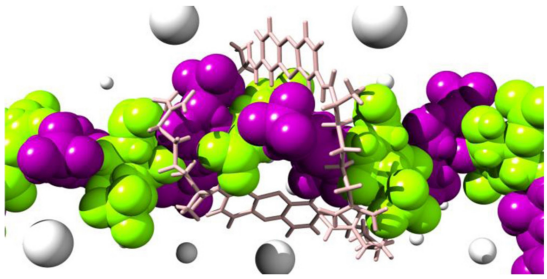


Figure 11. Complex of Mal-B with heparin polymer. Adapted with permission from: ref. [115]. © 2013, American Chemical Society.

7. Summary and Outlook

In the growing field of molecular recognition of carbohydrates by synthetic receptors, the targeting of glycans and their constituent elements is a major challenge. Recognition of these complex ligands requires a tailor-made receptor design, which has led several supramolecular chemists to explore different architectures, including temple structures, cages, podands, tweezer-like structures and foldamers. By relying solely on non-covalent interactions for biomimetic recognition, significant progress has been made in recent years and effective recognition has been achieved even in a very competitive medium such as water, not only towards monosaccharides, but also towards whole glycans both as oligo- and polysaccharides or glycoconjugates with proteins and lipids. Progress in glycan recognition has been summarized in this review, focusing on the recognition of important saccharides, such as D-mannose, D-galactose, L-fucose, N-acetyl-D-galactosamine and sialic acid and the glycans containing them. Recent successful results in this area will encourage the development of new receptors selectively recognizing a wider range of glycans and open the way to further applications ranging from glycosylation-based protein separation and detection to the study and modulation of protein-glycan recognition events for biofunctionality.

Conflict of Interests

The authors declare no conflict of interest.

Keywords: biomimetic receptors · carbohydrate-binding agents · glycans · host–guest systems · molecular recognition

[1] A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, P. H. Seeberger, R. L. Schnaar, Eds., *Essentials of Glycobiology, 4th Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), **2022**.
[2] T. K. Lindhorst, Eds., *Essentials of Carbohydrate Chemistry and Biochemistry*, John Wiley & Sons, **2007**.
[3] K. K. Palaniappan, C. R. Bertozzi, *Chem. Rev.* **2016**, *116*, 14277–14306.
[4] H.-J. Gabius, *BioSystems* **2018**, *164*, 102–111.
[5] X.-L. Zhang, H. Qu, *Adv. Exp. Med. Biol.* **2021**, *1325*, 219–237.
[6] C. J. Day, E. N. Tran, E. A. Semchenko, G. Tram, L. E. Hartley-Tassell, P. S. K. Ng, R. M. King, R. Ulanovsky, S. McAtamney, M. A. Apicella, J. Tiralongo, R. Morona, V. Korolik, M. P. Jennings, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, DOI 10.1073/pnas.1421082112.

[7] E. Rodríguez, S. T. T. Schetters, Y. van Kooyk, *Nat. Rev. Immunol.* **2018**, *18*, 204–211.
[8] S. S. Pinho, I. Alves, J. Gaifem, G. A. Rabinovich, *Cell. Mol. Immunol.* **2023**, *20*, 1101–1113.
[9] R. J. Woods, *Chem. Rev.* **2018**, *118*, 8005–8024.
[10] S. Mereiter, M. Balmaña, D. Campos, J. Gomes, C. A. Reis, *Cancer Cell* **2019**, *36*, 6–16.
[11] J. Tena, C. B. Lebrilla, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, DOI 10.1073/PNAS.2022238118.
[12] J. Y. Zhou, D. M. Oswald, K. D. Oliva, L. S. C. Kreisman, B. A. Cobb, *Trends Immunol.* **2018**, *39*, 523–535.
[13] B. Radovani, I. Gudelj, *Front. Immunol.* **2022**, *13*, DOI 10.3389/fimmu.2022.893365.
[14] S. S. Pinho, C. A. Reis, *Nat. Rev. Cancer* **2015**, *15*, 540–555.
[15] A. P. Davis, *Org. Biomol. Chem.* **2009**, *7*, 3629–3638.
[16] J. Z. J. I. A. P. J. H. Angulo, in *Essentials of Glycobiology [Internet]. 4th Edition.* (Eds.: A. Varki, R. Cummings, J. Esko, et al.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), **2022**.
[17] S. H. Liyanage, M. Yan, *Chem. Commun.* **2020**, *56*, 13491–13505.
[18] B. E. Collins, J. C. Paulson, *Curr. Opin. Chem. Biol.* **2004**, *8*, 617–625.
[19] S. Tommasone, F. Allabush, Y. K. Tagger, J. Norman, M. Köpf, J. H. R. Tucker, P. M. Mendes, *Chem. Soc. Rev.* **2019**, *48*, 5488–5505.
[20] D. Chettri, M. Boro, L. Sarkar, A. K. Verma, *Carbohydr. Res.* **2021**, *506*, DOI 10.1016/j.carres.2021.108367.
[21] O. Haji-Ghassemi, R. J. Blackler, N. Martin Young, S. V. Evans, *Glycobiology* **2015**, *25*, 920–952.
[22] E. Sterner, N. Flanagan, J. C. Gildersleeve, *ACS Chem. Biol.* **2016**, *11*, 1773–1783.
[23] G. Springsteen, B. Wang, *Tetrahedron* **2002**, *58*, 5291–5300.
[24] A. P. Davis, R. S. Wareham, *Angew. Chem. Int. Ed.* **1999**, *38*, 2978–2996.
[25] O. Francesconi, S. Roelens, *ChemBioChem* **2019**, *20*, 1329–1346.
[26] X. Wu, Z. Li, X. X. Chen, J. S. Fossey, T. D. James, Y. B. Jiang, *Chem. Soc. Rev.* **2013**, *42*, 8032–8048.
[27] W. Zhai, X. Sun, T. D. James, J. S. Fossey, *Chem. Asian J.* **2015**, *10*, 1836–1848.
[28] S. Jin, Y. Cheng, S. Reid, M. Li, B. Wang, *Med. Res. Rev.* **2010**, *30*, 171–257.
[29] A. Adamczyk-Woźniak, K. M. Borys, A. Sporzyński, *Chem. Rev.* **2015**, *115*, 5224–5247.
[30] Y. Zhao, *J. Mater. Chem. B* **2022**, *10*, 6607–6617.
[31] R. W. Gunasekara, Y. Zhao, *J. Am. Chem. Soc.* **2017**, *139*, 829–835.
[32] L. Duan, M. Zangiabadi, Y. Zhao, *Chem. Commun.* **2020**, *56*, 10199–10202.
[33] W. Sun, L. Du, M. Li, *Curr. Pharm. Des.* **2010**, *16*, 2269–2278.
[34] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.
[35] A. M. Yoshikawa, A. Rangel, T. Feagin, E. M. Chun, L. Wan, A. Li, L. Moeckl, D. Wu, M. Eisenstein, S. Pitteri, H. T. Soh, *Nat. Commun.* **2021**, *12*, 7106.
[36] A. Díaz-Fernández, R. Miranda-Castro, N. Díaz, D. Suárez, N. de-los-Santos-Álvarez, M. J. Lobo-Castañón, *Chem. Sci.* **2020**, *11*, 9402–9413.
[37] W. Li, S. Xu, Y. Li, J. Chen, Y. Ma, Z. Liu, *CCS Chemistry* **2023**, *5*, 497–509.
[38] Y. Aoyama, Y. Tanaka, H. Toi, H. Ogoshi, *J. Am. Chem. Soc.* **1988**, *110*, 634–635.
[39] A. P. Davis, *Chem. Soc. Rev.* **2020**, *49*, 2531–2545.
[40] Z. Sun, B. Fan, M. J. Webber, *ChemSystemsChem* **2023**, *5*, DOI 10.1002/syst.202200050.
[41] C. E. Miron, A. Petitjean, *ChemBioChem* **2015**, *16*, 365–379.
[42] M. Gentili, C. Nativi, O. Francesconi, S. Roelens, *Carbohydr. Chem.* **2016**, *41*, 149–186.
[43] Q. Wu, L. Wang, H. Yu, J. Wang, Z. Chen, *Chem. Rev.* **2011**, *111*, 7855–7875.
[44] X. Sun, T. D. James, *Chem. Rev.* **2015**, *115*, 8001–8037.
[45] 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.
[46] Y. Nakagawa, I. Yukishige, *Adv. Carbohydr. Chem. Biochem.* **2012**, *68*, 1–58.
[47] M. Mazik, *Chem. Soc. Rev.* **2009**, *38*, 935–956.
[48] K. O. François, J. Balzarini, *Med. Res. Rev.* **2012**, *32*, DOI 10.1002/med.20216.
[49] J. Balzarini, *Nat. Rev. Microbiol.* **2007**, *5*, 583–597.
[50] G. Lin, G. Simmons, S. Pöhlmann, F. Baribaud, H. Ni, G. J. Leslie, B. S. Haggarty, P. Bates, D. Weissman, J. A. Hoxie, R. W. Doms, *J. Virol.* **2003**, *77*, 1337–1346.
[51] S.-K. Wang, P.-H. Liang, R. D. Astronomo, T.-L. Hsu, S.-L. Hsieh, D. R. Burton, C.-H. Wong, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3690–3695.
[52] N. K. Routhu, S. D. Lehoux, E. A. Rouse, M. R. M. Bidokhti, L. B. Giron, A. Anzurez, S. P. Reid, M. Abdel-Mohsen, R. D. Cummings, S. N. Byrareddy, *Int. J. Mol. Sci.* **2019**, *20*, 5206.
[53] Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan, M. Crispin, *Science* **2020**, *369*, 330–333.
[54] H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li, S. Li, *Cell* **2020**, *183*, 730–738.
[55] N. C. Phillips, D. Robinson, B. G. Winchester, *Biochem. J.* **1976**, *153*, DOI 10.1042/bj1530579.
[56] E. Blattes, A. Vercellone, H. Eutamène, C.-O. Turrin, V. Théodorou, J.-P. Majoral, A.-M. Caminade, J. Prandi, J. Nigou, G. Puzo, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8795–8800.
[57] X. Liu, H. Nie, Y. Zhang, Y. Yao, A. Maitikabili, Y. Qu, S. Shi, C. Chen, Y. Li, *PLoS One* **2013**, *8*, e72704.
[58] M. L. A. de Leoz, L. J. T. Young, H. J. An, S. R. Kronewitter, J. Kim, S. Miyamoto, A. D. Borowsky, H. K. Chew, C. B. Lebrilla, *Mol. Cell. Proteomics* **2011**, *10*, M110.002717.
[59] M. Ferens-Sieczkowska, B. Kowalska, E. M. Kratz, *Biomarkers* **2013**, *18*, 10–22.
[60] T. J. Walsh, N. Giri, *Eur. J. Clin. Microbiol. Infect. Dis.* **1997**, *16*, 93–97.
[61] Y. Nakagawa, Y. Masuda, K. Yamada, T. Doi, K. Takegoshi, Y. Igarashi, Y. Ito, *Angew. Chem. Int. Ed.* **2011**, *50*, 6084–6088.
[62] a) Y. Nakagawa, T. Doi, Y. Masuda, K. Takegoshi, Y. Igarashi, Y. Ito, *J. Am. Chem. Soc.* **2011**, *133*, 17485–17493; b) Y. Nakagawa, T. Doi, K. Takegoshi, T. Sugahara, D. Akase, M. Aida, K. Tsuzuki, Y. Watanabe, T. Tomura, M. Ojika, Y. Igarashi, D. Hashizume, Y. Ito, *Cell Chem. Biol.* **2019**, *26*, DOI 10.1016/j.chembiol.2019.03.013.
[63] J. Cuntze, L. Owens, V. Alcázar, P. Seiler, F. Diederich, *Helv. Chim. Acta* **1995**, *78*, 367–390.
[64] A. Long, O. Perraud, M. Albalat, V. Robert, J. P. Dutasta, A. Martinez, *J. Org. Chem.* **2018**, *83*, 6301–6306.
[65] A. Ardá, C. Venturi, C. Nativi, O. Francesconi, F. J. Cañada, J. Jiménez-Barbero, S. Roelens, *Eur. J. Org. Chem.* **2010**, *2010*, 64–71.
[66] C. Nativi, M. Cacciarini, O. Francesconi, A. Vacca, G. Moneti, A. Ienco, S. Roelens, *J. Am. Chem. Soc.* **2007**, *129*, 4377–4385.
[67] C. Nativi, M. Cacciarini, O. Francesconi, G. Moneti, S. Roelens, *Org. Lett.* **2007**, *9*, 4685–4688.
[68] A. Vacca, O. Francesconi, S. Roelens, *Chem. Rec.* **2012**, *12*, 544–566; The BC^{0}_{50} (intrinsic median binding concentration) parameter, which can be calculated from the measured association constants, is defined as the total concentration of receptor necessary to bind 50% of the ligand when the fraction of bound receptor is zero and directly provides a quantitative measure of binding ability because, analogously to the IC_{50} parameter, the lower the value, the higher the affinity. It coincides with the dissociation constant K_d when the 1:1 association is the only equilibrium present in solution.
[69] C. Nativi, O. Francesconi, G. Gabrielli, A. Vacca, S. Roelens, *Chem. Eur. J.* **2011**, *17*, 4814–4820.
[70] O. Francesconi, C. Nativi, G. Gabrielli, M. Gentili, M. Palchetti, B. Bonora, S. Roelens, *Chem. Eur. J.* **2013**, *19*, 11742–11752.
[71] 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.
[72] O. Francesconi, C. Nativi, G. Gabrielli, I. De Simone, S. Noppen, J. Balzarini, S. Liekens, S. Roelens, *Chem. Eur. J.* **2015**, *21*, 10089–10093.
[73] 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.
[74] 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.
[75] N. Burrini, M. D'Ambrosio, M. Gentili, R. Giaquinto, V. Settimelli, C. Luceri, M. Cirri, O. Francesconi, *Pharmaceutica* **2023**, *15*, 235.
[76] M. F. Bravo, M. A. Lema, M. Marianski, A. B. Braunschweig, *Biochemistry* **2021**, *60*, 999–1018.
[77] S. Rieth, M. R. Miner, C. M. Chang, B. Hurlocker, A. B. Braunschweig, *Chem. Sci.* **2013**, *4*, 357–367.
[78] K. Palanichamy, M. F. Bravo, M. A. Shlain, F. Schiro, Y. Naeem, M. Marianski, A. B. Braunschweig, *Chem. Eur. J.* **2018**, *24*, 13971–13982.
[79] K. Palanichamy, A. Joshi, T. Mehmetoglu-Gurbuz, M. F. Bravo, M. A. Shlain, F. Schiro, Y. Naeem, H. Garg, A. B. Braunschweig, *J. Med. Chem.* **2019**, *62*, 4110–4119.
[80] S. Striegler, M. Dittel, *J. Am. Chem. Soc.* **2003**, *125*, 11518–11524; The apparent binding constants ($pK_{app} = -\log K_{app}$) is the constant

measured in the presence of other species (salt, buffer components) that may interfere with determination of the pure pH-dependent conditional dissociation constant.
[81] P. Mateus, B. Wicher, Y. Ferrand, I. Huc, *Chem. Commun.* **2018**, *54*, 5078–5081.
[82] Y. Ohishi, H. Abe, M. Inouye, *Chem. Eur. J.* **2015**, *21*, 16504–16511.
[83] H. Abe, C. Sato, Y. Ohishi, M. Inouye, *Eur. J. Org. Chem.* **2018**, *2018*, 3131–3138.
[84] S. S. Pinho, I. Alves, J. Gaifem, G. A. Rabinovich, *Cell. Mol. Immunol.* **2023**, *20*, 1101–1113.
[85] S. Duan, J. C. Paulson, *Annu. Rev. Immunol.* **2020**, *38*, 365–395.
[86] M. P. Lenza, U. Atxabal, I. Oyenarte, J. Jiménez-Barbero, J. Ereño-Orbea, *Cells* **2020**, *9*, 2691.
[87] a) M. Li, Y. Xiong, G. Qing, *Acc. Chem. Res.* **2023**, *56*, DOI 10.1021/acs.accounts.3c00397; b) X. Dong, H. Qin, J. Mao, D. Yu, X. Li, A. Shen, J. Yan, L. Yu, Z. Guo, M. Ye, H. Zou, X. Liang, *Anal. Chem.* **2017**, *89*, DOI 10.1021/acs.analchem.6b04394.
[88] A. Di Pasquale, S. Tommasone, L. Xu, J. Ma, P. M. Mendes, *J. Org. Chem.* **2020**, *85*, 8330–8338.
[89] M. Mazik, H. Cavga, *J. Org. Chem.* **2007**, *72*, 831–838.
[90] C. Geffert, M. Kuschel, M. Mazik, *J. Org. Chem.* **2013**, *78*, 292–300.
[91] 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.
[92] X. Schaapkens, R. N. van Sluis, E. O. Bobylev, J. N. H. Reek, T. J. Mooibroek, *Chem. Eur. J.* **2021**, *27*, 13719–13724.
[93] S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Science (1979)* **2002**, *297*, 590–593.
[94] M. Rauschenberg, S. Bandaru, M. P. Waller, B. J. Ravoo, *Chem. Eur. J.* **2014**, *20*, 2770–2782.
[95] O. Alptürk, O. Rusin, S. O. Fakayode, W. Wang, J. O. Escobedo, I. M. Warner, W. E. Crowe, V. Král, J. M. Pruet, R. M. Strongin, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9756–9760.
[96] F. Conte, N. van Buuringen, N. C. Voermans, D. J. Lefeber, *Biochim. Biophys. Acta Gen. Subj.* **2021**, *1865*, 129898.
[97] M. Schneider, E. Al-Shareffi, R. S. Haltiwanger, *Glycobiology* **2017**, *27*, 601–618.
[98] B. J. J. Timmer, A. Kooijman, X. Schaapkens, T. J. Mooibroek, *Angew. Chem. Int. Ed.* **2021**, *60*, 16178–16183.
[99] A. Pal, M. Bérubé, D. G. Hall, *Angew. Chem. Int. Ed.* **2010**, *49*, 1492–1495.
[100] Y. Jiao, J. Zhang, L. Zhang, Z. Lin, C. He, C. Duan, *Chem. Commun.* **2012**, *48*, 6022.
[101] O. Francesconi, M. Martinucci, L. Badii, C. Nativi, S. Roelens, *Chem. Eur. J.* **2018**, *24*, 6828–6836.
[102] O. Francesconi, F. Cicero, C. Nativi, S. Roelens, *ChemPhysChem* **2020**, *21*, 257–262.
[103] J. Ma, G. W. Hart, *Expert Rev. Proteom.* **2013**, *10*, 365–380.
[104] B. E. Lee, P.-G. Suh, J.-I. Kim, *Exp. Mol. Med.* **2021**, *53*, 1674–1682.
[105] Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G. J. Boons, S. Ingale, A. P. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 1775–1779.
[106] P. Rios, T. S. Carter, T. J. Mooibroek, M. P. Crump, M. Lisbjerg, M. Pittelkow, N. T. Supekar, G. J. Boons, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 3387–3392.
[107] O. Francesconi, F. Milanesi, C. Nativi, S. Roelens, *Angew. Chem. Int. Ed.* **2021**, *60*, 11168–11172.
[108] O. Francesconi, F. Milanesi, C. Nativi, S. Roelens, *Chem. Eur. J.* **2021**, *27*, 10456–10460.
[109] F. Milanesi, L. Unione, A. Ardá, C. Nativi, J. Jiménez-Barbero, S. Roelens, O. Francesconi, *Chem. Eur. J.* **2023**, *29*, DOI 10.1002/chem.202203591.
[110] 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.
[111] Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619–622.
[112] 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.
[113] P. Stewart, C. M. Renney, T. J. Mooibroek, S. Ferheen, A. P. Davis, *Chem. Commun.* **2018**, *54*, 8649–8652.
[114] N. S. Gandhi, R. L. Mancera, *Chem. Biol. Drug Des.* **2008**, *72*, 455–482.
[115] S. M. Bromfield, A. Barnard, P. Posocco, M. Fermeglia, S. Pricl, D. K. Smith, *J. Am. Chem. Soc.* **2013**, *135*, 2911–2914.
[116] M. D. Klein, R. A. Drongowski, R. J. Linhardt, R. S. Langer, *Anal. Biochem.* **1982**, *124*, 59–64.
[117] M. Corredor, D. Carbajo, C. Domingo, Y. Pérez, J. Bujons, A. Messeguer, I. Alfonso, *Angew. Chem. Int. Ed.* **2018**, *57*, 11973–11977.

Manuscript received: October 19, 2023
Revised manuscript received: November 3, 2023
Accepted manuscript online: November 9, 2023
Version of record online: November 27, 2023