

Chemistry A European Journal

 **Chemistry
Europe**
European Chemical
Societies Publishing

Accepted Article

Title: Synthesis and Binding of Mannose-Specific Synthetic Carbohydrate Receptors

Authors: Marcelo Fernando Bravo, Kalanidhi Palanichamy, Milan Shlain, Frank Schiro, Yasir Naeem, Mateusz Marianski, and Adam B Braunschweig

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Chem. Eur. J.* 10.1002/chem.202000481

Link to VoR: <https://doi.org/10.1002/chem.202000481>

WILEY-VCH

Synthesis and Binding of Mannose-Specific Synthetic Carbohydrate Receptors

M. Fernando Bravo,^[a,b,c] Kalanidhi Palanichamy,^[a,b] Milan A. Shlain,^[a,b] Frank Schiro,^[a,b] Yasir Naeem,^[a,b] Mateusz Marianski,^{*[a,b,c,d]} and Adam B. Braunschweig^{*[a,b,c,d]}

^[a] M. F. Bravo, Dr. K. Palanichamy, M. A. Shlain, F. Schiro, Y. Naeem, Prof. Dr. A. B. Braunschweig

Advanced Science Research Center at the Graduate Center of the City University of New York,
85 St Nicholas Terrace, New York, NY 10031

E-mail: abraunschweig@gc.cuny.edu

^[b] M. F. Bravo, Dr. K. Palanichamy, M. A. Shlain, F. Schiro, Y. Naeem, Prof. Dr. M. Marianski,
Prof. Dr. A. B. Braunschweig

Department of Chemistry and Biochemistry, Hunter College, 695 Park Ave, New York, NY
10065

^[c] M. F. Bravo, Prof. Dr. M. Marianski, Prof. Dr. A. B. Braunschweig

The PhD Program in Chemistry, The Graduate Center of the City University of New York, 365
5th Ave, New York, NY 10016

^[d] Prof. Dr. M. Marianski, Prof. Dr. A. B. Braunschweig

The PhD Program in Biochemistry, The Graduate Center of the City University of New York, 365
5th Ave, New York, NY 10016

Abstract

Synthetic Carbohydrate receptors (SCRs) that selectively recognize cell-surface glycans could be used for detection, drug delivery, or as therapeutics. Here we report the synthesis of seven new C_{2h} symmetric tetrapodal SCRs. The structures of these SCRs possess a conserved biaryl core, and they vary in the four heterocyclic binding groups that are linked to the biaryl core via secondary amines. Supramolecular association between these SCRs and 5 biologically relevant C^1 -O-octyloxy glycans, α/β -glucoside (**α/β -Glc**), α/β -mannoside (**α/β -Man**), and β -galactoside (**β -Gal**), was studied by mass spectrometry, 1H NMR titrations, and molecular modeling. These studies revealed that selectivity can be achieved in these tetrapodal SCRs by varying the heterocyclic binding group. We found that **SCR017** (3-pyrrole), **SCR021** (3-pyridine), and **SCR022** (2-phenol) bind only to **β -Glc**. **SCR019** (3-indole) binds only to **β -Man**. **SCR020** (2-pyridine) binds **β -Man** and **α -Man** with a preference to the latter. **SCR018** (2-indole) binds **α -Man** and **β -Gal** with a preference to the former. The glycan guests bound within their SCR hosts in one of three supramolecular geometries: center-parallel, center perpendicular, and off-center. Many host•guest combinations formed higher stoichiometry complexes, 2:1 **glycan•SCR** or 1:2 **glycan•SCR**, where the former are driven by positive allosteric cooperativity induced by glycan-glycan contacts.

Introduction

Interactions between glycan-binding proteins and carbohydrates in the glycocalyx have a critical role in immune response, cell-cell communication, cell-pathogen interactions, disease progression, and many other complex biological processes,^[1] with each cell and virus type presenting unique glycosylation pattern. For example, α -mannose is overexpressed on the surface of colorectal^[2] and ovarian^[3] cancer cells and several infectious viruses, including the *Flaviviridae*,^[4] HIV,^[5] and ebola,^[6] whereas β -galactose and N-acetylglucosamine are overexpressed on the surface of malignant melanoma cells.^[7] Although natural lectins and antibodies recognize specific glycans,^[8] their toxicity^[9] and immunological triggering have limited their development as therapeutic or imaging agents.^[10] So despite having a role in multiple high-priority diseases, there are currently only two glycan-binding antibodies in clinical trials and only one that has received FDA approval, which is approved for treating high-risk childhood neuroblastoma.^[11] As a result, cell- and viral-surface glycans are still considered “undruggable targets”,^[12] meaning they have a known role in disease progression, but no widely adopted therapeutic strategies exploit this information.^[13]

Small molecules that bind the glycans common to the surfaces of pathogens or diseased cells could have a transformational impact on understanding and treating disease. Synthetic carbohydrate receptors^[14] (SCRs) – synthetic molecules that form supramolecular complexes with glycans – could be used as substitute for lectins and antibodies in the context of drug delivery agents, disease or biofilm detection platforms, or as therapeutics. A major advantage of SCRs compared to antibodies or lectins is that SCR structures can be easily manipulated through the versatility of organic synthesis to tune substrate selectivity or mitigate toxicity. A challenge that has prevented the wider adoption of SCRs by the medical and research communities is that the majority of SCRs preferentially bind glucosides – all equatorial glycans that are prevalent in the blood and cytoplasm but are not major components of glycocalyx oligosaccharides – and, as a result, have limited uses.^[7b] Thus there is a pressing need for developing SCRs that preferentially, or even selectively, bind non-glucosidic glycans, which is a major challenge for supramolecular chemists because it requires designing receptors that can distinguish between structurally complex guests that sometimes differ by only the orientation of only a single stereocenter.

Nevertheless, substantial research efforts have been devoted to the development of SCRs, and these molecules can generally be subdivided into two classes: in the first,^[15] binding occurs through the formation of a covalent boronate ester on the SCRs with *syn* diols on the glycans, and in the second,^[14, 16] binding is noncovalent and is the result of H-bonding, C–H $\cdots\pi$ interactions, and van der Waals forces. Of the latter class, SCRs with some non-glucosidic

selectivity have been successfully developed by Davis,^[16ap-as, 16av] Roelens,^[16aa-ad, 16ah, 16ak-ao, 17] and Mazik.^[16f, 16h-j, l, o, q, u, v, 18] However, only a fraction of those SCRs are known to bind mannosides.^[16aa-ad, 16ah, ak, am, an]

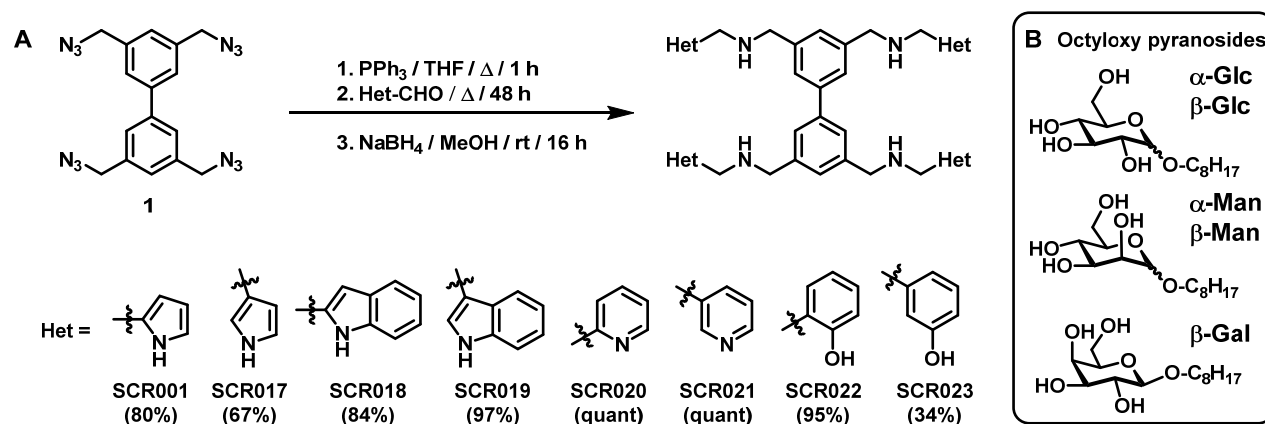
Like natural lectins, however, these noncovalent SCRs are promiscuous and bind many glycans, sometimes showing a preference for one glycan, and specific SCRS – those that only bind one non-glucosidic glycan – are still needed. Towards this goal, our group has developed a series of noncovalent SCRs based upon the biaryl-tetrapodal receptor **SCR001** that is selective towards mannosides^[19] – the C2 epimers of glucosides – as a result of multivalent and cooperative binding modes. Subsequently, we explored the role of the bonding and valency of a small library of tetrapodal SCRs on glycan association, and found that subtle changes in both have profound effects on their affinity and selectivity.^[19b] Further studies revealed that these tetrapodal SCRs inhibit Zika virus entry into Vero and HeLa cells,^[20] most likely by interrupting the clathrin-mediated endocytosis that follows glycan binding. Some of these tetrapodal receptors possessed nanomolar inhibition, making them amongst the most potent inhibitors yet reported against *Flaviviridae in vitro*, and, thus, supports continuing to explore how altering the structures of these SCRs affects glycan binding and antiviral activity.

The challenge of selectivity must be addressed before SCRs are widely adopted as sensors, therapeutics, or delivery agents. SCRs that bind non-glucosides, however, continue to resist rational design. To achieve selectivity the binding motifs found in their biological archetypes, lectins, where binding is characterized by shallow wells, promiscuity, multivalency, and cooperativity, must be reduced to a molecular design. Moreover, the binding pockets of SCR-glycan supramolecular structures exceed substantially the complexity of most host-guest complexes, and, as such, molecular modeling approaches are needed that can predict SCR selectivity and guide their design. Here we sought to address these challenges and create SCRs with increased selectivities towards non-glucosidic monosaccharides. An approach that has been adopted widely in modulating SCR affinity is to vary the heterocyclic units that form C–H \cdots π and H-bonding interactions with the glycan guests and their linkage regiochemistry.^[16j, l, ak, al, 18] Inspired by this previous work, here we adopt this approach of varying the heterocycle and the regiochemistry of attachment of the heterocycle to the biaryl core to study how this changes the binding of the resulting tetrapodal SCRs towards a series of biologically-relevant glycans. Of the seven new tetrapodal receptors we prepared, the result is two receptors that binds primarily α -mannosides, a receptor that is selective towards β -mannosides, and three receptors that are specific towards β -glucosides, all of which are biologically relevant targets, and these selectivities are explained by examining the molecular models. Importantly, the work described in these

studies lays out a viable pathway towards the development of selective SCRs, and, as a result, realizing their full potential in a variety of medical, research, and industrial applications.

Results and Discussion

The synthesis and binding of **SCR001** – **SCR009** and **SCR012** were reported previously.^[19] **SCR017–SCR023** (**Scheme 1A**) are new and were synthesized from common intermediate **1** in yields ranging from 34% to quantitative using our standard three-step one-pot protocol,^[19a] which involves a Staudinger amination of tetraazide **1** to give the corresponding iminophosphorane intermediate. A subsequent aza-Wittig reaction with the appropriate aryl/heteroaryl aldehyde is followed by reduction of the resulting imine with sodium borohydride. These new SCRs vary from **SCR001** in either the heterocycle structure or regiochemistry of heterocycle attachment, while maintaining the secondary amine groups and the C_{2h} symmetry of our previous tetrapodal SCRs. The heterocycle substituents explored here include 2- or 3-pyrrole, 2- or 3-indole, 2- or 3-pyridine, and 2- or 3-phenol. The SCR structures were characterized by ^1H NMR, ^{13}C NMR, and high-resolution mass spectrometry, and all spectroscopic data were consistent with the proposed structures (see Supporting Information for details).



Scheme 1. A) Synthesis of **SCR017** – **SCR023** from **1** and the corresponding heterocyclic aldehyde. **B)** Octyloxy pyranosides whose binding to the SCRs was studied.

Binding studies by mass spectrometry

Binding of the 7 new SCRs to the 5 glycans (**Scheme 1B**) was first studied by positive ion ESI mass spectrometry because the presence of ions corresponding to the **SCR•glycan** complex confirms supramolecular association.^[16aa, ae, ag, ak, 19b] As fragmentation peaks of the SCRs taken in the absence of glycan are necessary to interpret the mass spectra of the **SCR•glycan** complexes, we first subjected solutions containing only the SCRs to mass spectrometry analysis.

1 μM solutions of SCRs, were prepared by diluting 1 mM of the SCRs stock solutions in CH_2Cl_2 with 40% CH_2Cl_2 in CH_3CN . These solutions were then injected via direct infusion into the spectrometer with a syringe pump. Previously, in studying SCRs with pyrrole and furan substituents, the fragmentation patterns showed ions corresponding to the loss of heteroaryl arms via cleavage of the C–N bond,^[19b] which is a favored cleavage point for electron-rich heterocycles because of the stability of the resulting benzylic anions. Consistent with this, for example, the ESI mass spectrum of **SCR019** shows the $[\text{M}+\text{H}]^{1+}$ molecular ion in addition to $[\text{M}+\text{H}]^{n+}$ ions corresponding to loss of either one or more 3-indolebenzylic groups. Conversely, in the case of **SCR020**, which has electron poor pyridine heterocycles for instance, $[\text{M}+\text{H}]^{1+}$ and $[\text{M}+2\text{H}]^{2+}$ ions were observed, but ions corresponding to the cleavage of the C–N bond were not prominent, likely because the benzylic radical would not be stabilized in electron poor heterocycles. Similar trends were observed in all other SCR ionizations, showing that SCRs with electron rich heterocycles (**SCR001**, **SCR017** – **SCR019**) eject benzylic fragments, while SCRs with electron poor heterocyclic arms (**SCR020** – **SCR023**) do not (see Supporting Information).

With this understanding of how the SCRs fragment, the spectra of the **SCR•glycan** complexes were studied. 1 μM solutions of glycans in 40% CH_2Cl_2 in CH_3CN were prepared, and they were mixed one-to-one with the 1 μM solution of SCRs in 40% CH_2Cl_2 in CH_3CN . The mixture was then injected into the spectrometer via direct infusion with a syringe pump. Simulations of the expected masses and the isotopic distributions of the complexes, the individual components, and their fragmentation patterns were performed with Compass Data Analysis software (Bruker) to identify peaks corresponding to supramolecular association between the SCRs and the glycans. In the case of **SCR019• β -Man**, the ions corresponding to $[\text{SCR019}\cdot\beta\text{-Man}+\text{H}]^{1+}$, $[\text{SCR019}\cdot\beta\text{-Man}+2\text{H}]^{2+}$ and the ion $[\text{SCR019-Ind}\cdot\beta\text{-Man}+\text{H}]^{1+}$, resulting from loss of one indole-benzyl group, were observed (**Figure S22**). In the case of **SCR020• α -Man**, $[\text{SCR020}\cdot\alpha\text{-Man}+\text{H}]^{1+}$ and $[\text{SCR020}\cdot\alpha\text{-Man}+2\text{H}]^{2+}$ were observed (**Figure S22**). These studies were repeated for all **SCR•glycan** mixtures, and revealed that all SCRs bind all glycans assayed to some extent, forming 1:1 **SCR•glycan** complexes (**Figures S23-S29**). It should be noted, however, that these mass spectrometry experiments reveal little about strength and selectivity of association given the challenges related to the quantification of mass spectrometry binding data, and so other analytical techniques are required to determine association constants (K_{a} s) and selectivities of the SCRs towards the different glycans.

Determination of K_{a} s by NMR titrations

The K_a s between the glycans and the SCRs were determined by NMR titrations at 298 K in CD_2Cl_2 , since 1H NMR titrations are widely used for quantifying host-guest binding processes with K_a s ranging from 1 to $10^5 M^{-1}$.^[21] Also, synthetic carbohydrate receptor binding is commonly studied in non-aqueous solvents^[14a, b, 16c, 16f-ao, ar, 17-18, 22] because K_a s are generally higher than they would be in aqueous solvents, so changes in K_a s as a result of structural variations are amplified and more easily understood. Here CD_2Cl_2 was chosen as the solvent because it does not compete for H-bonds between the glycans and the receptors. For **SCR017**, **SCR019** and **SCR023**, 0.5% CD_3OD , 1% CH_3OH , and 4% CD_3OD , respectively, were added to the titration to increase the solubility of the SCRs. Prior to the **SCR•glycan** titration, dilution experiments were performed for all SCRs at a concentration range of 1 mM – 25 μM to determine if they undergo dimerization, and if the observed change in chemical shift ($\Delta\delta$) was >0.02 ppm,^[19b] the data were fit to a dimerization model^[19-20] to determine the dimerization constant, K_d (see Supporting Information). Dimerization was observed only for **SCR001**, which has been reported previously,^[19a] and **SCR023**. Following the dilution experiments, the 1H NMR titrations were performed by adding 6.25 μL aliquots of 16 mM solutions of glycans to 500 μL (1 mM) solutions of the SCRs, and the additions were continued to a 30:1 glycan:SCR ratio. As an illustrative example, the 1H NMR spectra of **SCR019• β -Man** is presented and discussed in detail here (**Figure 1**). For **β -Man**, the largest shift upon association was for the peak corresponding to the H^4 proton, with $\Delta\delta = 0.13$ ppm downfield, and the second largest shift was for the peak corresponding to H^6 , with $\Delta\delta = 0.11$ ppm downfield. The peak shifts are attributed to the change in chemical environment as a result of reversible supramolecular association between the glycan and the SCR that is occurring in the fast exchange regime. When involved in C–H $\cdots\pi$ interactions with aryl rings of SCRs, protons shift upfield,^[16f-ao, 18, 22a-e, 23] so these results suggest that these Hs do not form C–H $\cdots\pi$ interactions with **SCR019**. In contrast, the peaks corresponding to H^1 and H^5 of **β -Man** both shift upfield 0.04 ppm, upon association, suggesting the formation of C–H $\cdots\pi$ interactions with the aromatic rings of **SCR019**, which is corroborated by molecular modeling (**Figure S80**). Significant shifts were also observed for the SCR protons upon complexation. The largest shift was 0.20 ppm downfield for the peak corresponding to the indole N–H proton, indicating their participation in H–bonding with the glycans. The peak representing aromatic proton H^k shifted downfield 0.06 ppm, and the peak corresponding to H^f shifted 0.02 ppm upfield.

Titration experiments were repeated for all **SCR•glycan** combinations, and they are presented in the Supporting Information. Among all **SCR•glycan** complexes, the largest shift for a glycan C–H proton was $\Delta\delta = 0.15$ ppm downfield for the peak corresponding to H^6 in the **SCR017• α -Man** complex, and the largest shift for the aromatic proton was $\Delta\delta = 0.17$ ppm downfield for H^f of the

3-pyrrole ring of the **SCR017**•**α-Man** complex. Similarly, the largest shift for the N–H proton of pyrrole or indole heterocycle was $\Delta\delta = 0.74$ ppm downfield for N–H proton of 2-indole in the **SCR018**•**β-Gal** complex. ^1H NMR of most of the titrations presented significant shifts for SCR and glycan peaks upon mixing, indicating supramolecular association. It should be noted also that the spectra of several of the **SCR**•**glycan** combinations (e.g. **SCR018**•**β-Glc**, **SCR019**•**α-Glc**, **SCR021**•**α-Glc**, **SCR022**•**α-Glc**, **SCR021**•**β-Gal**, **SCR022**•**β-Gal**, **SCR023**•**α-Glc**, **SCR023**•**β-Man**, **SCR023**•**α-Man**, **SCR023**•**β-Gal**) did not show peak shifting upon mixing, indicating that no substantial binding was occurring under these experimental conditions. This result is significant because these data suggest that these SCRs do not bind all sugars – which is an important departure from previously studied tetrapodal SCRs, which were generally promiscuous binders, and associate with all glycans presented to them.

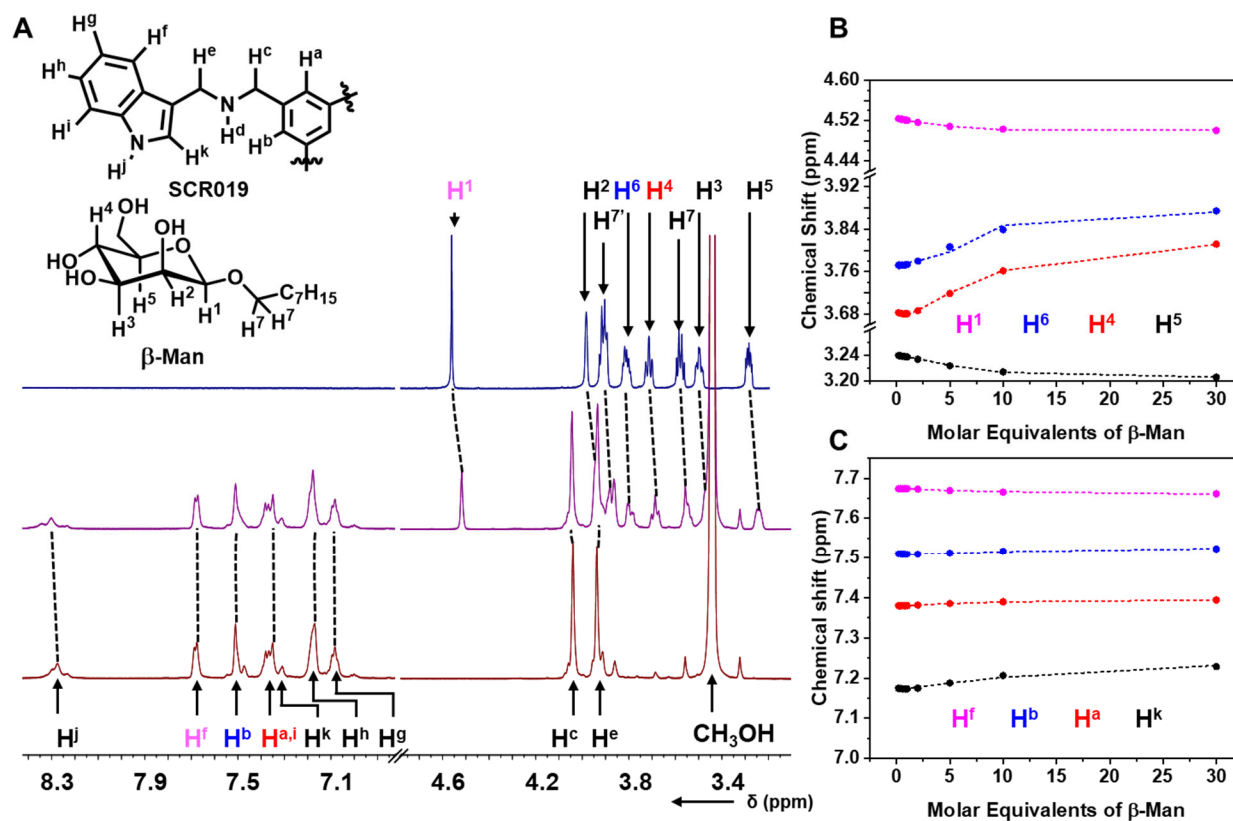


Figure 1. A) ^1H NMR (800 MHz, 1% CH_3OH in CD_2Cl_2 , 298 K) of $\beta\text{-Man}$ (16 mM, top), a 2:1 ratio of $\beta\text{-Man}$: **SCR019** (middle), and **SCR019** (1 mM, bottom). Dashed lines track the shifts of peaks upon mixing of **SCR019** and $\beta\text{-Man}$. B) The shift of the NMR peaks for protons H^{1} , H^{4} , H^{5} and H^{6} of $\beta\text{-Man}$ at 298 K, with bullets and lines representing the experimental data and the fit from a 1:2

SCR•glycan binding model, respectively. C) The shifts of the NMR peak of the H^a, H^b, H^f, and H^k protons of **SCR019** upon addition of **β-Man** in CD₂Cl₂ at 298 K, with bullets and lines representing the experimental data and the fit from a 1:2 **SCR•glycan** binding model, respectively.

These NMR titrations were used to quantify the K_a s for the supramolecular binding between the glycans and the SCRs. To determine the K_a s, ¹H NMR $\Delta\delta$ were fit to binding models that considered the different possible equilibria that can occur. For example, we have shown previously that the SCRs can dimerize, and that **SCR001** can form 1:1, 2:1, and 1:2 complexes with certain **β-Man** in CDCl₃^[19a] and CD₂Cl₂^[19b], and all these equilibria were considered when fitting the binding data. The K_a s and ΔG° for all **SCR•glycan** complexes and K_d for all SCRs were determined by minimizing the sum of squared residuals (SSRs) between the experimental data and the modelled fit (**Table 1**). To maximize the accuracy of the K_a values, NMR peak shift data of only clearly resolved peaks of glycans and SCRs that shifted a $\Delta\delta > 0.02$ ppm were fit simultaneously to an appropriate binding model, and the model that had the lowest error with the titration data was selected as the correct equilibrium. In an effort to avoid overestimation of K_a s, no binding values are reported for $K_a < 3.0 \times 10^1 \text{ M}^{-1}$, NMR peaks shift $< \Delta\delta = 0.02$ ppm, and unless 2 peaks in the ¹H NMR spectra have $\Delta\delta > 0.02$ ppm. To demonstrate the data analysis, the NMR peak shift data for the association of **SCR019** with **β-Man** are shown (**Figure 1B & C**). The peak shifts were best fit with a 1:2 **SCR•glycan** binding model with K_1 , K_2 and β of 2.3 M^{-1} , $3.2 \times 10^4 \text{ M}^{-1}$ and $7.4 \times 10^4 \text{ M}^{-2}$ where K_1 , K_2 and β correspond to 1:1 and 1:2 SCR:glycan K_a s and cumulative K_a ($K_1 \times K_2 \text{ M}^{-2}$), respectively. The higher stability of the 1:2 **SCR•mannoside** complex, **SCR019•(β-Man)₂**, compared to the 1:1 **SCR•mannoside** complex, **SCR019•(β-Man)** occurs as a result of positive allosteric cooperativity, as $K_2 > K_1$, ^[19a, 24] and the source of this cooperativity is explained with the molecular modeling results. Fitting of NMR peak shift data revealed that similar multiple equilibria also occur in the association of **SCR017** (3-pyrrole) with **β-Glc**, **SCR020** (2-pyridine) with **α-Man** and **SCR021** (3-pyridine) with **β-Glc**. On the other hand, the association of **SCR022** (2-phenol) with **β-Glc** showed formation of 1:1 and 2:1 **SCR•glycan** complexes in CD₂Cl₂. In all these cases, the K_2/K_1 ratio was >1 , indicating that the formation of higher stoichiometry complexes proceeds with positive cooperativity. For all other **SCR•glycan** systems, the best fit of the NMR peak shift data was obtained with a 1:1 **SCR•glycan** binding model indicating that higher stoichiometry complexes did not form.

Table 1. Complexation (K_a), dimerization (K_d), and cumulative (β) association constants of **SCR001** and **SCR017 – SCR023** with the five octyloxy pyranosides as determined from ^1H NMR titrations in CD_2Cl_2 at 298 K.^[a,b]

Receptor	β -Glc			α -Glc			β -Man			α -Man			β -Gal		Dilution
	K_1 (M^{-1})	K_2 (M^{-1})	β (M^{-2})	K_1 (M^{-1})	K_1 (M^{-1})	K_2 (M^{-1})	β (M^{-2})	K_1 (M^{-1})	K_2 (M^{-1})	β (M^{-2})	K_1 (M^{-1})	K_d (M^{-1})	K_1 (M^{-1})	K_d (M^{-1})	
SCR001	1.3×10^3	—	—	3.6×10^2	1.2×10^3	3.0×10^1	3.6×10^4 ^c	1.4×10^3	—	—	3.5×10^2	1.4×10^1	—	—	—
SCR017	1.2×10^0	6.8×10^3	8.4×10^3 ^d	— ^e	— ^e	—	—	— ^e	—	—	— ^e	— ^e	— ^e	— ^e	— ^e
SCR018	— ^f	—	—	— ^e	— ^e	—	—	1.1×10^4	—	—	6.6×10^1	— ^e	— ^e	— ^e	— ^e
SCR019	— ^g	—	—	— ^g	2.3×10^0	3.2×10^4	7.4×10^4 ^d	— ^e	—	—	— ^e	— ^e	— ^e	— ^e	— ^e
SCR020	— ^e	—	—	— ^e	1.1×10^2	—	—	2.6×10^2	1.1×10^3	2.8×10^5 ^d	— ^e	— ^e	— ^e	— ^e	— ^e
SCR021	2.7×10^0	1.2×10^4	3.3×10^4 ^d	— ^f	— ^e	—	—	— ^e	—	—	— ^f	— ^e	— ^f	— ^e	— ^e
SCR022	1.2×10^2	6.9×10^2	8.1×10^4 ^c	— ^f	— ^e	—	—	— ^e	—	—	— ^f	— ^e	— ^f	— ^e	— ^e
SCR023	— ^e	—	—	— ^f	— ^f	—	—	— ^f	—	—	— ^f	— ^e	— ^f	— ^e	8.5×10^3

^[a] Titrations were performed in triplicate for **SCR019**• **β -Man** to determine experimental error, and the standard deviations of K_a were $3.2 \times 10^2 \text{ M}^{-1}$ (15% error). ^[b] K_a s are based on 1:1 binding models that also consider K_d when appropriate. ^[c] Cumulative association constant ($\beta = K_1 K_2 (\text{M}^{-2})$) involving a 2:1 **SCR**•**glycan** binding model where K_1 and K_2 correspond to 1:1 and 2:1 **SCR**•**glycan** association constants, respectively. ^[d] Cumulative association constant ($\beta = K_1 K_2 (\text{M}^{-2})$) involving a 1:2 **SCR**•**glycan** binding model where K_1 and K_2 correspond to 1:1 and 1:2 **SCR**•**glycan** association constants, respectively. ^[e] No detectable binding/dimerization above the threshold of $K_a = 3.0 \times 10^1 \text{ M}^{-1}$. ^[f] No NMR peak shifts above the threshold of $\Delta\delta > 0.02 \text{ ppm}$. ^[g] Titration data fit unsatisfactory, although evidence of binding exists.

Analysis of the binding data of all tetrapodal SCRs reveals that varying the heterocycles has substantial impact on their K_a s and selectivities, and that, just by varying the heterocycle, binding can be varied from no binding, to promiscuous binding, to completely selective binding (**Figure 2**). The first tetrapodal SCR, **SCR001**, was promiscuous with a slight preference for **β -Man**.^[19a] Previous manipulations of the bonding between the biaryl core or the heterocycle^[19b] to include furan, thiophene, or imidazoles (**SCR002 – SCR009**) either lead to promiscuous binders (**SCR002, SCR004 – SCR006**), a slightly selective binder with weak overall binding (**SCR003**), or a lack of binding altogether (**SCR007 – SCR009**). In contrast, the heterocycles explored here resulted in receptors that were highly specific. **SCR017, SCR021, and SCR022** were completely specific and bind only **β -Glc**. **SCR018** and **SCR020**, alternatively, bind preferentially to **α -Man**, an important biological target for which there is no other specific SCR, and **SCR019** binds only **β -Man**. Of these new SCRs, all bound their preferred substrate with either $\text{Log}(K_1)$ or $\text{Log}(\beta) > 3.9$.

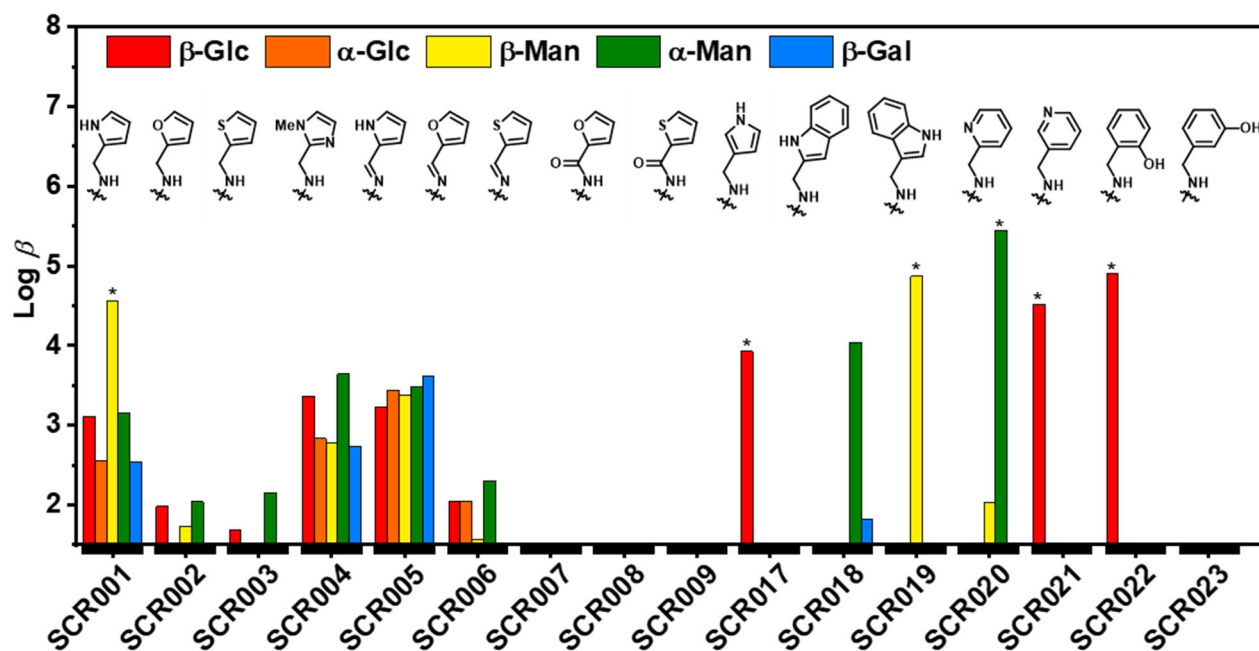


Figure 2. Log (β) values of the receptors towards different glycans. The baseline is set to Log (β) of 1.5 ($K_a = 3.0 \times 10^1 \text{ M}^{-1}$) as the threshold below which binding cannot be reported accurately from ^1H NMR titrations. Star over bars indicate a cumulative association constant ($\beta = K_1 \cdot K_2$), alternatively Log (β) = Log K_1 .

Computational modelling

Calculations were performed on all 40 **SCR•glycan** combinations to determine the structures of the supramolecular complexes and estimate their binding energies, ΔE s. In addition to clarifying why the changes to the heterocycles and linkage regiochemistry cause dramatic changes in the affinities and the selectivities of the SCRs towards different glycans, such computational guidance is invaluable in the design of future of SCRs because the laborious syntheses and binding studies on SCRs that have insignificant K_a s could be avoided. Estimating K_a s and host•guest structures must involve a long timescale molecular dynamics (MD) simulations married to the accuracy of *ab initio* potentials. While long-time MD simulations of the **SCR•glycan** complexes using an empirical potential function (force field) could reproduce the experimental K_a s, such modeling remains largely underexplored because of the lack of appropriate parameters for carbohydrate-receptor interactions in non-aqueous solutions, and, as such, is too preliminary to apply here. The alternative, *ab initio* MD, which utilizes density-functional theory (DFT) for propelling the equation of motions, is prohibitively expensive for systematically deriving the structures of molecular complexes.^[25] Another computation hurdle is the complexity of the **SCR•glycan** conformational space. The inherent flexibility of the glycosides and the SCRs,

encoded in ring puckering and rotatable bonds, respectively, results in a substantially greater computational challenge than, for instance, docking a conventional ligand into the active site of a rigid protein. A systematic search on the potential-energy landscape will underexplore the host•guest complex conformational space, so metadynamics techniques must be employed to accelerate the sampling process. To tackle these challenges, we applied a cascade-like protocol^[26] that starts by sampling using replica-exchange MD at the force field level and refines the promising candidates, immersed in the continuous solvation model, at the DFT level of theory.

With this approach, all 40 1:1 **SCR•glycan** complexes combinations and the 7 complexes with higher (2:1 or 1:2) binding stoichiometries were screened to determine ΔE (**Table 2**) and supramolecular structures. The computational ΔE s lack vibrational and entropic contributions and cannot therefore be compared directly to experimentally derived ΔG values. The ΔH values for association of **SCR001** with α -Glc, β -Gal, β -Glc, and β -Man in CDCl₃, have, however, been reported,^[19a] and these can be compared more directly to the ΔE values derived from the computational approach. Experimentally, the magnitude of ΔH of 1:1 binding of the glycans to **SCR001** follow the trend β -Man > β -Glc > β -Gal > α -Glc, and DFT-based computational methods reproduce this trend. In addition to this qualitative accord, quantitative agreement between the experimentally- and computationally-derived binding values is observed. For example, the difference in ΔH between **SCR001• β -Man** and **SCR001• α -Glc** is 7.4 kcal mol⁻¹, and in ΔE , by comparison, was 7.1 kcal mol⁻¹. Furthermore, experimentally β -Glc, β -Man and α -Man bind 4-5 times stronger to **SCR001** compared to β -Gal and α -Glc, which at room temperature translates to approximately 1 kcal mol⁻¹ difference in binding free energy, and theory predicts a 4 kcal mol⁻¹ energy gap between these two groups. Similar matches exist when comparing all 1:1 complexes, which have measurable K_a . Besides **SCR001**, only **SCR018• α -Man**, **SCR018• β -Gal** and **SCR020• β -Man** have substantial K_1 . Theory confirms that SCR018 binding energy with α -Man and β -Gal is at least 3.5 kcal mol⁻¹ greater than with the next glycan. Similarly, theory predicts that **SCR020** forms the strongest interaction with β -Man. These matches confirm the ability of our computational approach to predict accurately which host-guest pairs have ΔH of association.

Table 2. Computationally-derived, solvent-corrected binding energies (kcal mol⁻¹) of the **SCR•glycan** complexes. The underlined complexes have been determined experimentally to form supramolecular complexes, and the bolded numbers are experimentally determined.^[19]

	β-Glc		α-Glc	β-Man		α-Man		β-Gal
Receptor	ΔE₁	ΔE₂	ΔE₁	ΔE₁	ΔE₂	ΔE₁	ΔE₂	ΔE₁
SCR001	<u>-26.6</u>	-	<u>-21.9</u>	<u>-29.0</u>	<u>-26.3^a</u> <u>-25.4^b</u>	<u>-27.1</u>	-	<u>-22.1</u>
ΔH(exp)	-16.5±0.1		-13.1±0.5	-20.5±0.8				-15.4±0.8
SCR017	-24.0	<u>-23.0^a</u>	-24.7	-29.6	-	-24.2	-	-24.7
SCR018	-20.1	-	-19.3	-22.1	-	<u>-25.7</u>	-	<u>-26.5</u>
SCR019	-25.7	-	-26.5	-22.7	<u>-24.7^a</u>	-25.7	-	-19.3
SCR020	-28.2	-	-25.3	<u>-32.5</u>	-	-31.1	<u>-24.9^a</u>	-27.5
SCR021	-22.4	<u>-29.1^a</u>	-21.9	-23.9	-	-24.4	-	-21.1
SCR022	-18.4	<u>-31.8^b</u>	-20.8	-18.5	-	-19.3	-	-19.4
SCR023	-19.0	-	-25.9	-26.9	-	-26.5	-	-20.0

^a 1:2 **SCR•glycan** binding model, ^b 2:1 **SCR•glycan** binding model.

1:1 Binding Structures

Structural analysis was carried out on the 11 **SCR•glycan** complexes that were confirmed experimentally to bind. These analyses revealed three types of 1:1 supramolecular structures occurred, which are referred to as “center-parallel”, “center-perpendicular”, and “off-center”. The structures of **SCR001•β-Glc**, **SCR001•α-Glc**, and **SCR001•β-Man** are used to illustrate the differences in the three binding geometries, respectively (**Figure 3**). Structures of other complexes are shown in **Figures S80-S82** in the Supporting Information, including xyz Cartesian coordinates for all analyzed structures. The **SCR001•β-Glc** is a *center-parallel* binder in that the

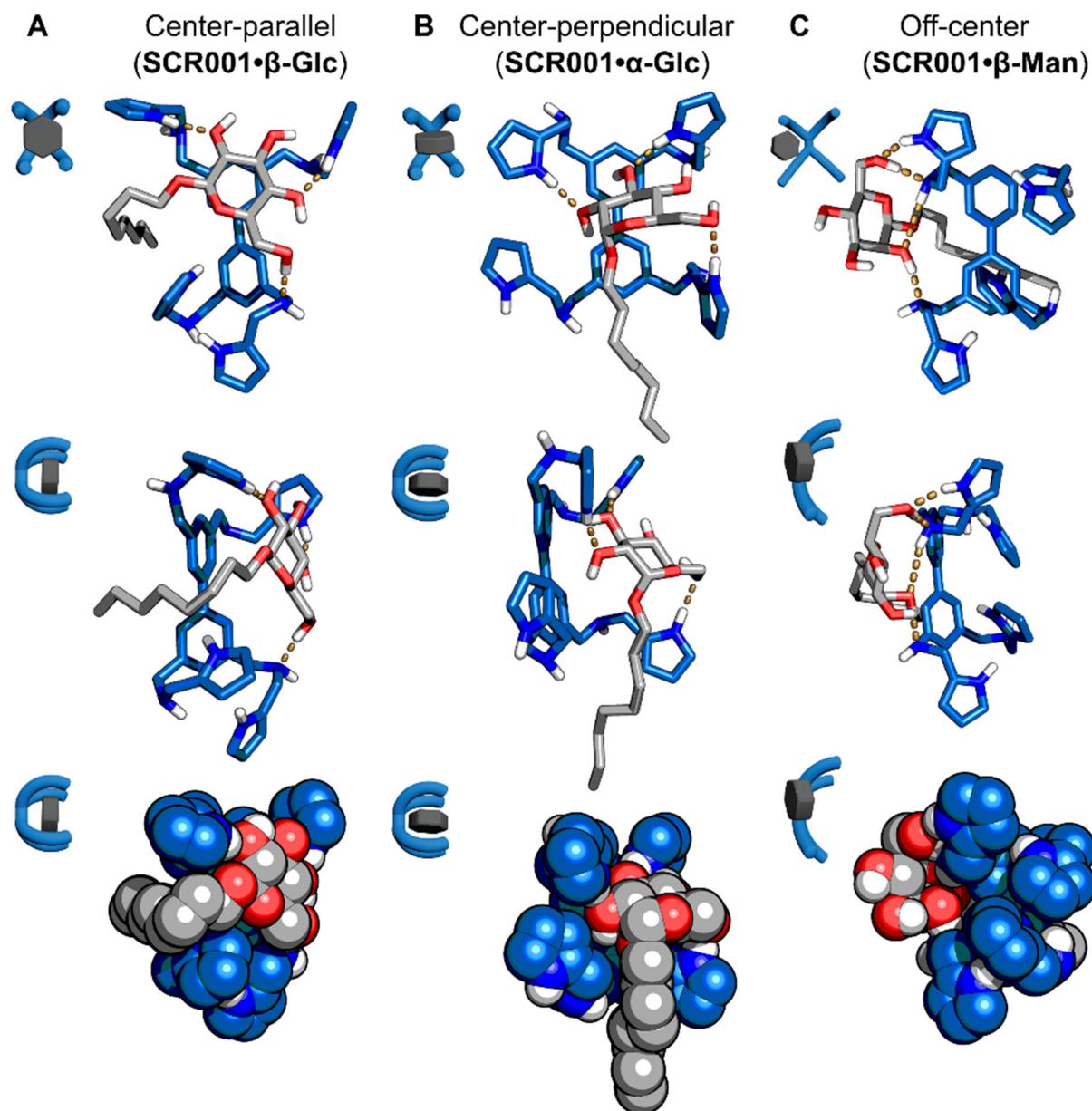


Figure 3. Three different types of 1:1 **SCR001•glycan** complexes. The top two rows show a model of the complex in two different perspectives (front and side), along with a simplified view in the top left corner. The gold dashed lines indicate H-bonding contacts. The bottom row shows a space-filling model of the host•guest complexes in the side-view perspective.

glycan rests above the center of the biaryl ring, where the diagonal axis of the hexose ring (a line connecting the C¹ and C⁴ carbons) and the biaryl group are parallel (**Figure 3A**). In this complex, several noncovalent interactions operate in concert to stabilize the supramolecular association. The β-glucoside accepts two H-bonds from the pyrrole rings – N_{pyr}–H•••O² and N_{pyr}–H•••O⁴ – and

donates an $O^6-H\cdots N_{amine}-H$ bond. The structure is further stabilized by $C-H^{1,3,5}\cdots\pi$ interactions between the hexose α -face and the benzene rings. The involvement of three of four receptor arms is consistent with previously reported spectroscopic and computational data,^[19a] providing further validation that our computational approach provides accurate structural data. The equatorial orientation of the octyloxy group appears to be structural prerequisite for the formation of a *center-parallel* complexes, although having an equatorial octyloxy group does not necessarily dictate the formation of a *center-parallel* structure. It should be noted that the interaction involving the α -face hydrogens of the hexose observed in the models agreed well with the NMR titration data reported previously,^[19] where $C-H^{1,3,5}\cdots\pi$ interactions caused an average upfield shift of 0.3 ppm of the hexose. Furthermore, the $N_{pyr}-H$ proton shifts downfield by 0.16 ppm validates the H-bonds observed in the model. Other complexes that adopt the *center-parallel* geometry include **SCR019• β -Man** and **SCR021• β -Glc** (see Supporting Information for details), both of which form strongly cooperative higher stoichiometry complexes.

In the *center-perpendicular* geometry adopted by **SCR001• α -Glc**, the diagonal axis of the hexose is perpendicular to the biaryl axis (**Figure 3B**). In this complex, the **α -Glc** accepts three H-bonds from the pyrrole rings – $N_{pyr}-H\cdots O^6$, $N_{pyr}-H\cdots O^3$ and $N_{pyr}-H\cdots O^2$. While the hexose is still centered over the biaryl core, the axial octyloxy group at the anomeric carbon forces the perpendicular orientation and prevents the formation of stabilizing $C-H\cdots\pi$ interactions. As a consequence, the **α -Glc** binds weaker than **β -Glc** to the **SCR001** by 4.2 kcal mol⁻¹. Again these structures agreed well with the NMR titration data, where a pyrrolic H-bond caused a 0.11 ppm downfield shift in $N_{pyr}-H$.^[19] The titration data also indicates that both H³ and H⁴ experience an upfield shift of 0.18 ppm and 0.29 ppm respectively. Because these Hs are on the opposite faces of the hexose ring, the **α -Glc** must adopt perpendicular orientation to the biaryl core so both protons can form $C-H\cdots\pi$ interactions, which, again, is entirely consistent with the lowest energy structure. Other complexes that adopt the *center-perpendicular* geometry include **SCR001• β -Gal**, **SCR020• β -Man** and **SCR020• α -Man**. In all cases, the K_1 is reported to be 100–350 M⁻¹, and, of these, only **SCR020• α -Man** forms a higher stoichiometry complex that is only weakly cooperative.

In the *off-center* geometry adopted by **SCR001• β -Man**, the hexose does not bind directly above the biaryl core, but instead it rests between the two arms on the same side of the long axis of the biaryl group of the receptor (**Figure 3C**). Although this geometry lacks stabilizing $C-H\cdots\pi$ interactions, the two amine groups and two 2-pyrrole heterocycles form a pocket that binds the axially oriented O^2-H of **β -Man**. The $N_{amine}-H\cdots O^2-H\cdots N_{amine}-H$ bonding motif, is further stabilized by adjacent $N_{pyr}-H\cdots O^6-H$ and $O^6-H\cdots N_{amine}-H$ bonds. As a result of these multiple interactions, the **SCR001• β -Man** has the strongest predicted ΔE amongst all the computed

SCR001•glycan complexes. As with the above geometries, this structure was in good agreement with the NMR data where an upfield 0.15 ppm shift is observed in $N_{\text{pyr}}\text{--H}$.^[19] The binding also causes upfield shifts of H^1 and $H^{3,4,5}$ protons. These protons are readily accessible for heterocycles to form $\text{C--H}\cdots\pi$ interactions, whereas access to the H^2 proton is blocked by the alkyl chain. Furthermore, the *off-center* structures can form symmetric higher-stoichiometry complexes by stabilizing another glycan in an identical binding pocket, as observed experimentally for **SCR001•(β -Man)₂** and **SCR001•(α -Man)₂** in CDCl_3 .^[19a] Other complexes that adopt the *off-center* geometry include **SCR017• β -Glc**, **SCR018• α -Man**, **SCR018• β -Gal**, and **SCR022• β -Glc**. Among these complexes, only **SCR018• α -Man** does not form higher stoichiometry complexes because the bulky heterocyclic groups block the second binding event.

Higher-Order Complexation

In the titrations, it was found that several **SCR•glycan** combinations formed higher-order (1:2 and 2:1) complexes that, in some cases, assemble with substantial positive cooperativity ($K_2 \gg K_1$). The origin of the cooperativity can be understood by examining the computed structures. **SCR001•(β -Man)₂**, **SCR017•(β -Glc)₂**, **SCR019•(β -Man)₂**, **SCR020•(α -Man)₂**, **SCR021•(β -Glc)₂** all form 1:2 **SCR•glycan** supramolecular complexes. Of those, only the **SCR001• β -Man** complex has $K_2 < K_1$. The 1:2 binding of **SCR001• β -Man** in CDCl_3 has been explained previously.^[19a] We had shown using computational and experimental data that the β -Man occupied identical binding pockets on opposite faces of the biaryl ring, and the more sophisticated computational approach presented here reproduces this result – the symmetric complex (**Figure 4A**) is more stable than any other complex of **SCR001•(β -Man)₂** found during the structural search. Most importantly, this molecular modeling reveals why the **SCR001•(β -Man)₂** has minimal cooperativity – there are no stabilizing interactions that form between the two glycan guests.

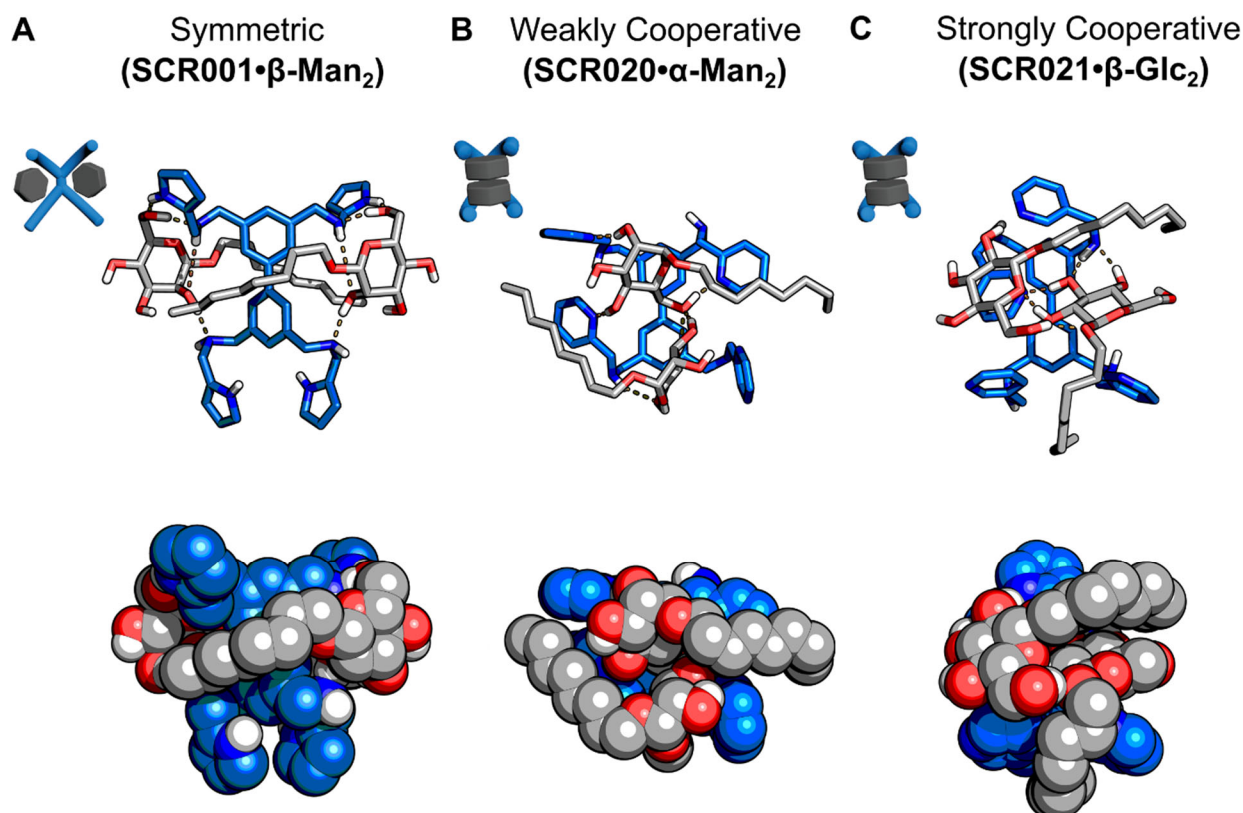


Figure 4. Three different 1:2 **SCR•glycan** complexes with varying degrees of cooperativity, along with a simplified view in the top left corner. The top row shows the stick models and the bottom row shows space-filling models. The gold dashed lines indicate H-bonding contacts.

Unlike the higher-order complex with weak cooperativity, computations revealed that all the 1:2 **SCR•glycan** complexes where $K_2 \gg K_1$ (**SCR020•(α-Man)₂**, **SCR021•(β-Glc)₂**, **SCR017•(β-Glc)₂**, **SCR019•(β-Man)₂**) are stabilized by sugar-sugar contacts (**Figure 4B** and **C** and **Figure S83**), and the ratio K_2/K_1 increases with increasing number of sugar - sugar contacts. In the case of **SCR020•(α-Man)₂**, the K_2/K_1 is approximately 4, and interactions between the two **α-Man** species are limited to one $O^3-H \cdots O^6-H$ contact that contributes 5 kcal mol⁻¹ of stabilization. The remaining 1:2 **SCR•glycan** complexes, **SCR017•(β-Glc)₂**, **SCR019•(β-Man)₂** and **SCR021•(β-Glc)₂**, however, feature K_2 s that are 3-4 orders of magnitude larger than K_1 , which means that the binding of the second hexose is strongly cooperative. Computational methods confirm that in each of these complexes the binding of the first hexose is relatively weak. The binding of the second hexose, however, is at least as strong as the binding of the first sugar. Furthermore, the interaction energy between two hexoses in these strongly cooperative

complexes is predicted to be $>10 \text{ kcal mol}^{-1}$ in magnitude because at least two H-bonds are formed between each pair of hexoses.

Titration also showed that **SCR001**•**β-Man** and **SCR022**•**β-Glc**, form 2:1 complexes in which multiple receptors bind to one sugar. Previously, we reported that **SCR001**₂•**β-Man** exists in complex equilibrium with **SCR001**•**β-Man** and **SCR001**•**β-Man**₂ species.^[19a] Here, the two **SCR001** receptors were wrapped around the glycan in a “hugging” complex (**Figure 5A**), where the axially oriented O²–H and O⁶–H that occur in a gauche conformation of **β-Man** stabilizes the complex by forming four H-bonds with both receptors. However, the second receptor adds only one more H-bond than those already present in the 1:1 **SCR001**•**β-Man** complex, and there is little interaction between the receptors. In effect, the binding energy of the second receptor is weaker by $3.4 \text{ kcal mol}^{-1}$ than the binding of the first, the experimentally determined K_2 is 2 orders of magnitude smaller than K_1 indicating a *non-cooperative* complex. In the case of **SCR022**₂•**β-Glc**, however, we observe that K_2/K_1 is ~ 6 , indicating a weakly *cooperative* complex (**Figure 5B**). Here, the theory predicts that the phenolic O–H groups, which serve as H-bond donors and acceptors, interact strongly with the glycans. In an unbound receptor, these O–H groups are locked by a strong H-bond with adjacent amine group, O_{ph}–H...N_{amine}–H. Upon interaction with **β-Glc**, however, two receptors interact with each other in a “spooning” complex, which creates a hydrophilic pocket for binding of the hexose. The **β-Glc** accepts a N_{amine}–H...O⁶–H H-bond from and donates a O⁶–H...O_{ph}–H H-bond to the same receptor. The other receptor supports this binding with an additional N_{amine}–H...O¹ H-bond and a pronounced C–H^{1,3,5}...π interaction between α-face of the hexose and a neighboring phenyl ring. Finally, theory predicts that the binding of the second receptor is stronger by more than 10 kcal mol^{-1} than the first, although our computational approach may be overestimating the strength of the van der Waals interactions between receptors.

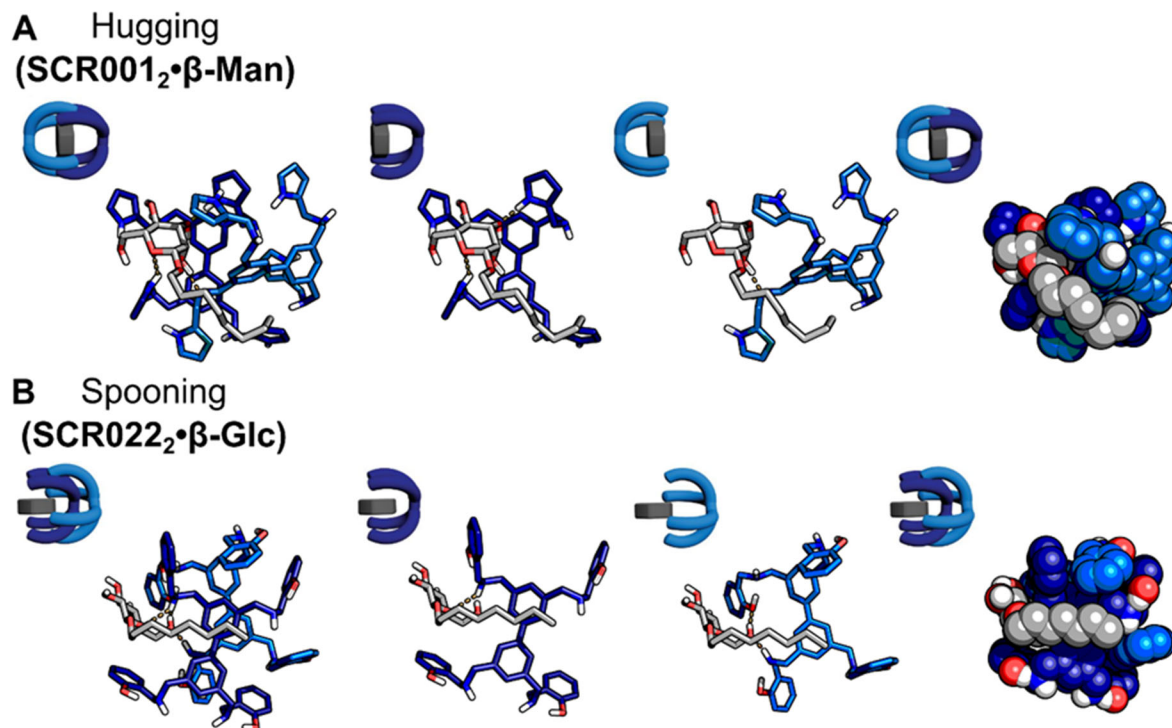


Figure 5. Two different 2:1 **SCR•glycan** complexes shown using stick and space-filling model, along with simplified view in the top left corner. The columns show whole complex, each pair of **SCR•glycan** components. The top row shows **SCR001₂•β-Man** complex adopting a “hugging” geometry. The bottom row shows 2:1 **SCR022₂•β-Glc** complex adopting a “spooning” geometry. The gold dashed lines indicate H-bonding contacts.

Conclusions

Here, we have reported the synthesis of seven new C_{2h} tetrapodal SCRs, and their binding against five biologically relevant octyloxy O-glycans was studied in CH₂Cl₂ by mass spectrometry, ¹H NMR titrations, and molecular modeling. These receptors varied only by the composition and attachment regiochemistry of the heterocyclic binding arms. Of these, three were selective binders of **β-Glc**, one was selective for **β-Man**, and two bound preferentially **α-Man**. By comparison with previously synthesized tetrapodal SCRs, we see that in only three generations, this class of receptors has been advanced from promiscuous, moderate binders to strong and highly selective binders. All 40 1:1 **SCR•glycan** supramolecular complexes and 7 additional higher stoichiometry (1:2 and 2:1) complexes were analyzed by molecular modeling, whose binding energies and structural models agreed well with experimental data. Binding was dictated by C–H···π contacts and H-bonding, and the 1:1 structures formed three general binding motifs that are

classified as “center-parallel”, “center-perpendicular”, or “off-center”. Modeling also revealed that when the 1:2 **SCR•glycan** complexes formed with substantial positive cooperativity, it was the result of new glycan-glycan contacts.

This work promises to advance the field of SCRs in several important ways. First, we echo the findings of others in the field who show the importance of the heterocyclic binding groups for manipulating selectivity and binding affinity. Secondly, the advent of accurate molecular modeling approaches provides the ability to predict accurately how a particular SCR will bind to a receptor, so arduous and time-consuming syntheses and binding studies can be avoided. The antiviral performance of these new receptors is currently being assessed, and how the increased selectivity and binding affinity affect antiviral activity and toxicity is a topic of substantial current interest. Moreover, having a library of SCRs with different selectivities could form the basis of powerful new sensor platforms for the *in situ* detection of glycans, a major scientific challenge for which there is no current solution.

Experimental Section

General Procedure. SCRs were synthesized following the procedure described below unless otherwise noted. PPh_3 (5 mmol, 5 eq) was added to a stirring solution of **1** (1 mmol, 1 eq) in THF (5 mL) at room temperature. The reaction was refluxed under Ar atmosphere for 1 h before the addition of the heteroarylaldehyde (5 mmol, 5 eq) at room temperature. The reaction mixture was refluxed for additional 48 h, cooled to room temperature, and concentrated under reduced pressure. The residue was dissolved in MeOH (5 mL), and NaBH_4 (10 mmol, 10 eq), was added in portions at room temperature under Ar atmosphere followed by stirring for 16 h. The reaction mixture was concentrated under reduced pressure, treated with CHCl_3 (30 mL) and H_2O (30 mL), and the organic layer was separated. The aqueous layer was extracted with CHCl_3 (3 x 30 mL), and the combined organic layers were dried over anhydrous Na_2SO_4 , filtered and concentrated under reduced pressure to give the crude product, which was purified by column chromatography (SiO_2 , 9:1:0.5 CHCl_3 : MeOH : NH_3 (aq)) to give the pure product.

Synthesis of 1,1',1'',1'''-([1,1'-biphenyl]-3,3',5,5'-tetrayl)tetrakis(*N*-((1*H*-pyrrol-3-yl)methyl)methanamine) (SCR017). Following the General Procedure, **SCR017** was synthesized from **1** and 1*H*-pyrrole-3-carbaldehyde and purified by column chromatography (SiO_2 , 9:1:0.5 CHCl_3 : MeOH : NH_3 (aq)) to provide a pale yellow solid (393 mg, 67%). ^1H NMR (300 MHz, CD_2Cl_2) δ 7.48 (s, 4H), 7.31 (s, 2H), 6.76 (d, J = 2.3 Hz, 8H), 6.21 (t, J = 2.15 Hz, 4H), 3.88 (s, 8H), 3.74 (s, 8H); ^{13}C NMR (75 MHz, CD_2Cl_2) δ = 140.93, 127.23, 125.58, 122.09, 117.68,

116.03, 108.25, 108.21, 52.93, 45.65; HRMS (ESI): m/z calcd for $C_{36}H_{42}N_8$ $[M+H]^+$: 587.3605, found 587.3606.

Synthesis of 1,1',1'',1'''-([1,1'-biphenyl]-3,3',5,5'-tetrayl)tetrakis(N-((1H-indol-2-yl)methyl)methanamine) (SCR018). Following the General Procedure, **SCR018** was synthesized from **1** and 1H-indole-2-carbaldehyde and purified by column chromatography (SiO_2 , 9:1:0.5 $CHCl_3$: MeOH : NH_3 (aq)) to provide a pale yellow solid (661 mg, 84%). 1H NMR (300 MHz, $CDCl_3$) δ 8.75 (s, 4H), 7.57 (d, J = 7.4 Hz, 4H), 7.40 (s, 4H), 7.30 (s, 4H), 7.27 (s, 2H), 7.12 (m, 8H), 6.38 (s, 4H), 3.97 (s, 8H), 3.81 (s, 8H), 2.10 (s, 4H); ^{13}C NMR (75 MHz, CD_2Cl_2) δ 141.47, 141.03, 137.95, 136.57, 128.85, 127.63, 126.14, 121.74, 120.38, 119.88, 111.16, 100.76, 77.96, 46.61.; HRMS (ESI): m/z calcd for $C_{52}H_{50}N_8$ $[M+H]^+$: 787.4231, found 787.4230.

Synthesis of 1,1',1'',1'''-([1,1'-biphenyl]-3,3',5,5'-tetrayl)tetrakis(N-((1H-indol-3-yl)methyl)methanamine) (SCR019). Following the General Procedure, **SCR019** was synthesized from **1** and 1H-indole-3-carbaldehyde and purified by column chromatography (SiO_2 , 9:1:0.5 $CHCl_3$: MeOH : NH_3 (aq)) to provide a pale yellow solid (763 mg, 97%). 1H NMR (300 MHz, DMSO) δ 10.83 (s, 4H), 7.61 (d, J = 7.7 Hz, 4H), 7.52 (d, J = 8.0 Hz, 4H), 7.37 - 7.27 (m, 6H), 7.24 (s, 4H), 7.04 (t, J = 7.0, 4H), 6.92 (t, J = 6.9, 4H), 3.88 (s, 8H), 3.81 (s, 4H); ^{13}C NMR (75 MHz, CD_2Cl_2) δ 141.52, 141.01, 136.79, 127.84, 127.41, 126.24, 123.68, 122.08, 119.49, 118.83, 113.84, 111.68, 78.05, 43.98; HRMS (ESI): m/z calcd for $C_{52}H_{50}N_8$ $[M+H]^+$: 787.4231, found 787.4227.

Synthesis of 1,1',1'',1'''-([1,1'-biphenyl]-3,3',5,5'-tetrayl)tetrakis(N-(pyridin-2-yl)methyl)methanamine) (SCR020). Following the General Procedure, **SCR020** was synthesized from **1** and picolinaldehyde and purified by column chromatography (SiO_2 , 9:1:0.5 $CHCl_3$:MeOH : NH_3 (aq)) to provide a pale yellow gum (628 mg, quant.). 1H NMR (300 MHz, $CDCl_3$) δ 8.50 (s, 4H), 7.56 (t, J = 7.6 Hz, 4H), 7.47 (s, 4H), 7.33-7.23 (m, 6H), 7.08 (t, J = 7.1 Hz, 4H), 3.91 (s, 8H), 3.85 (s, 4H), 2.53 (br s, 4H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.67, 149.24, 141.33, 140.71, 136.41, 127.14, 125.89, 122.37, 121.91, 54.56, 53.52; HRMS (ESI): m/z calcd for $C_{40}H_{42}N_8$ $[M+H]^+$: 635.3605, found 635.3607.

Synthesis of 1,1',1'',1'''-([1,1'-biphenyl]-3,3',5,5'-tetrayl)tetrakis(N-(pyridin-3-yl)methyl)methanamine) (SCR021). Following the General Procedure, **SCR021** was synthesized from **1** and nicotinaldehyde and purified by column chromatography (SiO_2 , 9:1:0.5 $CHCl_3$:MeOH

: NH₃ (aq)) to provide a pale yellow gum (630 mg, quant.). ¹H NMR (300 MHz, CDCl₃) δ 8.61 (s, 4H), 8.52 (d, *J* = 3.2 Hz, 4H), 7.73 (d, *J* = 7.8 Hz, 4H), 7.48 (s, 4H), 7.36-7.22 (m, 6H), 4.07-3.61 (m, 16H), 1.88 (br s, 4H); ¹³C NMR (75 MHz, CDCl₃) δ 149.80, 148.57, 141.46, 140.74, 135.86, 135.53, 127.00, 125.88, 123.42, 53.24, 50.66; HRMS (ESI): *m/z* calcd for C₄₀H₄₂N₈ [M+H]⁺: 635.3605, found 635.3603.

Synthesis of 2,2',2'',2'''-((((1,1'-biphenyl)-3,3',5,5'-tetrayltetrakis(methylene))tetrakis(azanediyl))tetrakis(methylene))tetraphenol (SCR022).

Following the General Procedure, **SCR022** was synthesized from **1** and 2-hydroxybenzaldehyde and purified by column chromatography (SiO₂, 9:1:0.5 CHCl₃: MeOH : NH₃ (aq)) to provide a yellow solid (660 mg, 95%). ¹H NMR (300 MHz, CDCl₃) δ 7.46 (s, 4H), 7.26-7.16 (m, 6H), 7.04 (d, *J* = 7.5 Hz, 4H), 6.92 – 6.76 (m, 8H), 4.08 (s, 8H), 3.91 (s, 8H); ¹³C NMR (75 MHz, CDCl₃) δ 158.07, 141.53, 139.56, 128.93, 128.65, 127.66, 126.55, 122.17, 119.25, 116.43, 52.51, 52.00; HRMS (ESI): *m/z* calcd for C₄₄H₄₆N₄O₄ [M+H]⁺: 695.3592, found 695.3586.

Synthesis of 3,3',3'',3'''-((((1,1'-biphenyl)-3,3',5,5'-tetrayltetrakis(methylene))tetrakis(azanediyl))tetrakis(methylene))tetraphenol (SCR023).

Following the General Procedure, **SCR023** was synthesized from **1** and 3-hydroxybenzaldehyde and purified by column chromatography (SiO₂, 8:2:2 CHCl₃:MeOH : NH₃ (aq)) to provide a white solid (236 mg, 34%). ¹H NMR (800 MHz, CD₂Cl₂) δ 7.52 (s, 4H), 7.27 (s, 2H), 7.18 (t, *J* = 7.6 Hz, 4H), 6.85 (s, 4H), 6.81 (d, *J* = 7.1 Hz, 4H), 6.74 (d, *J* = 7.4 Hz, 4H), 3.83 (s, 8H), 3.77 (s, 8H); ¹³C NMR (200 MHz, CD₂Cl₂) δ = 157.16, 141.01, 140.74, 140.14, 129.56, 127.77, 126.01, 119.65, 115.06, 114.34, 52.60, 52.49; HRMS (ESI): *m/z* calcd for C₄₄H₄₆N₄O₄ [M+H]⁺: 695.3592, found 695.3586.

ESI Mass Spectrometry Analysis of Receptor-Glycan Binding. A Bruker ultra-high resolution maXis-II / ETD ESI-q-TOF system was used to study receptor-glycan complex formation. The samples of the SCRs and the glycans were prepared in 1 mM in CH₂Cl₂, diluted to 1 μM with 40% CH₂Cl₂ in MeCN. The samples of SCR+glycan for complex mass screening were prepared using the aforementioned diluted solutions and mixed in a one-to-one fashion. All samples were analyzed via direct infusion into the spectrometer with a syringe pump. Theoretical isotopic distributions were calculated using the Chemistry tool in Bruker's Compass DataAnalysis software.

NMR Titrations and Peak Shift Fittings ^1H NMR titrations were performed in CD_2Cl_2 , unless otherwise noted, at a field strength of either 700 or 800 MHz at 298 K. The pyranosides' ^1H NMR peak assignments were previously reported^[2]. The experimental temperatures were verified through calibration with a 100% methanol standard^[4]. The addition of pyranoside to a SCR CD_2Cl_2 solution or vice versa resulted in the perturbation of the chemical shifts (δ) corresponding to resonances of both SCR and pyranoside. This is the result of an exchange process involving SCR (H) and pyranoside (G) equilibria products interchanging fast on the NMR timescale, resulting in the averaging of chemical shifts of protons in differing chemical environments. Accordingly, equilibrium constants (K) can be quantified by first defining a model that includes the correct set of equilibria, calculating the hypothetical concentrations of equilibrium species and the corresponding chemical shifts, and finally fitting the resulting data to the experimental results. For the 2:1 association of **SCR022• β -Man** and the 1:2 association of **SCR017• β -Glc**, **SCR021• β -Glc**, **SCR019• β -Man**, **SCR020• α -Man**, fitting was carried out as previously described.^[19b]

Theoretical Modeling. The search protocol for finding the minimum-energy binding structures involved the following steps:

1. The initial sampling of the conformational space has been performed using a replica-exchange molecular dynamics^[27] (REMD) protocol implemented in Gromacs-2018.2.^[28] First, the parameters for 8 SCRs and 5 sugars were derived from a CHARMM36 force field^[29] using the charmm-gui web-interface^[30] and combined into 40 **SCR•glycan** complexes. Initial hand-generated conformations were first equilibrated in the gas phase for 1 ns at 300 K, controlled by a v-rescaling thermostat,^[31] and the final structure was used as a starting geometry for REMD simulations. Each complex involved 16 replicas distributed in 300-400 K temperatures according to the Monte-Carlo criteria.^[32] The temperatures were chosen to yield an exchange rate of 0.65. The trajectories were simulated for a 500 ns each, with a 1 fs time step and an attempted exchange between trajectories neighboring in temperature space every 10 ps.
2. Each 300 K trajectory was subsampled every 100 ps to generate a set of initial structures visited during the simulations. The structures were then clustered using a single linkage algorithm with a loose 2.5 Å cutoff for root means square deviations (RMSD) between all atoms. In several cases in which the supramolecular complex was found dissociated at 300 K, we restarted the REMD from a different equilibrated structure. The clustering unraveled that some complexes (for instance **SCR018• β -Man** or **SCR019• β -Gal**) occupy only several conformational basins, while other (**SCR022• α -Man** or **SCR020• α -Glc**) visit

numerous lowly populated states. The overall distribution of cluster sizes and RMSD of the 20 most abundant clusters for each complex is shown in **Figure S79**. For each system we collected central structures of the 20 most abundant clusters, which were used as starting geometries for subsequent density-functional theory calculations.

3. The central structures of the most abundant complexes were next geometry-optimized using dispersion-corrected Generalized-Gradient Approximation DFT, PBE+vdW^{TS} [33] and *light* settings for a basis set, as implemented in FHI-aims numerical atomic orbitals code.^[34] Although the PBE functional, placed on the second rung of the Jacob's ladder, does not include the exact-exchange term in the functional form, as for instance B3LYP, it provides significant computational speed-up over more popular hybrid DFT methods suitable for screening of the molecular conformational space. Furthermore, the vdW^{TS} parameter-free dispersion correction has been successfully applied to bioorganic systems that require delicate balance between dispersion and electrostatic interactions.^[35] The potential energy of the geometry-optimized structures was collected.
4. To account for the solvent effects, we computed each conformer's solvation energies in CH₂Cl₂ using the SMD continuous solvation model^[36] available in Gaussian16 rev.D01 code.^[37] This model computes solvation energy from electrostatic interactions and appends it by dispersion and entropic terms. It has been shown to yield excellent solvation energies of small molecules in aqueous solution^[38] and best performance among other implicit in non-aqueous solvents^[39] The single-point energy of a complex in a gas-phase and in a solvent model were computed using PBE1PBE+D3/def2-SVP level of theory. The difference between these two values was adopted as a conformational-dependent solvation energy correction and added to the PBE+vdW^{TS} potential energy.

To evaluate whether sampling of the trajectory using REMD limited to 20 largest cluster is sufficient, we compared the lowest DFT-energy of five **SCR•β-Man** complexes (**SCR018-022**) with the lowest energy structures found among 638 conformations generated from low-mode dynamics, as implemented in MAESTRO^[40] Macromodel software, and reoptimized them at the respective DFT level of theory. The lowest-energy conformers found during REMD sampling were, after DFT-optimizations, on average 5.7 kcal mol⁻¹ more stable than those found in low-mode dynamics, which confirms that the REMD protocol appropriately samples the conformational space. The same search strategy was applied to derive solvation-corrected energies of free receptors. The most stable structures of free receptors were dominated by compact, dispersion-stabilized conformers that featured symmetry elements. For carbohydrates, the ⁴C₁ ring-puckering, is the most stable conformations in a gas phase, and a fully extended C1-

octyloxy chain was assumed. Finally, for the analysis we defined binding energy (ΔE) as a difference between energy of the complex (E_{complex}) and energies of relaxed SCR (E°_{SCR}) and glycan (E°_{G}):

$$\Delta E = E_{\text{complex}} - E^{\circ}_{\text{SCR}} - E^{\circ}_{\text{G}}$$

In addition, we defined interaction energy between two sugars ($\Delta E_{\text{sug-sug}}$) as a difference between energy of the two sugars in the binding geometry (E_{ss}) and the energy of each sugar individually ($E_{\text{s1}}, E_{\text{s2}}$):

$$\Delta E_{\text{sug-sug}} = E_{\text{ss}} - E_{\text{s1}} - E_{\text{s2}}$$

For full experimental procedures, synthetic protocols, analytical data, and copies of NMR spectra for all new compounds, ESI mass spectrometry data, NMR titrations and peak shift fittings, and computational modelling data, see the Supporting Information.

Acknowledgements

We thank the Air Force Office of Scientific Research (FA9550-19-1-0220) for financial support. The NMR data (Bruker Avance 300, 700, and 800 MHz) presented herein were collected in part at the CUNY ASRC Biomolecular NMR facility and the NMR facility of the City College of the CUNY. Mass spectra were collected at the CUNY ASRC Biomolecular Mass Spectrometry Facility and the Hunter College Mass Spectrometry Facility of the CUNY. We thank Dr. James Aramini (Biomolecular NMR Facility at ASRC), Dr. Rinat Abzalimov (Biomolecular Mass Spectrometry Facility at ASRC) and Dr. Barney Yoo (Hunter College Mass Spectrometry Facility) for their help in acquiring the spectral data. We thank Ms. Genrietta Yagudayeva (Hunter College) for her help in interpretation of the simulation data.

References

- [1] a) A. Barre, Y. Bourne, E. J. Van Damme, W. J. Peumans and P. Rougé, *Biochimie* **2001**, 83, 645-651; b) B. Ernst, *Lectins and Saccharide Biology* **2000**, 4; c) H. Gabius, **2009**; d) H. G. Garg, M. K. Cowman and C. A. Hales, *Carbohydrate chemistry, biology and medical applications*, Elsevier, **2011**, p; e) Q. Hao, E. J. Van Damme, B. Hause, A. Barre, Y. Chen, P. Rougé and W. J. Peumans, *Plant Physiol.* **2001**, 125, 866-876; f) D. C. Kilpatrick, *Handbook of Animal Lectins: Properties and Biomedical Applications*, Wiley, **2000**, p; g) T. K. Lindhorst, *Essentials of carbohydrate chemistry and biochemistry*, Wiley-VCH Weinheim, **2007**, p; h) R. Loris, H. De Greve, M.-H. Dao-Thi, J. Messens, A. Imberty and L. Wyns, *J. Mol. Biol.* **2000**, 301, 987-1002; i) M. C. Shelton, *J. Chem. Educ.* **2002**, 79, 562.
- [2] A.-S. Vercoutter-Edouart, M.-C. Slomianny, O. Dekeyser-Beseme, J.-F. Haeuw and J.-C. Michalski, *PROTEOMICS* **2008**, 8, 3236-3256.

- [3] a) L. R. Ruhaak, S. Miyamoto and C. B. Lebrilla, *Mol. Cell. Proteom.* **2013**, *12*, 846; b) S. R. Kronewitter, M. L. A. De Leoz, J. S. Strum, H. J. An, L. M. Dimapasoc, A. Guerrero, S. Miyamoto, C. B. Lebrilla and G. S. Leiserowitz, *PROTEOMICS* **2012**, *12*, 2523-2538; c) H. J. An, S. Miyamoto, K. S. Lancaster, C. Kirmiz, B. Li, K. S. Lam, G. S. Leiserowitz and C. B. Lebrilla, *Journal of Proteome Research* **2006**, *5*, 1626-1635; d) G. S. Leiserowitz, C. Lebrilla, S. Miyamoto, H. J. An, H. Duong, C. Kirmiz, B. Li, H. Liu and K. S. Lam, *International Journal of Gynecologic Cancer* **2008**, *18*, 470; e) R. Saldova, L. Royle, C. M. Radcliffe, U. M. Abd Hamid, R. Evans, J. N. Arnold, R. E. Banks, R. Hutson, D. J. Harvey, R. Antrobus, S. M. Petrescu, R. A. Dwek and P. M. Rudd, *Glycobiology* **2007**, *17*, 1344-1356.
- [4] a) S. Y. Kim, J. Zhao, X. Liu, K. Fraser, L. Lin, X. Zhang, F. Zhang, J. S. Dordick and R. J. Linhardt, *Biochemistry* **2017**, *56*, 1151-1162; b) Y. Chen, T. Maguire, R. E. Hileman, J. R. Fromm, J. D. Esko, R. J. Linhardt and R. M. Marks, *Nature Medicine* **1997**, *3*, 866-871; c) 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 and S. N. Byrreddy, *International journal of molecular sciences* **2019**, *20*, 5206.
- [5] C. N. Scanlan, J. Offer, N. Zitzmann and R. A. Dwek, *Nature* **2007**, *446*, 1038-1045.
- [6] G. Ritchie, D. J. Harvey, U. Stroeher, F. Feldmann, H. Feldmann, V. Wahl-Jensen, L. Royle, R. A. Dwek and P. M. Rudd, *Rapid Commun. Mass Spectrom.* **2010**, *24*, 571-585.
- [7] a) X. Han, Y. Zheng, C. J. Munro, Y. Ji and A. B. Braunschweig, *Curr. Opin. Biotechnol.* **2015**, *34*, 41-47; b) S.-C. Tao, Y. Li, J. Zhou, J. Qian, R. L. Schnaar, Y. Zhang, I. J. Goldstein, H. Zhu and J. P. Schneck, *Glycobiology* **2008**, *18*, 761-769.
- [8] a) U. L. Holmskov, *APMIS. Supplementum* **2000**, *100*, 1-59; b) I. P. Fraser, H. Koziel and R. A. B. Ezekowitz, *Seminars in Immunology* **1998**, *10*, 363-372; c) K. HÅKansson and K. B. M. Reid, *Protein Sci.* **2000**, *9*, 1607-1617.
- [9] a) I. M. Vasconcelos and J. T. A. Oliveira, *Toxicon* **2004**, *44*, 385-403; b) L. Barbieri, M. G. Battelli and F. Stirpe, *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1993**, *1154*, 237-282.
- [10] a) F. D. Vrionis, C. J. Wikstrand, P. Fredman, J.-E. Månsson, L. Svennerholm and D. D. Bigner, *Cancer Research* **1989**, *49*, 6645; b) M. Fukuda, S. Carlsson, J. Klock and A. Dell, *J. Biol. Chem.* **1986**, *261*, 12796-12806.
- [11] a) N. Cheung, B. H. Kushner, S. Yeh and S. M. Larson, *Int. J. Oncol.* **1998**, *12*, 1299-1605; b) A. L. Yu, A. L. Gilman, M. F. Ozkaynak, W. B. London, S. G. Kreissman, H. X. Chen, M. Smith, B. Anderson, J. G. Villablanca, K. K. Matthay, H. Shimada, S. A. Grupp, R. Seeger, C. P. Reynolds, A. Buxton, R. A. Reisfeld, S. D. Gillies, S. L. Cohn, J. M. Maris and P. M. Sondel, *N. Engl. J. Med.* **2010**, *363*, 1324-1334; c) N.-K. V. Cheung, I. Y. Cheung, K. Kramer, S. Modak, D. Kuk, N. Pandit-Taskar, E. Chamberlain, I. Ostrovnya and B. H. Kushner, *Int. J. Cancer* **2014**, *135*, 2199-2205.
- [12] a) Y. Nakagawa, T. Doi, T. Taketani, K. Takegoshi, Y. Igarashi and Y. Ito, *Chem.: Eur. J.* **2013**, *19*, 10516-10525; b) A. Dove, *Nat. Biotechnol.* **2001**, *19*, 913-917; c) in *Mammalian Glycan Biosynthesis: Building a Template for Biological Recognition*, pp. 1-32.
- [13] J. S. Lazo and E. R. Sharlow, *Annu. Rev. Pharmacool. Toxicol.* **2016**, *56*, 23-40.
- [14] a) M. Mazik, *ChemBioChem* **2008**, *9*, 1015-1017; b) M. Mazik, *RSC Advances* **2012**, *2*, 2630-2642; c) A. P. Davis and T. D. James, *Carbohydrate receptors*, Wiley-VCH: Weinheim, Germany, **2005**, p; d) A. P. Davis, *Org. Biomol. Chem.* **2009**, *7*, 3629-3638; e) S. Kubik, *Angew. Chem. Int. Ed.* **2009**, *48*, 1722-1725.
- [15] a) X. Wu, Z. Li, X.-X. Chen, J. S. Fossey, T. D. James and Y.-B. Jiang, *Chem. Soc. Rev.* **2013**, *42*, 8032-8048; b) T. D. James, K. S. Sandanayake and S. Shinkai, *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1910-1922; c) D. G. Hall, *Boronic Acids: Preparation, Applications in Organic Synthesis and Medicine*, Wiley, **2005**, p; d) S. Striegler, *Curr. Org. Chem.* **2003**, *7*, 81-102; e) T. D. James, M. D. Phillips and S. Shinkai, *Boronic Acids in Saccharide Recognition*, RSC Pub., **2006**, p.
- [16] a) S. Jin, Y. Cheng, S. Reid, M. Li and B. Wang, *Med. Res. Rev.* **2010**, *30*, 171-257; b) D. B. Walker, G. Joshi and A. P. Davis, *Cell. Mol. Life Sci.* **2009**, *66*, 3177-3191; c) M. Mazik, *Chem. Soc. Rev.* **2009**, *38*, 935-956; d) S. Penadés, *Host-Guest Chemistry: Mimetic approaches to study carbohydrate recognition*,

Springer, **2003**, p; e) A. P. Davis and R. S. Wareham, *Angew. Chem. Int. Ed.* **1999**, *38*, 2978-2996; f) J. Lippe and M. Mazik, *J. Org. Chem.* **2015**, *80*, 1427-1439; g) N. Koch, J.-R. Rosien and M. Mazik, *Tetrahedron* **2014**, *70*, 8758-8767; h) J. Lippe and M. Mazik, *J. Org. Chem.* **2013**, *78*, 9013-9020; i) C. Geffert, M. Kuschel and M. Mazik, *J. Org. Chem.* **2012**, *78*, 292-300; j) J.-R. Rosien, W. Seichter and M. Mazik, *Org. Biomol. Chem.* **2013**, *11*, 6569-6579; k) M. Mazik and C. Geffert, *Org. Biomol. Chem.* **2011**, *9*, 2319-2326; l) M. Mazik and C. Sonnenberg, *J. Org. Chem.* **2010**, *75*, 6416-6423; m) M. Mazik, A. C. Buthe and P. G. Jones, *Tetrahedron* **2010**, *66*, 385-389; n) M. Mazik, A. Hartmann and P. G. Jones, *Chem.: Eur. J.* **2009**, *15*, 9147-9159; o) M. Mazik and A. Hartmann, *J. Org. Chem.* **2008**, *73*, 7444-7450; p) M. Mazik and M. Kuschel, *Chem.: Eur. J.* **2008**, *14*, 2405-2419; q) M. Mazik and H. Cavga, *J. Org. Chem.* **2007**, *72*, 831-838; r) M. Mazik and A. C. Buthe, *J. Org. Chem.* **2007**, *72*, 8319-8326; s) M. Mazik, M. Kuschel and W. Sicking, *Org. Lett.* **2006**, *8*, 855-858; t) M. Mazik and H. Cavga, *J. Org. Chem.* **2006**, *71*, 2957-2963; u) M. Mazik, H. Cavga and P. G. Jones, *J. Am. Chem. Soc.* **2005**, *127*, 9045-9052; v) M. Mazik, W. Radunz and R. Boese, *J. Org. Chem.* **2004**, *69*, 7448-7462; w) M. Mazik and W. Sicking, *Tetrahedron Lett.* **2004**, *45*, 3117-3121; x) M. Mazik, W. Radunz and W. Sicking, *Org. Lett.* **2002**, *4*, 4579-4582; y) M. Mazik and W. Sicking, *Chem.: Eur. J.* **2001**, *7*, 664-670; z) M. Mazik, H. Bandmann and W. Sicking, *Angew. Chem. Int. Ed.* **2000**, *39*, 551-554; aa) A. Ardá, C. Venturi, C. Nativi, O. Francesconi, G. Gabrielli, F. J. Cañada, J. Jiménez-Barbero and S. Roelens, *Chem.: Eur. J.* **2010**, *16*, 414-418; ab) A. Ardá, C. Venturi, C. Nativi, O. Francesconi, F. J. Cañada, J. Jiménez-Barbero and S. Roelens, *Eur. J. Org. Chem.* **2010**, *2010*, 64-71; ac) A. Ardá, F. J. Cañada, C. Nativi, O. Francesconi, G. Gabrielli, A. Ienco, J. Jiménez-Barbero and S. Roelens, *Chem.: Eur. J.* **2011**, *17*, 4821-4829; ad) M. Cacciarini, E. Cordiano, C. Nativi and S. Roelens, *J. Org. Chem.* **2007**, *72*, 3933-3936; ae) M. Cacciarini, C. Nativi, M. Norcini, S. Staderini, O. Francesconi and S. Roelens, *Org. Biomol. Chem.* **2011**, *9*, 1085-1091; af) O. Francesconi, M. Gentili and S. Roelens, *J. Org. Chem.* **2012**, *77*, 7548-7554; ag) O. Francesconi, A. Ienco, G. Moneti, C. Nativi and S. Roelens, *Angew. Chem. Int. Ed.* **2006**, *45*, 6693-6696; ah) O. Francesconi, C. Nativi, G. Gabrielli, M. Gentili, M. Palchetti, B. Bonora and S. Roelens, *Chem.: Eur. J.* **2013**, *19*, 11742-11752; ai) O. Francesconi, M. Gentili, C. Nativi, A. Arda, F. J. Canada, J. Jimenez-Barbero and S. Roelens, *Chemistry* **2014**, *20*, 6081-6091; aj) O. Francesconi, C. Nativi, G. Gabrielli, I. De Simone, S. Noppen, J. Balzarini, S. Liekens and S. Roelens, *Chem.: Eur. J.* **2015**, *21*, 10089-10093; ak) C. Nativi, M. Cacciarini, O. Francesconi, G. Moneti and S. Roelens, *Org. Lett.* **2007**, *9*, 4685-4688; al) C. Nativi, M. Cacciarini, O. Francesconi, A. Vacca, G. Moneti, A. Ienco and S. Roelens, *J. Am. Chem. Soc.* **2007**, *129*, 4377-4385; am) C. Nativi, O. Francesconi, G. Gabrielli, A. Vacca and S. Roelens, *Chem.: Eur. J.* **2011**, *17*, 4814-4820; an) C. Nativi, O. Francesconi, G. Gabrielli, I. De Simone, B. Turchetti, T. Mello, L. D. C. Mannelli, C. Ghelardini, P. Buzzini and S. Roelens, *Chem.: Eur. J.* **2012**, *18*, 5064-5072; ao) A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli and S. Roelens, *J. Am. Chem. Soc.* **2004**, *126*, 16456-16465; ap) Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G. J. Boons, S. Ingale and A. P. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 1775-1779; aq) N. P. Barwell and A. P. Davis, *J. Org. Chem.* **2011**, *76*, 6548-6557; ar) A. P. Davis and R. S. Wareham, *Angew. Chem. Int. Ed.* **1998**, *37*, 2270-2273; as) H. Destecroix, C. M. Renney, T. J. Mooibroek, T. S. Carter, P. F. Stewart, M. P. Crump and A. P. Davis, *Angew. Chem. Int. Ed.* **2015**, *54*, 2057-2061; at) Y. Ferrand, M. P. Crump and A. P. Davis, *Science* **2007**, *318*, 619-622; au) B. Sookcharoenpinyo, E. Klein, Y. Ferrand, D. B. Walker, P. R. Brotherhood, C. Ke, M. P. Crump and A. P. Davis, *Angew. Chem. Int. Ed.* **2012**, *51*, 4586-4590; av) E. Klein, M. P. Crump and A. P. Davis, *Angew. Chem. Int. Ed.* **2005**, *44*, 298-302; aw) E. Klein, Y. Ferrand, E. K. Auty and A. P. Davis, *Chem. Commun.* **2007**, 2390-2392; ax) T. J. Ryan, G. Lecollinet, T. Velasco and A. P. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4863-4866; ay) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump and A. P. Davis, *Nat. Chem.* **2016**, *8*, 69; az) T. Velasco, G. Lecollinet, T. Ryan and A. P. Davis, *Org. Biomol. Chem.* **2004**, *2*, 645-647.

[17] O. Francesconi, M. Martinucci, L. Badii, C. Nativi and S. Roelens, *Chem.: Eur. J.* **2018**, *24*, 6828-6836.

[18] M. Mazik and A. Hartmann, *Beilstein J. Org. Chem.* **2010**, *6*, 9.

- [19] a) S. Rieth, M. R. Miner, C. M. Chang, B. Hurlocker and A. B. Braunschweig, *Chem. Sci.* **2013**, *4*, 357-367; b) K. Palanichamy, M. F. Bravo, M. A. Shlain, F. Schiro, Y. Naeem, M. Marianski and A. B. Braunschweig, *Chem.: Eur. J.* **2018**, *24*, 13971-13982.
- [20] K. Palanichamy, A. Joshi, T. Mehmetoglu-Gurbuz, F. Bravo, M. Shlain, F. Schiro, Y. Naeem, H. Garg and A. B. Braunschweig, *J. Med. Chem.* **2019**.
- [21] T. Ackermann, *Berichte der Bunsengesellschaft für physikalische Chemie* **1987**, *91*, 1398-1398.
- [22] a) M. Mazik, H. Bandmann and W. Sicking, *Angew. Chem.* **2000**, *112*, 562-565; b) M. Mazik and A. König, *J. Org. Chem.* **2006**, *71*, 7854-7857; c) M. Mazik and H. Cavga, *Eur. J. Org. Chem.* **2007**, *2007*, 3633-3638; d) M. Mazik and A. C. Buthe, *Org. Biomol. Chem.* **2008**, *6*, 1558-1568; e) M. Mazik and M. Kuschel, *Eur. J. Org. Chem.* **2008**, *2008*, 1517-1526; f) M. Mazik and A. C. Buthe, *Org. Biomol. Chem.* **2009**, *7*, 2063-2071.
- [23] a) A. P. Davis and R. S. Wareham, *Angew. Chem.* **1998**, *110*, 2397-2401; b) K. M. Bhattarai, A. P. Davis, J. J. Perry, C. J. Walter, S. Menzer and D. J. Williams, *J. Org. Chem.* **1997**, *62*, 8463-8473; c) N. Koch, J.-R. Rosien and M. Mazik, *Tetrahedron* **2014**, *70*, 8758-8767; d) M. Mazik and A. C. Buthe, *Org. Biomol. Chem.* **2009**, *7*, 2063-2071.
- [24] a) G. Ercolani, *J. Am. Chem. Soc.* **2003**, *125*, 16097-16103; b) M. Takeuchi, M. Ikeda, A. Sugasaki and S. Shinkai, *Acc. Chem. Res.* **2001**, *34*, 865-873.
- [25] V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo and J. Seo, *J. Am. Chem. Soc.* **2018**, *140*, 7554-7560.
- [26] a) F. Schubert, M. Rossi, C. Baldauf, K. Pagel, S. Warnke, G. von Helden, F. Filsinger, P. Kupser, G. Meijer, M. Salwiczek, B. Koksche, M. Scheffler and V. Blum, *PCCP* **2015**, *17*, 7373-7385; b) W. Hoffmann, M. Marianski, S. Warnke, J. Seo, C. Baldauf, G. von Helden and K. Pagel, *PCCP* **2016**, *18*, 19950-19954; c) M. Schneider, C. Masellis, T. Rizzo and C. Baldauf, *J. Phys. Chem. A* **2017**, *121*, 6838-6844.
- [27] D. van der Spoel and M. M. Seibert, *Phys. Rev. Lett.* **2006**, *96*, 238102.
- [28] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX* **2015**, *1-2*, 19-25.
- [29] O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison and A. D. Mackerell, *J. Chem. Theory Comput.* **2011**, *7*, 3162-3180.
- [30] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. Mackerell, J. B. Klauda and W. Im, *J. Chem. Theory Comput.* **2016**, *12*, 405-413.
- [31] G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
- [32] A. Patriksson and D. van der Spoel, *PCCP* **2008**, *10*, 2073-2077.
- [33] a) J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865-3868; b) A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.* **2009**, *102*, 073005.
- [34] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, *Comput. Phys. Commun.* **2009**, *180*, 2175-2196.
- [35] a) A. Tkatchenko, M. Rossi, V. Blum, J. Ireta and M. Scheffler, *Phys. Rev. Lett.* **2011**, *106*, 118102; b) N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler and L. Kronik, *J. Chem. Theory Comput.* **2011**, *7*, 3944-3951; c) M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, *J. Chem. Theory Comput.* **2016**, *12*, 6157-6168.
- [36] A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
- [37] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J.

Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox in *Gaussian 16 Rev. D.01, Vol.* Wallingford, CT, **2016**.

[38] A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 4538-4543.

[39] J. Zhang, H. Zhang, T. Wu, Q. Wang and D. van der Spoel, *J. Chem. Theory Comput.* **2017**, *13*, 1034-1043.

[40] S. Schrödinger Release 2019-4: MacroModel, LLC, New York, NY, 2019.