





Chemical synthesis of glycans up to a 128-mer relevant to the O-antigen of *Bacteroides vulgatus*

Qian Zhu ^{1,7}, Zhengnan Shen^{2,7}, Fabrizio Chiodo ³, Simone Nicolardi⁴, Antonio Molinaro⁵, Alba Silipo ⁵✉ & Biao Yu ^{1,6}✉

Glycans are involved in various life processes and represent critical targets of biomedical developments. Nevertheless, the accessibility to long glycans with precise structures remains challenging. Here we report on the synthesis of glycans consisting of [→4)-α-Rha-(1 → 3)-β-Man-(1 →] repeating unit, which are relevant to the O-antigen of *Bacteroides vulgatus*, a common component of gut microbiota. The optimal combination of assembly strategy, protecting group arrangement, and glycosylation reaction has enabled us to synthesize up to a 128-mer glycan. The synthetic glycans are accurately characterized by advanced NMR and MS approaches, the 3D structures are defined, and their potent binding activity with human DC-SIGN, a receptor associated with the gut lymphoid tissue, is disclosed.

¹State Key Laboratory of Bioorganic and Natural Products Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China. ²School of Physical Science and Technology, ShanghaiTech University, 393 Huaxia Middle Road, Shanghai 201210, China. ³Department of Molecular Cell Biology and Immunology, Amsterdam Infection and Immunity Institute, De Boelelaan 1108, 1081HZ Amsterdam, The Netherlands. ⁴Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden 2333 ZA, The Netherlands. ⁵Department of Chemical Sciences, University of Naples Federico II, Via Cintia 4, 80126 Napoli, Italy. ⁶School of Chemistry and Materials Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, 1 Sub-Lane Xiangshan, Hangzhou 310024, China. ⁷These authors contributed equally: Qian Zhu, Zhengnan Shen. ✉email: silipo@unina.it; byu@sioc.ac.cn

Glycans represent the most abundant biopolymers in nature, which constitute prominent parts of all living organisms and mediate fundamental biological processes. Short glycans (oligosaccharides) with less than 20 monosaccharide units, such as those being conjugated with lipids and proteins on the cell surfaces, are usually sufficient to fulfill various biological functions, including cell signaling, adhesion, and migration¹. Longer glycans, such as glycosaminoglycans, display both structural properties and various physiological and pathological functions². Polysaccharides, such as cellulose and starch/glycogen, serve for structural requirements or energy storage. The microbial glycans are structurally and functionally extremely diverse. For example, those on the microbe cell envelope mediate host–microbial cross-talk and play a critical role in the activation of the immune system. The development of robust strategies for the synthesis of well-defined carbohydrate structures is nowadays a top priority and has an enormous impact in various biomedical sectors; in this context, synthetic carbohydrate–vaccines, as well as glycoconjugates-based adjuvants, have been produced³. Paradoxically, functional studies on long glycans with precise structures have been rarely performed⁴, due to the extremely difficult acquisition of pure long glycans via either isolation from natural sources or via chemical synthesis from mono- and oligosaccharide units. Indeed, natural glycans occur as heterogeneous mixtures, or despite showing homogeneous repeating units exhibit a range of size distributions. The chemical synthesis of long glycans has been limited by the notorious low efficiency of the glycosidic coupling and protecting group manipulations. Consequently, glycans containing over 20 monosaccharide units have been synthesized only occasionally^{5–18}, contributing to the general idea that chemical synthesis of long glycans are not yet of practical usefulness¹⁹. In this regard, a paradigm shift in the functional studies of long glycans with precise structures requires the advancement of feasible approaches to their preparation, bringing innovation to the conventional synthetic carbohydrate chemistry^{20,21}.

Bacteroides vulgatus mpk is a commensal strain occurring commonly in the gastrointestinal tract of American and Western European population^{22,23}. Recent studies demonstrated that this bacterium exerted strong immune-modulating properties leading to the prevention of colitis-induction in mouse models, and the *Bacteroides* lipopolysaccharide (LPS) induced hyporesponsiveness towards subsequent LPS-stimuli. Also, the administration of *Bacteroides* LPS re-established intestinal immune homeostasis in a mouse model for experimental colitis, thus correlating the health-promoting effects to the weak agonistic properties of this LPS^{24,25}. Further, in the human innate immune system, we demonstrated a relevant capability of the LPS to induce anti-inflammatory cytokines²⁶. The *O*-antigen of *B. vulgatus* mpk LPS consists of a disaccharide repeating unit of [\rightarrow 4)- α -Rha-(1 \rightarrow 3)- β -Man-(1 \rightarrow)_n²⁷ (Fig. 1). The uniqueness of this structure and the relevance of glycan motifs in modulating the immune

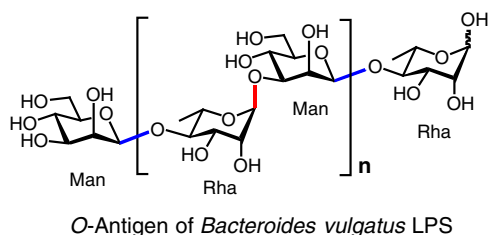


Fig. 1 The *O*-antigen glycans of *Bacteroides vulgatus* LPS. The α -(1 \rightarrow 3) glycosidic bond between L-rhamnopyranose (Rha) and D-mannopyranose (Man) is highlighted in red, and the β -(1 \rightarrow 4) glycosidic bond between Man and Rha in blue. Source data are provided as a Source Data file.

functions prompted us to synthesize homogeneous glycoforms containing *Bacteroides* *O*-antigen repeating units for structural and functional studies.

Herein, we report the syntheses of glycans up to a 128-mer of the *O*-antigen of the LPS of *B. vulgatus* mpk, and the structural and functional studies of the synthetic, pure, and active glycoforms.

Results

Syntheses of the glycans. The target glycans consist of two types of glycosidic linkages, i.e., Rha-(1 α \rightarrow 3)-Man and Man-(1 β \rightarrow 4)-Rha. The β -mannopyranoside linkages are known to be difficult to construct in a highly stereoselective manner, while the α -rhamnopyranoside linkages are among the easiest to synthesize²⁸. This knowledge forced us to fix the requisite β -mannopyranoside linkage at a disaccharide level and to extend the glycan resorting to α -rhamnopyranosylation. Such a disaccharide building block required orthogonal protection at the Rha 1-OH and Man 3-OH, which could be transformed into the corresponding donor and acceptor for a convergent glycan elongation. Disaccharide **8**, bearing MP (*p*-methoxyphenyl) and TBS (*tert*-butyldimethylsilyl) group at the reducing and non-reducing end, respectively, was found to be effective in the present work, in which the 4',6'-O-benzylidene and 2'-*O*-benzyl group were used to facilitate the construction of the β -mannopyranoside linkage and the 2-*O*-benzoyl group to secure the α -rhamnopyranosylation in the subsequent glycan assembly (Fig. 2).

Although the bulky 3'-*O*-TBS group installed at the mannosyl donors was known to be detrimental to the β -selectivity in mannosylation^{28,29}, we managed to prepare the desired disaccharide building block **8** in decagram scales with a convenient and inexpensive procedure (Supplementary Fig. 1 & Supplementary Table 1). Removal of the TBS group with TBAF (THF, rt) led to disaccharide acceptor **8^A** smoothly; at this stage, the β/α isomers resulted from the mannosylation step could be easily separated. The desired *o*-hexynylbenzoate donor **8^D** was prepared from **8** via two steps, i.e., oxidative removal of the anomeric MP group with CAN (CH₂Cl₂, MeCN, H₂O, 0 °C) and condensation with *o*-hexynylbenzoic acid (EDCI, DMAP, CH₂Cl₂, rt). As expected, the glycosidic coupling between disaccharide donor **8^D** and acceptor **8^A** proceeded smoothly under the standard gold(I)-catalyzed conditions (0.1 equiv. Ph₃PAuNTf₂, 5 Å MS, CH₂Cl₂, 0 °C)³⁰, leading to tetrasaccharide **9** in 92% yield at a decagram scale, with the α -rhamnopyranoside linkage being exclusively formed.

The convenient synthesis of 4-mer **9** from 2-mer **8** established an effective approach toward assembly of the 2^{*n*+1}-mer glycans via repetition of a cycle of three transformations (Fig. 2), including (1) selective cleavage of the anomeric MP group (with CAN) and condensation of the resulting hemiacetal with *o*-hexynylbenzoic acid to provide the donor (steps **a** and **b**); (2) selective removal of the non-reducing end TBS ether (with TBAF) to provide the acceptor (step **c**); and (3) coupling of the donor and acceptor (under the catalysis of Ph₃PAuNTf₂) to furnish the 2^{*n*+1}-mer glycan (step **d**). These transformations were found robust, so that the syntheses of 8-mer **10**, 16-mer **11**, and 32-mer **12** met with no incident in gram-scales.

Further transformations starting from 32-mer **12** (molecular weight (Mw) = 11,354) were challenged by the decreased solubility of the macromolecular substrates. Thus, removal of the TBS ether on **12** was carried out at a low concentration of 14.0 mM in THF; hydrolysis of the anomeric MP group at 2.2 mM in CH₂Cl₂/MeCN/H₂O and condensation with *o*-hexynylbenzoic acid at 2.1 mM in CH₂Cl₂. These transformations were clean so that the desired 32-mer acceptor **12^A** (88%) and donor

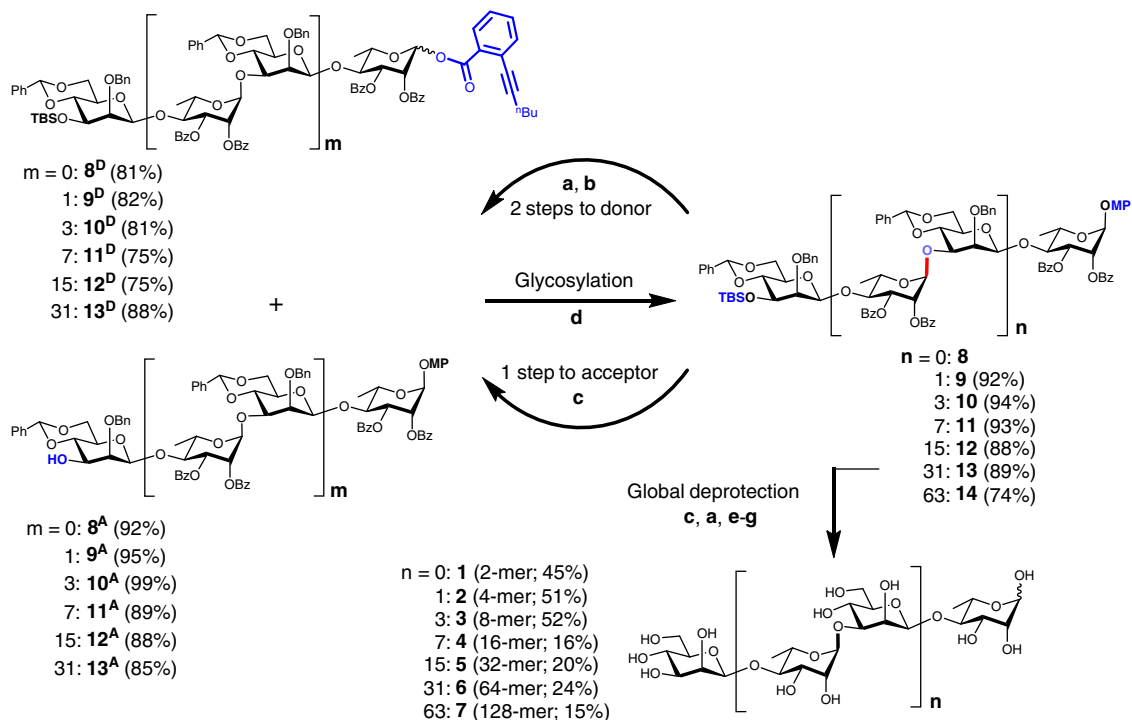


Fig. 2 Iterative assembly to the preparation of glycans up to 128-mer (**7**). **a** CAN, CH₂Cl₂, MeCN, H₂O, 0 °C; **b** *o*-hexynylbenzoic acid, EDCl, DMAP, CH₂Cl₂, rt; **c** TBAF, HOAc, THF, rt; **d** Ph₃PAuNTf₂ (0.1 or 0.2 equiv.), 5 Å MS, CH₂Cl₂, 0 °C; **e** BzCl, Et₃N, CH₂Cl₂, rt; **f** H₂, 10% Pd/C (and 20% Pd(OH)₂), solvent, rt; **g** NaOCH₃, CH₃OH, rt. The experimental procedures varied slightly for substrates of different sizes (Supplementary Tables 2–5), and the hydrogenolysis (step **f**) was repeated a couple of times after step **g** until the benzyl groups were fully removed. CAN cerium (IV) diammonium nitrate, DMAP 4,4-dimethylaminopyridine, EDCl 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride, MS molecular sieves, rt room temperature, TBAF tetrabutylammonium fluoride, THF tetrahydrofuran. The newly synthesized glycosidic bond is highlighted in red, and reacting functional groups in blue. Source data are provided as a Source Data file.

12^D (75%) were obtained with high purity after flash column chromatography on silica gel. The coupling of **12^A** and **12^D** was carried out at 7.4 mM in CH₂Cl₂ (0.2 equiv. Ph₃PAuNTf₂, 5 Å MS, 0 °C), leading to the coupled 64-mer **13** in 89% yield after purification on gel permeation chromatography (GPC).

The ready availability of 64-mer **13** (at a hundred mg scale) encouraged us to carry on the next synthetic cycle to prepare 128-mer **14**. The removal of the TBS group on **13** could only be carried out at a concentration of 1.7 mM in THF (10 equiv. TBAF, 4 equiv. HOAc, rt); hydrolysis of the anomeric MP group at 0.9 mM in CH₂Cl₂/MeCN/H₂O (36 equiv. CAN, 0 °C) and subsequent condensation with *o*-hexynylbenzoic acid (7 equiv.) at 1.4 mM in CH₂Cl₂ (13 equiv. EDCl, 13 equiv. DMAP, rt). Upon flash chromatography on short silica gel columns to remove the excess reagents, 64-mer acceptor **13^A** and donor **13^D** were obtained with satisfactory purity (~88% yields), as indicated by nuclear magnetic resonance (NMR) analysis. The coupling of 64-mers **13^A** (80 mg, 1 equiv.) and **13^D** (1.4 equiv.) was then conducted at a concentration of 3.6 mM in CH₂Cl₂ under otherwise similar conditions as used in the previous glycosylation reactions, and furnished 128-mer **14** in a decent 74% yield after GPC purification.

The remaining synthetic challenge was the global deprotection of the fully protected long glycans. For 128-mer **14**, cleavage of 64 benzylidene, 64 benzyl, and 128 benzoyl groups, besides the terminal TBS and MP group, are required, involving scission of 322 C–O bonds. Fortunately, a procedure optimized on di- and tetrasaccharides **8** and **9** was successfully adapted to the deprotection of the longer congeners, that involved five sequential steps: (1) removal of the TBS ether with TBAF; (2) cleavage of the anomeric MP with CAN; (3) protection of the resulting

hemiacetal with benzoyl group; (4) removal of the benzyl and benzylidene groups via hydrogenolysis; and (5) removal of the benzoyl groups with NaOCH₃ in MeOH. The seemingly redundant step (3) for capping the resultant hemiacetal was found crucial to the subsequent hydrogenolysis (especially for the long glycan substrates), since it enabled complete removal of the detrimental impurities from the proceeding steps and avoided degradation of the aldehyde. Nevertheless, hydrogenolysis of the longer glycans starting from 16-mer **11** was found difficult to complete, because the gradually releasing hydroxyl groups could cause an unwanted decrease of the solubility in the reaction solvents. Thus, repetitive hydrogenolysis was performed after the removal of the benzoates to cleave the remaining benzyl groups in a mixture solvent of MeOH/H₂O/HOAc. Moreover, a combination of 10% Pd/C and 10% Pd(OH)₂ was used as the catalyst³¹. The complete and clean reactions enabled acquiring the free glycans with satisfactory purity via simple filtration through Sephadex G-25 (for 1–5) or LH-60 (for **6** and **7**).

Structural characterization. The fully protected and free glycans were well characterized by a combination of NMR and MS analyses. As for the free glycans 1–7, the ¹H-NMR spectra (Fig. 3a) showed a characteristic set of signals of the disaccharide repeating unit and were fully assigned (Supplementary Figs. 3–7; Supplementary Table 6). The NMR resonances of the internal repeating units were coincident and clearly distinguishable from the terminal reducing and nonreducing units. The molecular size of the glycans was estimated by the integration of the ¹H-NMR signals (ratio of internal and terminal units) and by the retention time on GPC (Supplementary Fig. 2); as for the longer size

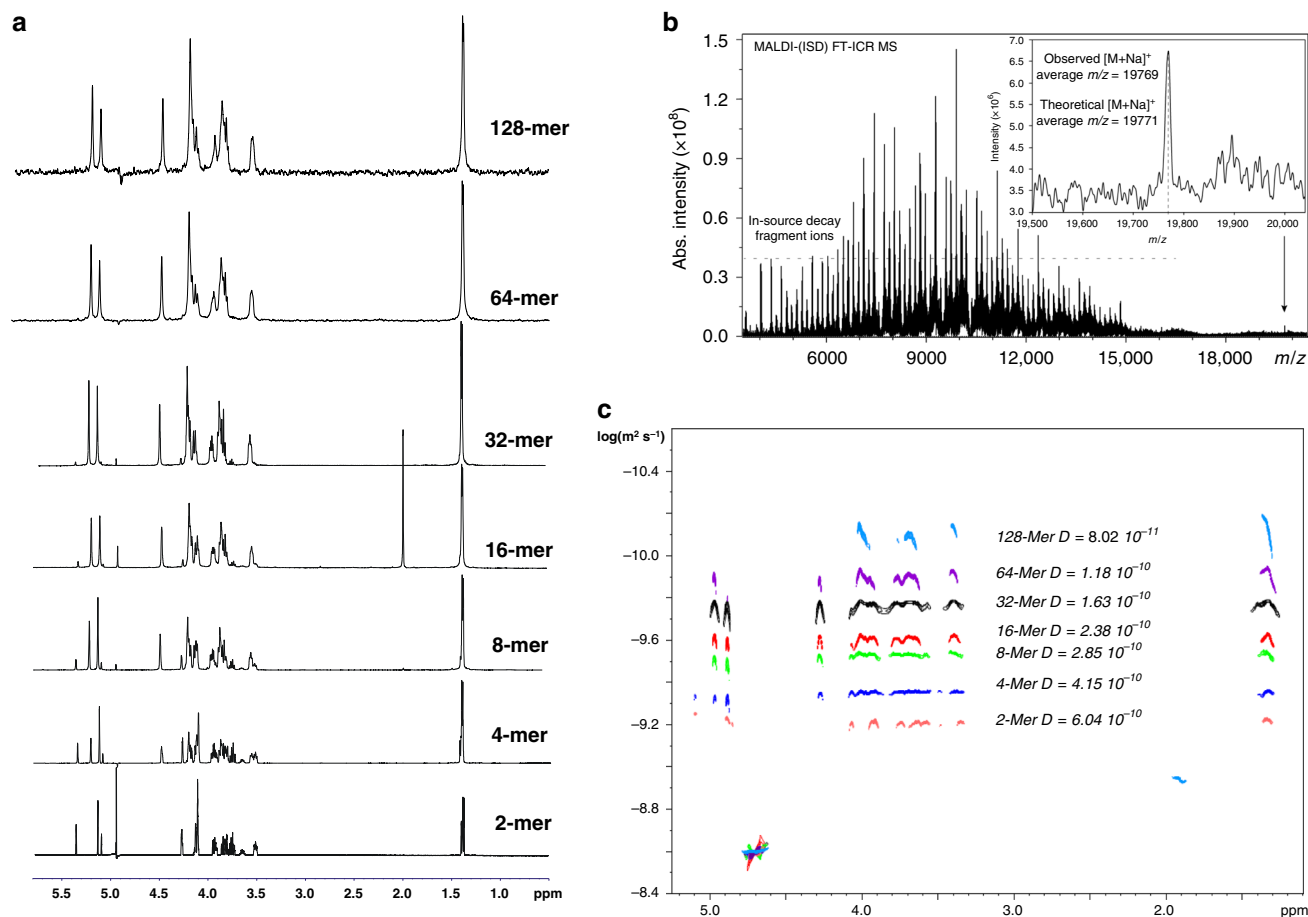


Fig. 3 Characterization of the glycans. | **a** Overlaid ^1H -NMR spectra (acquired on Bruker 600 MHz with cryogenic probe and analyzed with Bruker Topspin) of the free glycans 1–7. **b** MALDI-FT-ICR-MS spectrum of 128-mer **7**. The observed average m/z of 128-mer **7**, detected as $[\text{M} + \text{Na}]^+$, is in agreement with the calculated theoretical value (m/z 19771.1). **c** 2D diffusion ordered spectroscopy (DOSY) NMR spectra of glycans 1–7. The x-axis represents the ^1H -NMR dimension; the y-axis represents the diffusion dimension; the double-logarithmic plot of D against Mw (Supplementary Fig. 13) provided a calibration curve described by the least-squares fitted linear equation $\text{Log } D = -7873 - (0.486 \cdot \text{Log Mw})$. Source data are provided as a Source Data file.

glycans, a combination of DOSY NMR and advanced MALDI-MS analyses was applied. In particular, DOSY-NMR measurements, capable of determining the molecular self-diffusion coefficients^{32,33}, were successfully applied to estimate the molecular weights of the longest glycans (Fig. 3c; Supplementary Fig. 13). In particular, DOSY-derived Mw for 64-mer **6** was well consistent with the theoretical value; the average molecular mass of 128-mer **7**, complicated by its size and lability during ionization, was instead determined by MALDI-FT-ICR-MS (Fig. 3b; Supplementary Fig. 14)³⁴.

Conformation and immunological activities. Together with the complete NMR assignments (Supplementary Table 6; Supplementary Figs. 3–7), molecular mechanics and dynamic simulation were carried out to assess the conformational behavior of the free glycans (Supplementary Figs. 8–12; Supplementary Table 7). The glycosidic linkages adopted Φ values in accordance with the *exo*-syn anomeric conformation and were characterized by a certain degree of flexibility. The oligomers tend to form extended, flexible structures, as confirmed by the *inter*-residue NOE contacts, diagnostic of the absence of a compact disposition of the sugar backbone (Fig. 4a). On longer glycans (i.e., 64- and 128-mer), this may lead to a tendency to pack into supramolecular assemblies, this latter also accounting for the low solubility of the longer

glycans and the low sensitivity of MALDI technique for the high Mw 128-mer.

The synthetic approach allowed us to deploy a full set of pure synthetic different sized *O*-antigen fragments, which were tested for immunological activity. Given the presence of rhamnose and mannose in *B. vulgatus* *O*-antigen and considering the niche and the physiological roles of these commensal bacteria, we used three immune receptors of the gut-associated lymphoid tissues in the study. Thus, *B. vulgatus* LPS and their fragments were examined on a direct ELISA with three human C-type lectins (i.e., DC-SIGN, Langerin, and Dectin-1)^{35,36}, showing a qualitative binding to DC-SIGN and Langerin (Supplementary Fig. 15a). Subsequently, the binding of *B. vulgatus* LPS to DC-SIGN and Langerin was tested in a quantitative ELISA competition experiment; a better binding to DC-SIGN was observed (Supplementary Fig. 15b). Following up, the synthetic glycans were tested on a competition ELISA experiment against DC-SIGN³⁷. Expectedly, glycans 1–7 were shown to be able to inhibit the LPS binding to DC-SIGN in a typical dose-response curve (Fig. 4b). A better affinity for the longer glycans (**6** and **7**) was observed compared with the shorter ones. Hence, we additionally proved a selective binding of the synthetic *O*-antigens from *B. vulgatus* to DC-SIGN, adding unreported information that could explain the tolerogenic signaling induced by the commensal *B. vulgatus* in gut-associated lymphoid tissues.

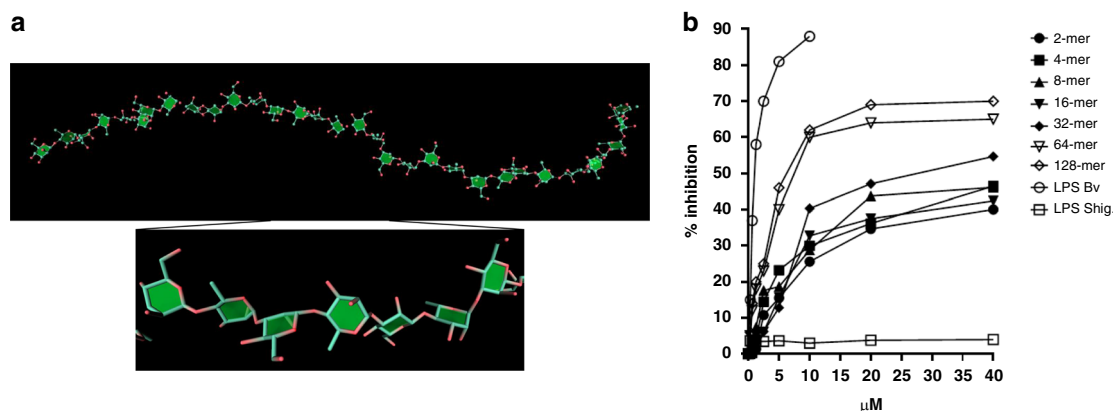


Fig. 4 Conformation and immunological activities. **a** A representative solution conformation of the 32-mer as determined by NMR and molecular simulation (Maestro Schrödinger), prepared with SweetUnitMol³⁹; **b** ELISA analysis of the competition binding of the glycans and LPS from *B. vulgatus* to the human C-type lectin DC-SIGN, with LPS from *Shigella flexneri* as a negative control. The competition experiments have been performed three times showing similar results; the graph shows data of one of these experiments⁴⁰. Source data are provided as a Source Data file.

Discussion

The synthesis of glycans up to a 128-mer (7) consisting of disaccharide repeating unit of $[\rightarrow 4)\text{-}\alpha\text{-Rha-(1}\rightarrow 3)\text{-}\beta\text{-Man-(1}\rightarrow]_n$ have been achieved. The present success is attributable to the following aspects: (1) The strategic $[2^n + 2^n]$ glycosylation enables exponential growth of the glycan sizes¹⁷, leading to 128-mer via only six repetitive cycles of transformations, and ready separation of the double-sized products from the remaining and decomposed substrates. The heroic syntheses of a linear 50-mer¹³ and a 100-mer¹⁸ via step-wise glycosylation demanded exhaustive reactions (glycosylation and capping on a solid support) at each step; while the synthesis of the 151-mer¹⁸ and 92-mer¹⁴ took advantage of multifold glycosylations to install the identical branches. (2) The overall protecting group arrangement enables robust transformations into donors and acceptors, provides appropriate solubility, and allows global deprotection at the macromolecular level. It is noteworthy that the 15% yield achieved for the deprotection of 128-mer ($14\rightarrow 7$) corresponds to an average $\sim 99.4\%$ yield for each cleavage of the 322 C–O bonds. Thus far, hydrogenolysis of benzyl ethers has been employed as the final step in all the synthesis of long glycans over 50-mer, however, cautiousness remains for this transformation which could require tedious exploration of the reaction conditions³⁸. (3) The gold(I)-catalyzed glycosylation reaction, involving α -rhamnopyranosylation, is both stereospecific (in the presence of a neighboring-participating benzoyl group) and high-yielding. The 64-mer donor (**13^D**) represents the longest glycosyl donor that has ever been employed for successful glycosylation; this might be attributed to the merits of glycosyl *ortho*-alkynylbenzoates, including easy preparation, stability, and selective activation under mild conditions³⁰.

Besides the synthetic achievement, the availability of pure homogeneous glycans has allowed us to test the limits of analytical tools for the characterization of polysaccharides. Thus, the accurate confirmation of the molecular size of 64-mer (6) represents one of the longest glycan determined so far by DOSY-NMR, and the average molecular mass of 128-mer (7) represents the highest mass of a synthetic linear polysaccharide being measured by MS. Moreover, the synthetic *O*-antigen showed a selective affinity with human lectin DC-SIGN and this might be relevant information for therapeutic use against different gut-associated inflammatory diseases. We expect that the present report can inspire researches on the syntheses and thus functional studies on pure and homogenous polysaccharides for different therapeutic uses.

Methods

General methods for the syntheses. Reactions were carried out in glassware. Crushed 4 or 5 Å molecular sieves were activated through flame-drying under high vacuum immediately prior to use. All chemicals were purchased as reagent grade and used without further purification, unless otherwise noted. Analytical thin-layer chromatography was performed using Merck precoated silica gel 60 F-254 plates. Compound spots were visualized by UV light (254 nm) and immersion into a solution of 5% H_2SO_4 in ethanol, followed by hot air gun heating. Column chromatography was performed on silica gel (200–300 mesh). Gel filtration was performed on Sephadex LH-60, LH-20, or G25 (J&K). Optical rotations were obtained on Anton Paar MCP 5500 polarimeter at 589 nm (Na).

Synthesis of 64-mer donor **13^D.** A solution of 64-mer **13** (320 mg, 14.2 μmol) in $\text{CH}_2\text{Cl}_2/\text{MeCN}$ (10 mL/4 mL) was stirred at 0 °C. Ammonium ceric nitrate (281 mg, 0.51 mmol) was dissolved in H_2O (1 mL) at 0 °C. The latter solution was added into the former via syringe, and the mixture was stirred at 0 °C for 2 h. The mixture was diluted with CH_2Cl_2 and washed with sat. aq. NaHCO_3 and brine, respectively. The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by silica gel column chromatography (petroleum ether/ $\text{EtOAc}/\text{CH}_2\text{Cl}_2$, 4:1.5:3) to give the corresponding crude hemiacetal (301 mg) as a yellowish solid.

To a mixture of the crude hemiacetal (301 mg, $\sim 13.5\ \mu\text{mol}$), EDCI (32.1 mg, 0.17 mmol), DMAP (20.6 mg, 0.170 mmol), and *ortho*-hexynylbenzoic acid (20.4 mg, 0.10 mmol), was added CH_2Cl_2 (10 mL). The mixture was stirred at rt for 10 h, and was then diluted with CH_2Cl_2 and washed with sat. aq. NaHCO_3 . The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by silica gel column chromatography (petroleum ether/ $\text{EtOAc}/\text{CH}_2\text{Cl}_2$, 12:1.5:3 then 4:1.5:3) to give compound **13^D** (283 mg, 88% for two steps, $\alpha/\beta = 2.3:1$) as a white solid.

Synthesis of 64-mer acceptor **13^A.** To a solution of 64-mer **13** (190 mg, 8.5 μmol) in anhydrous THF (5 mL), were added HOAc (2 μL , 34.9 μmol) and TBAF (1.0 M in THF, 86.0 μL , 86.0 μmol). The mixture was stirred at rt for 48 h, and was then diluted with CH_2Cl_2 and washed with sat. aq. NH_4Cl and brine, respectively. The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by silica gel column chromatography (petroleum ether/ $\text{EtOAc}/\text{CH}_2\text{Cl}_2$, 4:1.5:3) to give **13^A** (179 mg, $\sim 85\%$; $>90\%$ purity based on $^1\text{H-NMR}$ analysis) as a white solid.

Synthesis of protected 128-mer **14.** A mixture of 64-mer donor **13^D** (113 mg, 5.00 μmol) and acceptor **13^A** (80 mg, 3.58 μmol) in a Schlenk flask equipped with a Teflon-coated magnetic stir bar was dried in high vacuum at 40 °C overnight. And then 5 Å MS (100 mg) and anhydrous CH_2Cl_2 (1 mL) were added. The mixture was stirred at rt for 15 min, and was then cooled to 0 °C, to which $\text{Ph}_3\text{PAuNTf}_2$ (0.53 mg, 0.72 μmol) was added. The resulting mixture was stirred at 0 °C for 15 h. After addition of Ph_3P (0.5 mg) and Et_3N (4 μL), the mixture was filtrated through a pad of Celite. After removal of solvent, the resulting residue was dissolved in CHCl_3 and filtrated through Millipore filter, and then GPC (CHCl_3) was performed to give 128-mer **14** (114 mg, 74%) as a white solid.

Synthesis of 128-mer **7.** To a solution of 128-mer **14** (45 mg, 1.0 μmol) in anhydrous THF (5 mL), were added HOAc (2 μL , 34.9 μmol) and TBAF (1.0 M in THF, 86.2 μL , 86.2 μmol). The mixture was stirred at rt for 46 h, and was then

diluted with CH_2Cl_2 and washed with sat. aq. NH_4Cl and brine, respectively. The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo to give a residue.

A solution of the residue above (45 mg, 1.0 μmol) in $\text{MeCN}/\text{CH}_2\text{Cl}_2$ (8 mL/5 mL) was stirred at 0 °C. Ammonium ceric nitrate (242 mg, 0.44 mmol) was dissolved in H_2O (0.5 mL) at 0 °C. The latter solution was added into the former via syringe. The mixture was stirred at 0 °C for 1.5 h, and was then diluted with CH_2Cl_2 and washed with sat. aq. NaHCO_3 and brine, respectively. The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was used in the next step without further purification.

To a mixture of the residue above (45 mg), DMAP (20 mg, 0.16 mmol), and Et_3N (30 μL , 0.21 mmol) in CH_2Cl_2 (2 mL), was added BzCl (12 μL , 104 μmol). The mixture was stirred at rt overnight, and was then diluted with CH_2Cl_2 and washed with sat. aq. NaHCO_3 and brine, respectively. The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by silica gel chromatography ($\text{EtOAc}/\text{CH}_2\text{Cl}_2$, 1:2) to give a white solid (40 mg).

The solid above (40 mg) was dissolved in a mixture of $\text{MeOH}/\text{CHCl}_3/\text{H}_2\text{O}/\text{HOAc}$ (1.5 mL/3 mL/0.1 mL/0.15 mL) containing 10% Pd/C (80 mg, wetted with 55% H_2O) and 20% $\text{Pd}(\text{OH})_2/\text{C}$ (80 mg, wetted with 50% H_2O). The resulting mixture was stirred under H_2 atmosphere (1 atm) at rt for 48 h, and was then filtered through a pad of Celite, and the Celite pad was washed with $\text{CH}_2\text{Cl}_2/\text{MeOH}$ (2:1, v/v) three times. The filtrates were concentrated to give a white solid (30 mg).

The residue above (30 mg) was added into a solution of $\text{NaOCH}_3/\text{CH}_3\text{OH}$ (5 mL, pH = ~11) and CH_2Cl_2 (4 mL). The mixture was stirred at rt for 24 h, and was then neutralized with Amberlyst 15 H^+ resin and filtered. The filtrate was concentrated to give a residue (12 mg). $^1\text{H-NMR}$ analysis indicated remaining of a few benzyl groups.

Thus, the solid above (6 mg) was dissolved in a mixture of $\text{MeOH}/\text{H}_2\text{O}/\text{HOAc}$ (1.5 mL/1.5 mL/0.1 mL) containing 10% Pd/C (12 mg, wetted with 55% H_2O) and 20% $\text{Pd}(\text{OH})_2/\text{C}$ (6 mg, wetted with 50% H_2O). The resulting mixture was stirred under H_2 atmosphere (1 atm) at rt for 24 h, and was then filtered through a pad of Celite, and the Celite pad was washed with H_2O and MeOH three times. The filtrate was concentrated. The residue was purified by gel filtration (Sephadex LH-60, H_2O) to afford 128-mer 7 (3 mg, 15% over six steps) as a glassy solid or white powder after lyophilization.

NMR analysis. For structural assignments, 1D and 2D $^1\text{H-NMR}$ spectra were recorded in D_2O at 308 K on Bruker 600 MHz equipped with a cryo probe. ROESY and NOESY experiments were recorded using data sets ($t_1 \times t_2$) of 4096×800 points with mixing times between 100 and 700 ms. Double quantum-filtered phase-sensitive COSY experiments were performed using data sets of 4096×800 points. TOCSY experiments were performed with spinlock times of 100 ms, using data sets ($t_1 \times t_2$) of 4096×800 points. In all homonuclear experiments the data matrix was zero-filled in both dimensions to give a matrix of $4\text{K} \times 2\text{K}$ points and was resolution enhanced in both dimensions by a cosine-bell function before Fourier transformation. HSQC, HSQC-ROESY and HMBC experiments were measured in the ^1H -detected mode via single quantum coherence with proton decoupling in the ^{13}C domain, using data sets of 2048×600 points. Experiments were carried out in the phase-sensitive mode. A 60 ms delay was used for the evolution of long-range connectivities in the HMBC experiment. In all heteronuclear experiments the data matrix was extended to 4096×4096 points using forward linear prediction extrapolation.

PFQ-NMR analysis. All the experiments were performed on 100–200 μg of glycans solved in 500 μL of D_2O . 1D $^1\text{H-NMR}$ experiments for diffusion measurements were performed using a stimulated echo sequence with bipolar gradient pulses and one spoil gradient (stebppg1s1d) with a diffusion time, Δ from 50 to 150 ms and a gradient duration varying from 1 to 4 ms. These parameters were optimized, kept as short as possible to minimize T_2 and T_1 losses, the sequence was run as 2D NMR experiment with a linear gradient incremented, in 32 steps, from 2% to 95% (between 1.8 and 32.9 G cm^{-1}). The values of diffusion coefficients were calculated from the decay of the signal intensity of different protons. The normalized integral intensities of each proton were fitted to an exponential decay, according to: $I = I_0 \exp - D_k$ with $k = (G^* \gamma \delta)^2 (\Delta - \delta/3 - \tau/2)$, where I is the integral intensity, normalized to the integral obtained at the lowest gradient amplitude; I_0 is the signal intensity in the absence of an applied magnetic field gradient; γ is the magnetogyric ratio of the proton; G is the strength of the magnetic field gradient pulses; δ is their duration; Δ is the distance between the leading edges of the gradient pulses; and τ is the gradient recovery time. Diffusion rates were extracted from the slope of the straight lines obtained by plotting $\ln(I/I_0)$ against the complex abscissa k .

Conformational studies. Molecular mechanics calculations were performed using the MM3* force field, a dielectric constant of 80 was used. For the disaccharide structure, both Φ and Ψ were varied incrementally using a grid step of 18°, each (Φ , Ψ) point of the map was optimized using 2000 P.R. conjugate gradients. Molecular dynamic simulations were run by using the MM3* force field, bulk water solvation was simulated by using MacroModel generalized Born GB/SA continuum solvent model. All simulations were performed at 298 K, structures were initially subjected to an equilibration time of 300 ps, then a 10,000 ps molecular dynamic simulation was performed with a dynamic time-step of 1.5 fs, a bath constant of 0.2 ps and the

SHAKE protocol to the hydrogen bonds. Trajectory coordinates were sampled every 2 ps, and a total of 5000 structures were collected for every simulation. Ensemble average-interproton distances were calculated using the NOEPROM program by applying the isolated spin pair approximation as described. Solvent-accessible surfaces were calculated with the Surface utility of Maestro. Conformers were visualized with Maestro, Discovery Studio Visualizer and SweetUnitMol.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary information files, including experimental details, characterization data, and ^1H and ^{13}C NMR spectra (Supplementary Figs. 16–78) of all new compounds and MS spectra of the glycans (Supplementary Figs. 79–91). All the data are available from the authors upon reasonable request. Source data are provided with this paper.

Received: 1 May 2020; Accepted: 29 July 2020;

Published online: 18 August 2020

References

- Bertozzi, C. R. & Kiessling, L. L. Chemical glycobiology. *Science* **291**, 2357–2364 (2001).
- Capila, I. & Linhardt, R. J. Heparin-protein interactions. *Angew. Chem. Int. Ed.* **41**, 390–412 (2002).
- Verez-Bencomo, V. et al. A synthetic conjugate polysaccharide vaccine against Haemophilus influenzae type b. *Science* **305**, 522–525 (2004).
- Delbianco, M. et al. Well-defined oligo- and polysaccharides as ideal probes for structural studies. *J. Am. Chem. Soc.* **140**, 5421–5426 (2018).
- Matsuzaki, Y., Ito, Y., Nakahara, Y. & Ogawa, T. Synthesis of branched poly-N-acetyl-lactosamine type pentaantennary pentacosaccharide: glycan part of a glycosyl ceramide from rabbit erythrocyte membrane. *Tetrahedron Lett.* **34**, 1061–1064 (1993).
- Pozsgay, V. A new strategy in oligosaccharide synthesis using lipophilic protecting groups: synthesis of a tetracosasaccharide. *Tetrahedron* **11**, 151–172 (2000).
- Fraser-Reid, B., Lu, J., Jayaprakash, K. N. & López, J. C. Synthesis of a 28-mer oligosaccharide core of Mycobacterial lipoarabinomannan (LAM) requires only two n-pentenyl orthoester progenitors. *Tetrahedron* **17**, 2449–2463 (2006).
- Cheon, H.-S., Lian, Y. & Kishi, Y. Highly stereoselective and iterative synthesis of α -(1-4)-linked polysaccharides composed of 3-O-methyl-D-mannose. *Org. Lett.* **9**, 3323–3326 (2007).
- Joe, M., Bai, Y., Nacario, R. C. & Lowary, T. L. Synthesis of the docosanasaccharide arabinan domain of Mycobacterial arabinogalactan and a proposed octadecasaccharide biosynthetic precursor. *J. Am. Chem. Soc.* **129**, 9885–9901 (2007).
- Ishiwata, A. & Ito, Y. Synthesis of decosaccharide arabinan motif of Mycobacterial cell wall. *J. Am. Chem. Soc.* **133**, 2275–2291 (2011).
- Calin, O., Eller, S. & Seeberger, P. H. Automated polysaccharide synthesis: assembly of a 30mer mannoside. *Angew. Chem. Int. Ed.* **52**, 5862–5865 (2013).
- Hansen, S. U., Miller, G. J., Cliff, M. J., Jayson, G. C. & Gardiner, J. M. Making the longest sugars: a chemical synthesis of heparin-related [4]n oligosaccharides from 16-mer to 40-mer. *Chem. Sci.* **6**, 6158–6164 (2015).
- Naresh, K., Schumacher, F., Hahn, H. S. & Seeberger, P. H. Pushing the limits of automated glycan assembly: synthesis of a 50mer polymannoside. *Chem. Commun.* **53**, 9085–9088 (2017).
- Wu, Y., Xiong, D.-C., Chen, S.-C., Wang, Y.-S. & Ye, X.-S. Total synthesis of mycobacterial arabinogalactan containing 92 monosaccharide units. *Nat. Commun.* **8**, 14851 (2017).
- Islam, M., Shinde, G. P. & Hotha, S. Expedient synthesis of the heneicosaccharyl mannose capped arabinomannan of the Mycobacterium tuberculosis cellular envelope by glycosyl carbonate donors. *Chem. Sci.* **8**, 2033–2038 (2017).
- Pasari, S., Manmode, S., Walke, G. & Hotha, S. A versatile synthesis of pentacosafuranoside subunit reminiscent of Mycobacterial arabinogalactan employing one strategic glycosidation protocol. *Chem. Eur. J.* **24**, 1128–1139 (2018).
- Li, X. et al. Gold(I)-catalyzed glycosylation with glycosyl ynenates as donors. *Org. Lett.* **21**, 9693–9698 (2019).
- Joseph, A. A., Pardo-Vargas, A. & Seeberger, P. H. Total synthesis of polysaccharides by automated glycan assembly. *J. Am. Chem. Soc.* **142**, 8561–8564 (2020).

19. Wu, Y. & Ye, X.-S. Recent advances in chemical synthesis of polysaccharides. *Acta Chim. Sin.* **77**, 581–597 (2019).
20. Krasnova, L. & Wong, C.-H. Oligosaccharide synthesis and translational innovation. *J. Am. Chem. Soc.* **141**, 3735–3754 (2019).
21. Boltje, T. J., Buskas, T. & Boons, G. –J. Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat. Chem.* **1**, 611–622 (2009).
22. Di Lorenzo, F., De Castro, C., Silipo, A. & Molinaro, A. Lipopolysaccharide structures of Gram-negative populations in the gut microbiota and effects on host interactions. *FEMS Microbiol. Rev.* **43**, 257–272 (2019).
23. Wexler, A. G. & Goodman, A. L. An insider's perspective: bacteroides as a window into the microbiome. *Nat. Microbiol.* **2**, 17026 (2017).
24. Müller, M. I. et al. Intestinal colonization of IL-2 deficient mice with non-colitogenic *B. vulgatus* prevents DC maturation and T-cell polarization. *PLoS ONE* **3**, e2376 (2008).
25. Steimle, A. et al. Weak agonistic LPS restores intestinal immune homeostasis. *Mol. Ther.* **27**, 1974–1991 (2019).
26. Di Lorenzo, F. et al. Pairing Bacteroides vulgatus LPS structure with its immunomodulatory effects on human cellular models. *ACS Cent. Sci.* <https://doi.org/10.1021/acscentsci.0c00791> (2020).
27. Hashimoto, M. et al. Structural study on lipid A and the O-specific polysaccharide of the lipopolysaccharide from a clinical isolate of Bacteroides vulgatus from a patient with Crohn's disease. *Eur. J. Biochem.* **269**, 3715–3721 (2002).
28. Crich, D. Mechanism of a chemical glycosylation reaction. *Acc. Chem. Res.* **43**, 1144–1153 (2010).
29. Zhu, Y. & Yu, B. Highly stereoselective β -mannopyranosylation via the 1- α -glycosyloxy-isochromenylium-4-gold(I) intermediates. *Chem. Eur. J.* **21**, 8771–8780 (2015).
30. Yu, B. Gold(I)-catalyzed glycosylation with glycosyl o-alkynylbenzoates as donors. *Acc. Chem. Res.* **51**, 507–516 (2018).
31. Li, Y., Manickam, G., Ghoshal, A. & Subramaniam, P. More efficient palladium catalyst for hydrogenolysis of benzyl groups. *Synth. Commun.* **36**, 925–928 (2006).
32. Groves, P. Diffusion ordered spectroscopy (DOSY) as applied to polymers. *Polym. Chem.* **8**, 6700–6708 (2017).
33. Iqbal, S. et al. Enzymatic and acidic degradation of high molecular weight dextran into low molecular weight and its characterizations using novel diffusion-ordered NMR spectroscopy. *Int. J. Biol. Macromol.* **103**, 744–750 (2017).
34. van der Burgt, Y. E. M. et al. Structural analysis of monoclonal antibodies by ultrahigh resolution MALDI in-source decay FT-ICR mass spectrometry. *Anal. Chem.* **91**, 2079–2085 (2019).
35. Stambach, N. S. & Taylor, M. E. Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology* **13**, 401–410 (2003).
36. Mitchell, D. A., Fadden, A. J. & Drickamer, K. A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *J. Biol. Chem.* **276**, 28939–28945 (2001).
37. Zhang, Q. et al. Synthetic, zwitterionic Sp1 oligosaccharides adopt a helical structure crucial for antibody interaction. *ACS Cent. Sci.* **5**, 1407–1416 (2019).
38. Crawford, C. & Oscarson, S. Optimized conditions for the palladium-catalyzed hydrogenolysis of benzyl and naphthylmethyl ethers: preventing saturation of aromatic protecting groups. *Eur. J. Org. Chem.* 3332–3337 (2020).
39. Perez, S., Tubiana, T., Imbert, A. & Baaden, M. Three-dimensional representations of complex carbohydrates and polysaccharides. SweetUnityMol: a video game based computer graphics software. *Glycobiology* **25**, 483–491 (2014).
40. Chiodo, F. et al. High sensitive detection of carbohydrate binding proteins in an ELISA-solid phase assay based on multivalent glyconanoparticles. *PLoS ONE* **8**, e73027 (2013).

Acknowledgements

This work is financially supported by: the National Key Research & Development Program of China (2018YFA0507602), the National Natural Science Foundation of China (21621002 and 91753000), the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (ZDBS-LY-SLH030), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB20020000); PRIN 2017 “Glytunes” (2017XZ2ZBK and 2019-2022) to A.S.; progetto POR SATIN and Progetto POR Campania Oncoterapia (to A.M.); the European Commission (H2020-MSCA-814102-SWEET CROSSTALK project) to A.M. and A.S. We thank Dr. Zhi Ma and Weishi Li for help on the GPC and Wenmin Wu on NMR measurements.

Author contributions

B.Y. and A.M. designed the research and experiments. Q.Z. and Z.S. conducted the synthetic work. A.S. conducted the NMR and computational studies. F.C. performed the bioassays. S.N. measured the MS of long glycans. B.Y. and A.S. wrote the paper. All authors discussed the results and commented on the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-17992-x>.

Correspondence and requests for materials should be addressed to A.S. or B.Y.

Peer review information *Nature Communications* thanks Steven Townsend and the other, anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020