

# Exploration of expanded carbohydrate chemical space to access biological activity using microwave-induced acid condensation of simple sugars

James A London, Sarah L Taylor, Igor L Barsukov, Alan Cartmell and Edwin A Yates\*

*Department of Biochemistry & Systems Biology, ISMIB, University of Liverpool, Liverpool L69 7ZB, UK*

\*Corresponding author: [eyates@liv.ac.uk](mailto:eyates@liv.ac.uk)

Complex glycans are ubiquitous in Nature and essential to life; glycosaminoglycans surround all mammalian cells and are key players in development and cell signalling, while plant cell walls provide structural rigidity, even during phases of growth, and plant cells account for the vast majority of carbon fixation on earth. Despite their diverse roles, however, only a fraction of the potential chemical space hypothetically available to complex glycans has been explored. New regions of this chemical space can, nevertheless, be accessed by generating new ones which do not occur in Nature or, by modifying naturally-occurring polysaccharide structures. Collectively, we term these *new polysaccharides* (NPs). Two synthetic routes to NPs are described here; the *de novo* route, directly from monosaccharide starting materials and the *functionalization* route, involving glycosylation of existing polysaccharides. The reaction in both cases involves a simple condensation step under microwave heating, catalysed by benign organic acids. The approach is illustrated by the generation of structures with biological activities ranging from cell signalling and inhibition of bacterial growth to mimicking carbohydrate antigens from pathogenic microorganisms. The method is as applicable to fine chemicals as it is to industrial waste, for example, biotechnologically-derived D-allulose (D-psicose) or biofermentation waste, and enables rapid access to extensive chemical space, encompassing both existing biological structures and those previously unattainable. Accessing this chemical space unlocks new functionalities and applications for complex glycans in the biological, medical, biotechnological and materials science arenas.

## 1. Introduction

Polysaccharides are some of the largest and most complex macromolecules on earth as well as the most abundant. They perform essential biological roles in every kingdom of life. For instance, the extracellular matrix surrounding the cells of all mammals contains the complex glycan, heparan sulfate (HS), which has staggering potential complexity; an octasaccharide offering thousands of hypothetical sequence permutations, and its improper formation is embryonically lethal<sup>1</sup>. The surfaces of many bacterial cells are also coated in glycans which form capsules known to protect the bacteria from viral entry and help protect their host against viral infection whilst complex bacterial communities produce carbohydrate-rich biofilms contributing to antibiotic resistance, enabling their persistence in unfavourable environments. The cell walls of plants consist of a variety of complex glycans and the most abundant molecule in the biosphere is cellulose, a homopolymer of  $\beta$ 1,4 linked D-glucose, forming the primary constituent of secondary cell walls and providing plants much of their structural rigidity. Additionally, plant primary cell walls contain a complex variety of hemicellulose, pectin, rhamnogalactans and others<sup>2,3</sup> whose coordinated modification enable growth, provide support and flexibility, while also influence flowering.

Mankind has long isolated and re-purposed naturally occurring complex glycans for 'non-natural' uses of technological and societal value. Currently, cellulose is employed in filtration and dialysis membranes, the hemicellulose, xyloglucan, is used in industrial surfactants, the HS-like glycan, heparin, whose biological role is still not fully understood, is exploited as a major anticoagulant drug and the polysaccharides xanthan gum, carrageenan and gum arabic (from bacterial, algal and plant sources, respectively) find applications that include thickeners for food. Purification of such polysaccharides can involve significant effort and natural batch-to-batch variation also presents challenges to product consistency, thus, approaches that enable biotechnologically relevant complex carbohydrates to be accessed which are both economically viable and reproducible, would be desirable.

Despite the diverse and essential roles that complex carbohydrates serve in nature, and their considerable industrial value through deliberate re-purposing, the sequence and conformational landscape of complex carbohydrates that are explored by nature is, in fact, very small in comparison to that which is hypothetically available. Complex carbohydrates comprise more individual building blocks than proteins, their sugar building blocks may be D- or L- configured, possess either  $\alpha$  or  $\beta$  glycosidic bonds, involve 5- or 6-membered (furanose and pyranose) rings, and can be attached through several potential linkage positions to their

neighbours in both a linear or branched manner. Even ignoring other potential modifications, such as phosphorylation, sulfation and acetylation, this complexity provides this class of molecule much more structural variety than is available to proteins. Just as in proteins, however, structure is related intimately to the function of the molecule, and can be finely adjusted through subtle structural changes.

Accessing this unexplored sequence and conformational carbohydrate space would provide new polysaccharides (NPs) that do not exist in nature, with unique chemical and physical properties of potential value. Documented approaches have been reported for the synthesis of polysaccharides starting from the reaction of suitably protected intermediate glycosyl donor and acceptor molecules<sup>4</sup>, as well as reports of pseudo-polysaccharides with hybrid or unnatural linkages (e.g. amines, orthoesters and carbamates)<sup>5</sup>, and also those that employ ring opening polymerisation<sup>6</sup>. The use of protected donors and acceptors, however, can be expensive and time consuming, and is usually restricted to the synthesis of oligosaccharides.

The approach adopted here to generate NPs is the acid-catalysed condensation of simple sugars. This reaction has a long history and is based on Emil Fischer's 1895 acid catalysed formation of methyl glycosides from monosaccharides and methanol<sup>7</sup> to form polysaccharides<sup>8</sup> which was later adapted to employ catalysts such as phosphoric acid<sup>8,9</sup>. These approaches have subsequently been adapted further<sup>10,11</sup> and here, we apply microwave heating to drive the condensation of simple, unprotected, sugars to generate a series of exemplar NPs (**Figure 1A, B**), catalyzed by environmentally benign organic acids, such as citric acid, and explore their potential in a range of biological assays. This approach avoids the environmentally damaging metal salts, such as tin or zinc chlorides and silver triflate of some conventional Lewis acid catalysed glycosylations, and can generate NPs of several kDa within minutes. The NPs formed can then be purified easily by dialysis or gel permeation chromatography before processing for analysis or downstream applications. The approach generates polysaccharides following the most favourable thermodynamic route and foregoes formal control over regio- and stereoselectivity. The properties of the resulting NPs can still be influenced to some extent through the choice of the initial reactants and their complexity can be increased by conducting repeated rounds of synthesis introducing varied reactants. It should be noted, however, that the primary goal in this initial exploration is to create *de novo* NPs which do not exist in nature and whose biological activities, or biotechnological applications, are initially unknown, thus, the generation of pre-determined structures is not a formal aim.

Furthermore, in addition to the *de novo* approach, a *functionalisation* approach can be adopted. Existing polysaccharides can be modified using this second route to alter their properties. The starting material for functionalisation could vary from existing materials with

the aim of modifying their properties, to waste materials such as those sourced from biotechnological processes (e.g. biofuel reactors) with the aim of generating products with useful activities and, hence, added value.

We describe the application of microwave induced acid-catalysed condensation to generate NPs through both routes. We demonstrate that, without further optimization, *de novo* polysaccharides possess diverse biological activities that include influencing biofilm formation and acting as mimetics of tuberculosis arabinans and of epitopes on the cell surface of parasites. Illustrating the functionalisation route, chondroitin sulfate that has been fucosylated and chemically sulfated, inhibits cell growth in the presence of heparin-dependent growth factors. These data demonstrate that NPs are not biologically inert and can interact with biological systems in new ways. New polysaccharides can be produced quickly, simply, and the method is amenable to operating at scale. Thus, NPs have enormous potential to drive an expansion in the use of complex carbohydrates in both biological and industrial settings, utilising a simple, rapid and environmentally-friendly synthetic process.

## 2. Results

We describe below five examples of the generation of novel polysaccharides (NPs) using the present approach without further optimisation. Four examples employ the *de novo* synthetic route, starting from monosaccharide building blocks to construct non-natural glycans, thereby exploring novel chemical space. The fifth and final example involves functionalisation of the naturally occurring glycan, chondroitin sulfate-A, with L-fucose followed by chemical sulfation, to introduce new biological functionality into an existing polysaccharide.

### *Generation of polysaccharides able to influence biofilm formation by Pseudomonas aeruginosa.*

*Pseudomonas aeruginosa* (*P.aeruginosa*), an opportunistic pathogen, is responsible for infection in vulnerable individuals and is particularly common in clinical settings. For instance, *P.aeruginosa* contamination of medical devices is a widespread problem, and can cause pneumonia in immunocompromised patients, as well as persistent infections of the lungs of cystic fibrosis sufferers. *Pseudomonas aeruginosa* exhibits extensive drug resistance and is also able to form biofilms, presenting additional resistance to drug treatment and cleaning procedures. Biofilms are complex structures containing polysaccharides, proteins, nucleic acids and lipids, that are able to adhere to the surfaces of tissues and implanted devices and thereby form an environment that protects the bacteria from both antibiotics and the immune system of the host, as well as providing a site in which other bacteria may become resident<sup>12</sup>.

To investigate the use of NPs as novel reagents in disrupting biofilm formation we employed the *de novo* synthetic route to generate several NPs, each using a simple sugar (D-galactose, L-arabinose, L-rhamnose, D-glucose, D-arabinose and D-mannose respectively) as the building block catalyzed by citric acid and formed complex, branched polymers typically of ~ 5 kDa. These were added at a concentration of 100  $\mu\text{g mL}^{-1}$  to bacterial cell cultures of *P.aeruginosa*, and the effect on biofilm formation monitored using a crystal violet- based method of biofilm detection<sup>13</sup>. The effects on biofilm formation of *P.aeruginosa* by the NPs was varied; the D-galactose NP caused a ca.36 % reduction in biofilm formation, while others reduced biofilm formation less dramatically (L-arabinose, L-rhamnose and D-glucose NPs) or not at all (D-arabinose and D-mannose NPs) (**Figure 2**). The mechanisms behind the action of these NPs, which have not been optimized, are unknown, but the potential utility is clear.

*Poly D-arabinose NPs are recognised by anti-lipoarabinomannan (LAM) polysaccharide antibodies.*

The review by Peltier *et al.*<sup>14</sup> summarizes over fifty polysaccharides from microorganisms that contain both pyranose and furanose ring forms of hexoses. Of these, about one third were recognized as playing a role in pathogenic processes and included major disease-causing genera such as *Mycobacterium*, *Shigella* and *Trypanosoma*. During the mutarotation process undergone by free monosaccharides in aqueous solution, five-membered furanose rings form more quickly than six-membered pyranose rings but, unlike pyranoses, furanose ring forms do not favour significantly the formation of particular anomeric configurations, suggesting that complex NP structures are likely to result during the condensation process; a possibility that was examined in more detail by NMR. Some of the structures formed are likely to resemble naturally occurring sequences, while others are expected to be entirely new to nature.

In *Mycobacterium tuberculosis*, two polysaccharides with immune-modulatory properties predominate; lipoarabinomannan (LAM) and arabinogalactan (AG), which both contain furanose ring forms. In the branched LAM polysaccharide,  $\beta$  (1-6) and  $\beta$  (1-5) linked D-galactofuranose, as well as  $\beta$  D-arabinofuranose, have been identified<sup>15-17</sup>. Structures comprising these furanose ring forms are of particular interest, since they are not found in mammals and thus they could act, or be modified to act, as potential enzyme inhibitors, against Mycobacterial transferases or, as a source of antigenic structures to prime the immune system. The LAM structures include terminal sections that represent epitopes responsible for the antibody response evident commonly in mycobacterial infections<sup>17</sup>. The availability of an anti-LAM antibody, used as a diagnostic for the disease, offers the possibility of identifying epitopes based on D-arabinose rapidly, for future exploitation, using the *de novo* synthetic

route. To this end, an NP was produced from D-arabinose and both direct binding of the anti-LAM antibody to the arabinose NP, as well as the ability of the NP to compete with binding of the antibody to the positive control, were demonstrated (**Figure 3**). This D-arabinose NP was also subjected to analysis by NMR which confirmed its complex structure, comprising anomeric linkages of both  $\alpha$  and  $\beta$  pyranose and furanose ring forms (**Figure 4 and inset table**). NMR revealed signals for both free primary hydroxyl groups and those linked by glycosidic bonds, evident in the HSQC spectrum with  $^{13}\text{C}$  chemical shifts around 64 ppm and 69-70 ppm respectively [18]. The analysis highlights the complexity of the NP products formed, even from a single starting sugar, in this case revealing both  $\alpha$  and  $\beta$  D-arabinose in both furanose and pyranose forms [19]. Measurement of one ( $^1J_{\text{CH}}$ ) and three bond ( $^3J_{\text{HH}}$ ) NMR coupling constants can provide information regarding anomeric configuration and conformation and measurements of free D-arabinose, provided  $^3J_{\text{H}_1\text{H}_2}$  and  $^1J_{\text{C}_1\text{H}_1}$  values for reference of 3.15/179.5 Hz and 4.53/180.3 Hz for D-arabinofuranose and 3.66/181.2 Hz and 7.78/178.7 Hz for  $\alpha$  and  $\beta$  D-arabinopyranose anomeric forms respectively. The  $^1J_{\text{CH}}$  couplings for the pyranose forms in the D-arabinose NP (**Figure 3C**) corresponded relatively closely to those of free D-arabinose (**Figure 3E**), consistent with  $^4\text{C}_1$  chair forms but, those of the furanose forms varied, most likely indicating a conformation for these within the polymer that differs from the expected  $^2\text{T}_3$  form of free D-arabinofuranose in solution<sup>18</sup>. Arabinan structures extracted from *Mycobacterium spp.*<sup>15-17</sup> lipoarabinomannan (LAM) that are considered responsible for the immunogenic response, contain a range of linkages and configurations, including  $\beta$  D-arabinofuranose (1-2),  $\alpha$  D-arabinofuranose, and -3 and -5 linked  $\alpha$  D-arabinofuranose residues, as well as terminal D-arabinofuranoside units<sup>17</sup>; signals consistent with some of these structures were also evident in the NMR spectra of the D-arabinose NP.

*Generation of polysaccharides recognized by a lectin that binds  $\alpha$  D-galactose structures that are characteristic of the surfaces of trypanosomes.*

Oligo- and polysaccharide structures terminating with  $\alpha$  D-galactose are characteristic glycan structures on the cell surfaces of trypanosomes and, since  $\alpha$  D-gal linked structures do not occur in humans, they have been proposed as both diagnostic tools and as components of possible vaccines<sup>20</sup>. Indeed, a conjugate of an  $\alpha$  (1-6)-D-galactose disaccharide and BSA was shown to be protective against *Leishmania major* infection<sup>21</sup>. The  $\alpha$ -D-galactose epitope has also provided an antigenic obstacle to xenotransplantation, since it is present in most other mammalian species. Following bites by ticks,  $\alpha$  (1-3)-D-galactose disaccharides can also be the source of acquired allergy to meat in the diet. It has been suggested that the loss of  $\alpha$ -D-galactose biosynthetic capability in humans and apes, which occurred 28 million years ago, bestowed a selective advantage by providing immunogenic protection against a number of pathogens (discussed in <sup>22</sup>).

Here, the *de novo* synthesis route was used to generate NPs with potential utility in these applications, and the detection of  $\alpha$  D-gal epitopes was demonstrated using a lectin from *Griffonia simplicifolia* with high affinity for  $\alpha$  linked D-galactose units (**Figure 2C**).

New polysaccharides from a biotechnologically-derived starting material, D-allulose.

The approach also provides a route to generate NPs from biotechnologically-derived starting materials and can include those produced from rare or novel monosaccharides. An example is provided by the condensation of D-allulose (D-psicose), the C-3 epimer of D-fructose, which is used as a low-calorie dietary sweetener, and is found only very rarely in nature. D-allulose is produced on a large scale by biotechnological approaches<sup>23</sup>, however, and the *de novo* route generates NPs with potentially novel attributes that presumably include polysaccharide structures to which existing carbohydrate degrading enzymes have never been exposed. D-allulose was condensed using two different organic acid catalysts, citric ( $\text{pK}_{\text{a}(1)}$  3.1) and *cis*-aconitic (a dehydrated derivative of citric acid,  $\text{pK}_{\text{a}(1)}$  2.8) acids. The NP produced with citric acid catalysis exhibited a molecular weight by viscosity measurements against dextran standards of ca. 5.3 kDa, while that of the *cis*-aconitic acid catalyst was lower, at ca. 4 kDa (**Table 1**).

Generation of a fucosylated glycosaminoglycan via functionalisation and its subsequent sulfation enabled modulation of FGF cell signaling.

Finally, we illustrate the generation of a NP made by the *functionalisation* route. As an example, L-fucose was reacted with an existing polysaccharide, the glycosaminoglycan, chondroitin-4-sulfate (chondroitin sulfate-A), followed by an O-sulfation step and the sulfated product was tested for its ability to compete against heparin in a cell-based assay of signalling. In this assay, heparin acts as a proxy for the naturally occurring heparan sulfate. The BaF cell line is devoid of endogenous heparan sulfate, which is the obligatory co-receptor in FGF-FGFR tyrosine kinase driven cell signalling<sup>24</sup>, and cells were transfected with a single, selected FGFR (FGFR1c), then cultured in the presence of a chosen FGF (FGF-1) together with heparin which served as the positive control. The test NP, chemically sulfated fucosylated-CS-A, was then co-incubated under these conditions as a potential inhibitor and its ability to compete the cell proliferation capability of the FGF-heparin combination was determined (**Figure 6**).

The FGF-driven coordination of wound and nerve repair as well as developmental processes, are all areas of extensive research. Derivatisation of existing materials such as GAGs, many of which have already found application in practical devices, offer a direct means of further tailoring their properties.

### 3. Discussion

It is well-known that the number of hypothetical sequence arrangements of the amino acids that comprise a typical protein is vast, but what is not appreciated so widely is that this complexity is dwarfed by the possible variation provided by carbohydrates of much shorter lengths. Complex carbohydrates fulfil numerous, varied and essential roles throughout nature, indeed, the emergence of eukaryotic life closely tracks the emergence and complexity of sulfated glycans at the cell surface<sup>25</sup>. Despite this, to facilitate life, Nature has only tapped into a small subset of the plethora of complex glycans that potentially could be generated. Furthermore, mankind has been able to utilise this limited source of naturally occurring glycans for a wide range of further uses for which they did not evolve originally. This leads to the tantalizing speculation as to what could be achieved were the vast chemical and conformational space of complex carbohydrates to be explored more comprehensively. There are undoubtedly carbohydrate materials, yet to be conceived, which are capable of serving known carbohydrate functions, as well as those capable of fulfilling purposes for which carbohydrates have, hitherto, not been proposed as a solution.

Among the more obvious and immediately applicable uses would be the generation of NPs to replace existing polysaccharides for which purification and natural variation issues arise. For example, xyloglucan, which can be used in surfactants<sup>26</sup> is known to vary from batch to batch depending on the source<sup>27</sup>, necessitating regular optimisation of processes. It would be feasible to generate NPs with similar properties from either the same sugar building blocks as xyloglucan, or from others bearing no obvious relation, but that still yielded a polymer exhibiting similar or enhanced properties. The use of polysaccharides as food thickeners is another area in which NPs, generated through the present approach, could readily provide replacement polysaccharides and thereby avoid time-consuming methods of purification, whether the polysaccharides in current use originate from marine, terrestrial, or bacterial sources. For example, the extraction of xylan from terrestrial plants is known to alter the structure of the original material under highly alkaline conditions, which are also environmentally unfriendly, by removing acetyl and ferulic acid groups<sup>28, 29</sup>. For some applications, these groups need to be replaced chemically. Again, by beginning to explore carbohydrate chemical space through microwave assisted acid-catalysed condensation, it may be possible to generate substitutes for xylan with similar properties in an environmentally responsible manner.

The generation of NPs that contain entirely new structures, such as those produced here from D-allulose that are, as far as is known, alien to nature, raises a further interesting

possibility. This is to test both the degree to which extant enzymes are specific and to explore the possibility of evolving new enzymes experimentally that are capable of degrading such new structures, for example, from the polysaccharide degrading enzymes of bacteria. Two types of potential substrate can be envisaged. The first would comprise NPs made from naturally occurring monosaccharides whose polysaccharides contain both naturally occurring and alien polymerised structures, while the second would employ a sugar as the starting material that does not exist in nature (or only as a very rare monosaccharide), such as D-allulose, to generate polysaccharides that have never been encountered in nature. This would provide an experimental regime under which the evolution of carbohydrate active enzymes could be monitored from a given starting point and with defined materials, providing a unique insight into how new carbohydrate active enzymes can be evolved. It would also be a method of applying a bespoke glycan toolbox to newly generated NPs, enhancing their utility as the ability to manipulate them improves.

The ability to inhibit biofilm formation could provide a role for optimised NPs as a protective coating on surfaces and devices but, those able to support biofilm formation could be used in a sacrificial role or perhaps even to promote biofilm formation at a location of choice should this be desirable. A further advantage is that the ability of *P. aeruginosa* to evolve to the effects of a NP would be far reduced compared to antibiotics, as the genetic tools to degrade these NPs do not exist in nature. Should resistance begin to develop, the ability to either generate completely new NPs or *functionalise* existing ones (see final results section below) is a rapid process.

It should also be appreciated that the biological activities of the NPs generated in this work serve as a proof of principle and they required no optimisation to demonstrate their biological functions. This supports the view that it is not a question of *whether* newly generated NPs will have biological and industrial applications but, rather, *what* those will (or could) be. Several general routes by which the structures of NPs could be influenced include the timed addition of reactants with contrasting properties, for example, adding a charged sugar to an already formed branched but neutral scaffold, or the deliberate manipulation of the equilibrium composition (furanose/pyranose) of the starting materials through timing (exploiting the relatively rapid equilibration of furanose and pyranose forms or their slower anomersation), or manipulation through pH, and the use of cations. For higher levels of control over the degree of branching, the use of conventionally protected sugars can also be envisaged as initial scaffolds for polymerisation. This would enable, for example, NPs with some pre-designed characteristics to be synthesized.

Despite it being based on a reaction that harks back to the early days of carbohydrate chemistry, the technique for microwave assisted acid catalysed condensation offers a rapid, facile method for beginning to explore the enormous untapped potential of carbohydrate

sequence and conformational space. Starting from widely-available simple sugars and employing environmentally benign catalysis, it promises new materials, some of which will be able to replace current chemicals in products such as cosmetics, as well as generating new materials with completely novel functionality.

#### 4. Materials and Methods

General synthetic procedure: - A chosen reducing sugar (monosaccharide hexose (500 mg), or pentose (417 mg), 2.78 mmoles) was dissolved in an aqueous solution of the organic acid catalyst (citric (50 mg) or *cis*-aconitic acid (45 mg (0.26 mmoles) in 2.5 mL water) in a 10 mL conical flask plugged with glass wool, and was degassed using nitrogen (5 mins). The mixture was microwaved (medium-high power (4 mins) followed by high power (1 additional minute) in an atmosphere of nitrogen. The thick syrup, often exhibiting colouration ranging from very pale yellow to dark brown, was then allowed to cool and was re-dissolved in water (10 mL), dialysed against water (2L, 3.5 kDa cut-off) and dried by rotary evaporation before further analysis by NMR and other physico-chemical techniques, or use in biological assays. As an example, the NP made using this method from 417 mg of D-arabinose with citric acid catalysis, yielded 60 mg with  $M_r > 3,500$  Da (nominal cut-off of the dialysis tubing).

Generation of fucosylated chondroitin sulfate-A and its subsequent sulfation: - The GAG oligosaccharide, chondroitin sulfate-B ( $M_w$  26 kDa) was functionalised by reaction of CS-A (200 mg) and L-fucose (30 mg), citric acid (40 mg) in water (2.5 mL) over 4 mins in a commercial microwave oven to generate fucosylated CS (Fuc-CS) with  $M_w$  34 kDa (determined by viscosity measurements relative to dextran standards). Subsequent O-sulfation was achieved by reacting the fucosylated product (52 mg) with 3 eq. pyr. $SO_3$  complex in DMF (3 mL) overnight at room temperature (16°C), followed by neutralisation to pH 8 (0.1 N NaOH), precipitation in cold ethanol and dialysis (3.5 kDa cut-off v  $H_2O$ ) to recover the product and the dialysate was dried for further use.

Assay of biofilm formation by *Pseudomonas aeruginosa*: - *Pseudomonas aeruginosa* was cultured under standard conditions<sup>30</sup>. Novel polysaccharides, made from the condensation of L-arabinose, L-rhamnose (see General Synthetic procedure above) and Fuc-CS were formed as described above purified by dialysis (>3.5 kDa (dialysis) and assayed (100  $\mu g \cdot mL^{-1}$ ) for their influence on biofilm formation (*Pseudomonas aeruginosa*), in a 96-well assay that detects biofilm formation through staining with crystal violet<sup>13</sup>.

Inhibition assay of FGF signaling in BaF cells: - BaF3 cells (murine lymphocytes) were transfected with the FGF receptor (FGFR1c) and maintained in RPMI-1640 supplemented with 10% fetal calf serum, 2 mM glutamine, 100 U nL<sup>-1</sup> pen G, 1 nM FGF1, 50 µg mL<sup>-1</sup> streptomycin sulfate and 2 ng mL<sup>-1</sup> IL-3. Assays for cell function were as described [19]. Positive control experiments included addition of porcine mucosal heparin (3 – 1000 µg mL<sup>-1</sup>) and inhibition involved addition of the test NP, sulfated L-fucosyl chondroitin sulfate-A (M<sub>w</sub> 34 kDa, 3 – 1000 µg mL<sup>-1</sup>) in the presence of heparin at 1000 µg mL<sup>-1</sup>.

Inhibition of anti-LAM antibody binding: - Assays were conducted using a commercial lipoarabinomannan ELISA kit (Aviva Biosystems Biology, San Diego, USA) in two formats – one in which direct antibody binding to an NP (made by condensation of D-arabinose and previously coated onto an ELISA plate) was measured. The D-arabinose NP was coated at 100 µL in PBS per well (72 h, room temp) followed by washing and blocking (4% BSA in PBS) and antibody probing (1/1000 dilution in PBS from 1 mg mL<sup>-1</sup> in 4% BSA in PBS) incubation (1 h), washing and detection by probing with HRP conjugated secondary antibody and subsequent detection of tetramethylbenzidine at 450 nm. The second assay was a competition ELISA, in which the D-arabinose NP served as a potential competitor for the binding of the anti-LAM antibody to the positive control (manufacturer's instructions).

Agglutination assay using a lectin from *Griffonia simplicifolia* for α-D-galactose epitopes: - Lectin-coated latex beads (Sigma, UK, polystyrene beads (0.8 µm, 0.5 mL suspension) were prepared by incubating (overnight, 4°C) with a solution of a lectin able to bind α D-galactose units (isolectin GS-IB4 from *Griffonia simplicifolia*) using 100 µL of a 1 mg mL<sup>-1</sup> in PBS stock solution (Invitrogen, UK) followed by extensive washing in PBS. Comparisons of agglutination were made by addition of a small volume (10 µL) of a suspension of the beads to a solution of the test NP (10 µL, 0.1 mg mL<sup>-1</sup>) on a glass slide.

NMR characterisation of new polysaccharides: - NMR experiments were performed on products (20 mg) in 700 µL of D<sub>2</sub>O at 343 K using a 600 MHz Bruker Avance II+ spectrometer fitted with a TCI CryoProbe and chemical shift values are quoted relative to DSS (0 ppm). In addition to 1-dimensional (<sup>1</sup>H and <sup>13</sup>C) spectra, heteronuclear (HSQC and HMBC) 2-dimensional spectra, and coupled HSQC, coupled HSQC and COSY spectra were also recorded on a sample of D-arabinose (5 mg mL<sup>-1</sup>) to measure <sup>1</sup>J<sub>CH</sub> and <sup>3</sup>J<sub>HH</sub> coupling constants respectively. Spectra were processed using Bruker TopSpin and <sup>1</sup>H/<sup>13</sup>C HSQC integration was performed using the INFOS spectrum fitting software<sup>31</sup>). Proton spectrum integration was performed using Bruker TopSpin software.

Measurement of molecular mass by viscosity: - Viscosity measurements offer one way to estimate the molecular mass of NPs. This can be done by measuring the flow rates of the solvent ( $t_0$ ) and a series test solutions ( $t$ ) of polysaccharides of varying, but known, molecular weights at a given concentration ( $5 \text{ mg mL}^{-1}$ ). Plotting the relative viscosity,  $\eta_r (= t/t_0)$ , against concentration,  $c$  (g/ml), the intrinsic viscosity,  $\eta$ , can be recovered using the Mark-Houwink equation (**equation 1**) and, by reference to the constants  $K$  and  $\alpha$ , the molecular weight can be found.

$$\text{Equation 1. } [\eta] = KM^\alpha \text{ hence, } \ln[\eta] = \alpha K \ln M$$

(where  $[\eta]$  is the intrinsic viscosity,  $M$  is the molecular weight,  $K$  and  $\alpha$  are constants for the solvent system and type of molecule, available from the literature).

Literature values for the Mark-Houwink constants,  $K$  and  $\alpha$ , for dextran in water at  $20^\circ\text{C}$  (molecular weight range 70 kDa to 2 MDa)<sup>32</sup> are available but, a more direct approach is to employ a series of dextrans of defined molecular weight and to construct a calibration curve to convert the experimental measurement of  $t/t_0$  to molecular weight ( $M_w$ ). Dextran standards of  $M_w$ ; 7, 12, 16, 21.4 and 196 kDa (5% w/v solutions in water; Pharmacosmos, Denmark) were employed for this purpose.

## Funding

A.C. was supported by the Academy of Medical Sciences/Wellcome Trust through the springboard grant SBF005/1065 163470. J.A.L. and E.A.Y. acknowledge support from the BBSRC grant BB/M011186/1.

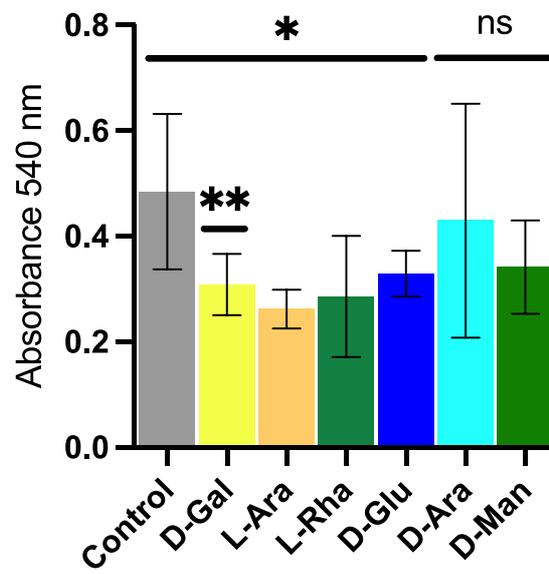
## Acknowledgements

Dr. Scott E Guimond (Keele University, UK) and Prof. David Ornitz (Washington University in St. Louis, Missouri, USA) are thanked for provision of BaF cell lines.

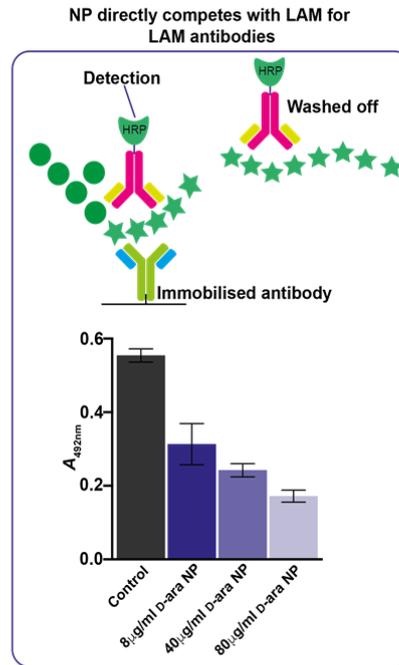
## Abbreviations

D-Gal, D-galactose; L-Ara, L-arabinose; L-Rha, L-rhamnose; D-Glu, D-glucose; D-Ara, D-arabinose; D-man, D-mannose; LAM, lipoarabinomannan.

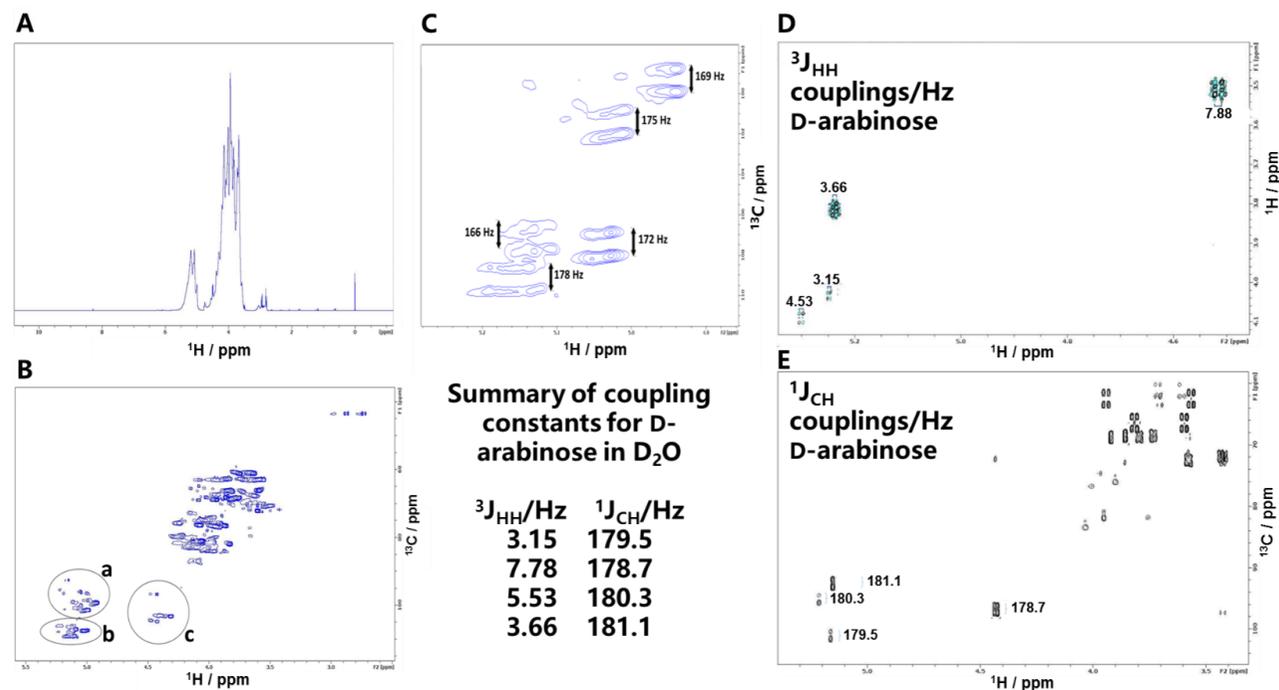




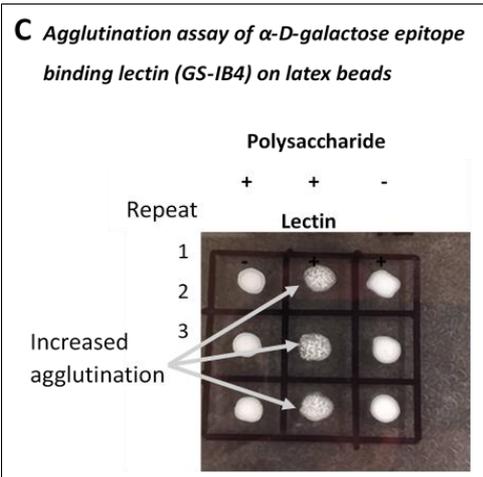
**Figure 2. The ability of novel polysaccharides to alter biofilm formation.** New polysaccharides formed by condensation of D-galactose, L-arabinose, L-rhamnose or D-glucose were all able to inhibit biofilm formation at 100  $\mu\text{g mL}^{-1}$ . ns indicates no significant difference.



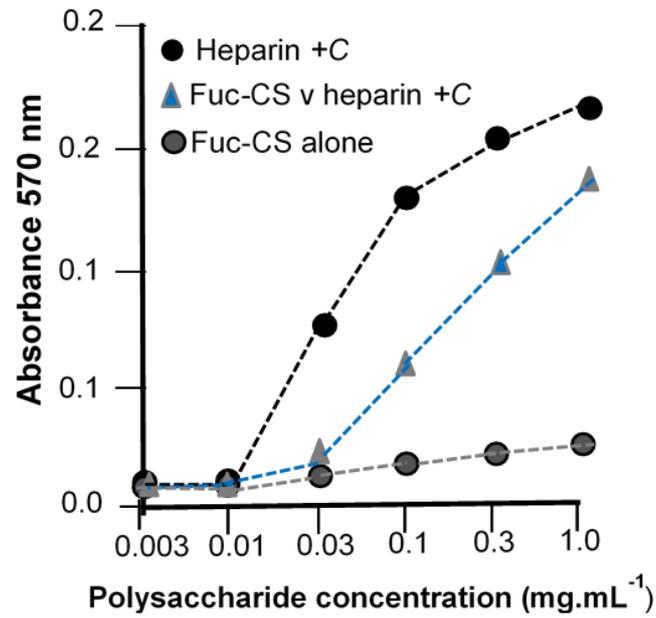
**Figure 3. Novel polysaccharides generated from D-arabinose can act as mimics of lipoarabinomannan from *Mycobacterium tuberculosis*.** The arabinose NP is able to compete (applied at 8, 40 and 80  $\mu$ g mL) for binding to the anti-LAM primary antibody.



**Figure 4.** **A.**  $^1\text{H}$  NMR spectrum of the NP formed by condensation of D-arabinose. **B.** Decoupled HSQC NMR spectrum of the NP of D-arabinose revealing its complex structure. Inset: anomeric regions containing signals provisionally assigned to; *a.*  $\beta$  D-arabinopyranose and terminal  $\beta$  D-arabinofuranose. *b.* -2 linked  $\alpha$  D-arabinofuranose, -3,5 linked D-arabinofuranose, -2,3,5 linked D-arabinofuranose and terminal D-arabinofuranose. *c.*  $\alpha$  D-arabinopyranose<sup>33,34</sup>. **C.** Coupled HSQC NMR spectrum of anomeric region for D-arabinose NP in water (D<sub>2</sub>O) showing  $^1J_{CH}$  couplings between H-1 and C-1. **D.** Coupled COSY NMR spectrum of anomeric region for D-arabinose in D<sub>2</sub>O showing  $^3J_{HH}$  couplings between H-1 and H-2. **E.** Coupled HSQC NMR spectrum of anomeric region for D-arabinose showing  $^1J_{CH}$  couplings between C-1 and H-1. The inset summarizes the anomeric coupling constants for D-arabinose.



**Figure 5. Agglutination assay employing an NP formed by condensation of D-galactose.** The NP formed by condensation of D-galactose is recognized by a lectin (isolectin GS-1B4 from *Griffonia simplicifolia*) specific for a D-galactose epitopes such as those on the surface of Trypanosomes. The lectin was coated on latex beads and addition of the NP resulted in agglutination.



**Figure 6. Functionalised chondroitin sulfate is a competitive inhibitor of fibroblast growth factor signaling.** The functionalized chondroitin sulfate NP was titrated against the positive control, heparin. Heparin serves as mimic of heparin sulfate, the endogenous and obligatory co-receptor of FGF signaling, and supports signaling between FGF1 and FGFR1c in BaF cells that have been modified to lack the ability to synthesise heparan sulfate.

	Flow times (t/t <sub>0</sub> ) in viscosity apparatus.				
Molecular weight dextran standards/kDa	7	12	16	21.4	196
t/t <sub>0</sub> (rel. to H <sub>2</sub> O = 1.00)	1.29	1.51	1.59	1.84	4.37
D-allulose NP (citric acid) (SD)	1.22 (0.3)	M <sub>w</sub> ca. 5.3 kDa: dp 24/25			
D-allulose NP ( <i>cis</i> -aconitic acid) (SD)	1.17 (0.2)	M <sub>w</sub> ca. 4 kDa: dp 33			

**Table 1. Molecular weight determination of D-allulose NPs formed under catalysis by citric acid and *cis*-aconitic acid by viscosity measurements relative to dextran standards.**

## References

- 1 M Ringvall, J Ledin, K Holmborn, A-M Olofsson, L Kjellen and E Forsberg. Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1. *J Biol Chem* (2000) **275**, 25926-30.
- 2 NC Carpita. Structure and Biogenesis of the cell walls of grasses. *Ann.Rev.Plant Physiol. Plant Mol. Biol.*, 47, (1996) **47**, 445-476.
- 3 J Vogel, *Curr. Opin. Unique aspects of the grass cell wall. Plant Biol.*, (2008) **11**, 301-307.
- 4 SS Nigudkar & AV Demchenko. Stereocontrolled 1,2 *cis* glycosylation as the driving force of progress in synthetic carbohydrate chemistry. *Chem.Sci.*, (2015) **6**, 2687.
- 5 R Xiao & MW Grinstaff. Chemical synthesis of polysaccharides & polysaccharide mimetics. *Prog. Polym.Sci.*, (2017) **74**, 78.
- 6 GL Gregory, EM Lopez-Vidal and A Buchard. Polymers from sugars: cyclic monomer synthesis, ring-opening polymerisation, material properties & applications. *Chem.Commun.*, (2017) **53**, 2198.
- 7 E Fischer, *Über die verbindungen der Zucker mit den Alkoholen und Kenonen. Ber. Dtsch. Chem. Ges.*, (1895) **28**, 1145-1167.
- 8 E Pascu and PT Mora, Polycondensation of D-glucose and other simple sugars in presence of acids. *J.Am.Chem.Soc.*, (1950) **72**, 1045.

- 9 PT Mora and JW Wood, Synthetic polysaccharides. I. Polycondensation of glucose. *J. Am. Chem. Soc.*, (1958) **80**, 685-692.
- 10 PR Murray (1988). Polydextrose. In G. G. Birch, & M. G. Lindley (Eds.), *Low-calorie products* (pp. 83–100). New York: Elsevier Applied Science.
- 11 AM Daines, Z Smart, IM Sims, GW Tannock and SFR Hinkey. Synthesis, characterization & microbial utilization of amorphous polysugars from lactose. *Carbohydr. Polym.*, (2015) **117**, 223-229.
- 12 N Høiby, O Ciofu and T Bjarnsholt. *Pseudomonas aeruginosa* biofilms in cystic fibrosis. *Future Microbiology* (2010) **5**, 1663-1667.
- 13 GA O'Toole, Microtiter dish biofilm formation assay, *J. Vis. Exp.*, (2011) **47**, 2437.
- 14 new 14P Peltier, R Euzen, R Daniellou, C Nugier-Chauvin and V Ferrieres. Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. *Carbohydr. Res.*, (2008) **343**, 1897-1923.
- 15 GS Besra, K-H Khoo, MR McNeill, A Dell, HR Morris and PJ Brennan. A new interpretation of the structure of the mycolyl-arabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and <sup>1</sup>H nuclear magnetic resonance spectroscopy. *Biochem.*, (1995) **34**, 4257-66.
- 16 M Daffe, M McNeill and PJ Brennan. Major structural features of the cell wall arabinogalactans of *Mycobacterium*, *Rhodococcus*, and *Nocardia* spp. *Carbohydr. Res.*, (1993) **249**, 383-398.
- 17 D Chatterjee, CM Bozic and M McNeill. Structural features of the arabinan component of the lipoarabinomannan of *Mycobacterium tuberculosis*. *J. Biol. Chem.*, (1991) **266**, 9652-60.
- 18 RS Shallenberger, *Advanced Sugar Chemistry*. Ellis Horwood, Chichester UK, 1982.
- 19 HN Cheng and TG Neiss. Solution NMR spectroscopy of food polysaccharides. *Polymer Reviews* (2012) **52**, 81-114.
- 20 APV Moura, LBC Santos, CRN Brito *et al.*, Virus-like particle display of the alpha-gal carbohydrate for vaccination against *Leishmania*. *ACS Central Science* (2017) **3**, 1026-1031.
- 21 E Iiguez, NS Schocker, K Subramaniam, S Portillo, AL Montoya, WS Al-Salem, CL Torres, F Rodriguez, OC Moreira, A Acosta-Serrano, K Michael, IG Almeida and RA Maldonado. An  $\alpha$ -Gal-containing neoglycoprotein vaccine partially protects against murine cutaneous leishmaniasis caused by *Leishmania major*. *PLOS Negl Trop Dis* (2017) **11**, e0006039.
- 22 JM Wilson, AJ Schuyler, N Schroeder and TAE Platts-Mills. Galactose- $\alpha$ -1,3-Galactose: Atypical Food Allergen or Model IgE Hypersensitivity? *Curr. Allergy Asthma Rep.*, (2017) **17**, 8.
- 23 W Mu, W Zhang, Y Feng, B Jiang and L Zhou. Recent advances on applications and biotechnological production of D-psicose, *Appl. Microbiol. Biotechnol.*, (2012) **94**, 1461-7.
- 24 DM Ornitz, A Yaron, JG Flanagan, CM Svahn, E Levi and P Leder. Heparin is required for cell-free binding of basic FGF to a soluble receptor and for mitogenesis in whole cells. *Mol. Cell. Biol.*, (1992) **12**, 240-247.

- 25 A Ori, MC Wilkinson and DG Fernig, A systems biology approach for the investigation of the heparin/heparin sulfate interactome. *J.Biol.Chem.*, (2011) **286**, 19892-19904.
- 26 L Greffe, L Bessueille, V Bulone and H Brumer. Synthesis, preliminary characterization, and application of novel surfactants from highly branched xyloglucan oligosaccharides. *Glycobiol.*, (2005) **15**, 437-445.
- 27 A Schultink, L Liu, L Zhu and M Pauly. Structural diversity and function of xyloglucan sidechain substituents. *Plants* (2014) **3**, 526-542.
- 28 X Chen, J Shekuro, R Elander and M Tucker. Improved xylan hydrolysis of corn stover by deacetylation with high solids dilute acid pretreatment. *Ind. Eng. Chem. Res.*, (2012) **51**, 70-76.
- 29 P M-A Pawar, S Koutaniemi, M Tenkanen and EJ Mellerowicz. Acetylation of woody ligocellulose: significance and regulation. *Front. Plant Sci.*, (2013) **4**, 118.
- 30 AE LaBauve and MJ Wargo, *Curr. Growth and Laboratory Maintenance of Pseudomonas aeruginosa*. *Protocol. Microbiol.*, (2012) mc06e01s25.
- 31 AA Smith. INFOS: spectrum fitting software for NMR analysis. *J. Biomol. NMR* (2017) **67**, 77-94.
- 32 E Antoniou and M Tsianou. Solution properties of dextran in water and formamide. *J.App.Polym.Sci.*, (2012) **125**, 1681-92.
- 33 F Dourado, SM Cardoso, AMS Silva, FM Gama and MA Coimbra. NMR structural elucidation of the arabinan from *Prunus dulcis* immunological active pectic polysaccharides. *Carbohydr.Polym.*, (2006) **66**, 27-33.
- 34 BA Wolucka, MR McNeill, E de Hoffmann, T Chojnacki and PJ Brennan. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of etambutol on mycobacteria. *J.Biol.Chem.*, (1994) **269**, 23328-35.