Glycodendriproteins: a synthetic glycoprotein mimic enzyme with branched sugar-display potently inhibits bacterial aggregation.

Phillip M. Rendle,¹¶ Andreas P. Seger,¹§ Joao Rodrigues,¹¥ Neil J. Oldham,¹ Richard R. Bott,² J. Bryan Jones,³ Marjorie M. Cowan,⁴* and Benjamin G. Davis¹*

Supporting Information

Further Discussion

Accessibility to the branched glycan structures required for this study: Branched structures have also previously been introduced to proteins using valuable but nonselective chemistry;¹ the structure-function relationships for the resulting heterogeneous protein products are thus difficult to characterize. Advances in the synthesis of oligosaccharides have made available small quantities of some highly-branched carbohydrate structures, their total synthesis can be difficult and their attachment to proteins in a controlled site-selective manner to create single pure glycoprotein glycoforms has not been achieved. Indeed, despite impressive developments,² the fact that the synthesis of branched oligosaccharides is not yet routine is illustrated by the fact that, to the best of our knowledge, tri-antennary oligosaccharides are the most elaborately branched oligosaccharide structures synthesized to date³⁻⁶ and yet these fall short of the higher valency levels that we required for our study. Moreover, these have not been incorporated into proteins. The current limitations of glycoprotein construction have therefore until now prevented the examination of structures beyond those carrying a single carbohydrate antenna. Higher valencies have been achieved with glycodendrimers then with oligosaccharides and therefore we considered that these would serve as more accessible mimics of the N-linked glycan structures on glycoproteins. Representative mono- and multi-antennary carbohydrate reagents 1-4 were therefore used to chemical modify a cysteine mutant created through mutagenesis in this study.

Rendle et al

page S2

Importance of Gram Positive Bacteria as a Pathogen Target and Anti-infective Strategies: As a pathogen *A. naesludii* is a member of the bacterial class currently defined by many in the public health and infectious disease communities as the greatest challenge in hospital infections.⁷ Bacterial adhesion is a complex process involving more than just fim-type adhesins. Moreover, so-called virulence inhibition may not always be a successful strategy.⁸ Yet the interactions with Gal-binding lectins and subsequent sugar-dependent inhibition of aggregation shown here demonstrate the effectiveness of glycodendriproteins as synthetic mimics of glycoproteins and suggest that in certain cases antiinfective activity may be generated using glycoconjugates.

Observed Multivalent Inhibition Pattern of Glycodendriprotein-enzymes: The novel class of glycoconjugates disclosed here, glycodendriproteins, has allowed us to assess the role of carbohydrate multivalency when present on protein surfaces and in adhesion processes of the gram-positive pathogenic bacteria, *A. naeslundii*. This technique has identified that the optimal binding needed to target the adhesin lectins in *A. naeslundii* can be best achieved in this system through the use of bi-antennary structures that mimic the bi-antennary complex structural class of glycoproteins that are postulated to be their natural ligands.⁹ This is supported by suggestions that the *A. naeslundii* Gal-binding adhesin is functionally active as a dimer.¹⁰ As is observed in nature, the branched display of oligosaccharide structures (Additional Figure 1) is shown to be associated with a higher affinity for lectins than for those containing single monosaccharide structures.

In contrast to the increasing levels of lectin binding observed with increasing branching in ELLA assays, optimal inhibitory potency was observed for S156C-**2b** that has a biantennary carbohydrate structure. Although an initially surprising result since S156C-**2b**,**3**,**4** all have the potential for divalent display of Gal, it parallels previous observations that a bi-antennary structure and not a higher level of valency gave highest lectin sequestration by a kinetically-controlled process.¹¹ In this regard, it may be that the lack of increased efficacy beyond this bi-antennary structure might relate to an increased ability of the more highly multivalent conjugates to cross-link adhesins thereby potentially limiting access by the protease component or indeed increasing coaggregation. Alternatively, the reduced activity of S156C-**3**,**4** towards fimA may be

due to a combination of increased distance of this divalent motif from the protease active site in S156C-**3**,**4** as compared to S156C-**2b**, additional conformational entropic cost of binding such an extended, higher valency ligand and/or incorrect topology.

Optimal reduction in aggregation was obtained here for the display of only two carbohydrates (bi-antennary), despite the greater affinity determined in our in vitro studies using the Gal-binding lectin PNA for tetra-antennary glycoprotein structures: tetra-antennary structures are rarer in natural glycoproteins than bi-antennary¹² and this suggests that A. naeslundii preferentially might target this carbohydrate motif over more highly branched structures to increase its chances of success in infection. Alternatively, the differences in binding may well be due to a different specificity of lectins¹³ (PNA vs FimA) which despite sharing a common carbohydrate ligand might have a spatial specificity that is dependent on carbohydrate density¹⁴ and quaternary structure. This is supported by suggestions that the A. naeslundii Gal-binding adhesin is functionally active as a dimer¹⁰ (the expression of recombinant A. naeslundii fimA has not yet been realised and is an important current goal to explore its behaviour its behaviour in ELLA assays), whereas, the Gal-binding lectin PNA is functionally active as a tetramer.¹⁵ Molecular mechanics analysis suggests that S156C-4 is capable of bridging carbohydrate binding sites in PNA tetramer.¹⁶ It cannot, however, be discounted that the glycoprotein mimic structures described here may be too short and therefore might alternatively or additionally operate through aggregative processes.¹⁷ Unfortunately, little is known about critical distances between binding sites in A. naeslundii FimA.

Nonetheless, the fact that we have determined different optimal glycodendriprotein ligands for these two different proteins highlights the need for a flexible chemical synthesis, such as that adopted here, that in turn allows flexible glycoprotein mimicry; what may be the best ligand for one lectin may not be for another and general methods are therefore needed to determine glycoprotein-like ligands on a case-by-case basis. The rationale for the design of the synthetic analogues of glycoproteins stems from biomimesis of the inherent structural similarity of the core, but not the tips, of the branched glycan structures in naturally-occurring glycoproteins. All *N*-linked glycoproteins contain the core pentasaccharide Man $\alpha(1,6)$ [Man $\alpha(1,3)$]Man $\beta(1,4)$ GlcNAc $\beta(1,4)$ GlcNAc (Additional Figure 1). Since, it is typically the tips of the glycan

arms and not the core that are recognised by lectins we replaced this common central structure with one that effectively mimics it but that can be far more rapidly and readily constructed, **2b**. It should be possible to replace the common pentasaccharide core found in all *N*linked glycoproteins. The TREN core in **2b** goes someway towards this mimickry but can be more readily constructed. This works also demonstrates that biomimetic assays are not necessarily a good measure of lectin binding^{11,18,19} or of the efficacy of methods that exploit that binding.²⁰ The inadequacies of lectins as models of carbohydrate binding structures on intact organisms are also highlighted by the discrepancy in our studies between the multivalent requirement for lectin binding and for microbial aggregation inhibition.

Materials and Methods:

Glycodendrimer Glycosylating Reagent Syntheses. In brief, the TREN-trivalent amine core 6 was Boc-protected at one amine terminus and the chloroacetamide group introduced at the remaining two amine termini through treatment with chloroacetic anhydride to form a bivalent TREN-core 7. D-Galactose was introduced to the bivalent TREN-core through displacement of the chloride with 1-thio-β-D-galactopyranose to create a di-galactose-TREN sulfhydryl core 8. The di-galactose-TREN core was elaborated through treatment with CF₃COOH to remove the Boc group, and then reaction of the liberated amine terminus with the thiobutyrolactone to create digalactose-TREN-sulfhydryl 9. Tri-galactose-TREN core 10 was constructed through sequential reaction of bivalent TREN core 7 with 1-thio-β-D-galactose (Gal-SH) and then di-galactose-TREN-sulfhydryl 9. Tetra-galactose-TREN core was constructed through reaction of di-galactose-TREN-MTS core 11 with two di-galactose-TRENsulfhydryl 9 cores. The reactive MTS group was introduced to 2-bromoethyl β-Dgalactoside \rightarrow 1 through displacement of bromide with sodium methanethiosulfonate as described.²¹ Di-galactose aromatic reagent **2a** was created from mesitylene tribromide **5** through exhaustive displacement of bromide with sodium methanethiosulfonate and reaction of two MTS groups with Gal-SH to give 2a. The reactive MTS group was introduced to the di-galactose-TREN core $9 \rightarrow 2b$, the tri-galactose-TREN core $10 \rightarrow 3$ and tetra-galactose-TREN core $12 \rightarrow 4$ at the remaining Boc-protected terminus, through treatment with CF₃COOH to remove the Boc group, and then reaction of the liberated

amine terminus with the NHS-butyl-MTS linker reagent. While reagents 2a, 2b and 4 are symmetrical, this methodology also allowed the synthesis of an asymmetric triantennary reagent 3, designed to mimic the asymmetric carbohydrate display that is observed in tri-antennary *N*-linked glycoproteins.¹² It is noteworthy that this chemistry was conducted completely without the use of carbohydrate protecting groups. Finally as a control conjugate, use of 1-thioglucopyranose (Glc-SH) in an essentially analogous way to the use of Gal-SH also allowed the synthesis from 7 of a Glc-analogue of 2b, 2b(Glc).

2 a : 1, 3 - b i s (5 - t h i o - β -D-galactopyranosyldisulfanylmethyl)-5methanethiosulfonatomethyl-2,4,6-trimethylbenzene: A solution of 1,3,5*tris*(methanethiosulfonatomethyl)-2,4,6-trimethylbenzene (0.238 g, 0.48 mmol) in DMF (20 ml) was cooled under N₂ to 0°C and the sodium salt of 1-thio- β -D-galactose (0.209 g, 0.96 mmol) in water (10 ml) was added over a period of 2h using a syringe pump. After warming to room temperature over night, the solvents were removed under reduced pressure, and the resulting yellow oil was purified by flash chromatography (SiO₂, CHCl₃ : MeOH : AcOH : H₂O, 60 : 30 : 3 : 5) to afford 1,3-bis(5-thio- β -Dgalactopyranosyldisulfanylmethyl)-5-methanethiosulfonatomethyl-2,4,6-

trimethylbenzene (**2a**; 0.193g, 55%) as a white solid, mp 118-120°C; $[\alpha]^{24}_{D}$ = -62 (c=0.59 in MeOH); IR (KBr) v=3421 (OH) 1635, 1559 (Ar), 1311, 1080 (S-SO₂) cm⁻¹; ¹H NMR (500 MHz, CD₃OD): δ 2.53 (s, 6H, CH₃), 2.59 (s, 3H, CH₃), 3.35 (s, 3H, SO₂CH₃), 3.56 (dd, 2H, *J* 9.3 and 3.3 Hz, 2 × H-3'), 3.62 (t, 2H, *J* 6.1 Hz, 2 × H-5'), 3.77 (dd, 2H, *J* 11.3 and 5.6 Hz, 2 × H-6a'), 3.82 (dd, 2H, *J* 11.4 and 6.5 Hz, 2 × H-6b'), 3.93-3.97 (m, 4H, 2 × H-2'/4'), 4.27 (d, 2H, *J* 11.4 Hz, 2 × CHSS-Gal), 4.35 (d, 2H, *J* 11.5 Hz, 2 × CHSS-Gal), 4.40 (d, 2H, *J* 9.4 Hz, 2 × H-1'), 4.58 (s, 2H, CH₂SSO₂); ¹³C NMR (125 MHz, CD₃OD): δ 16.7 (CH₃), 17.3 (CH₃), 37.5 (CH₂SSO₂), 42.4 (CH₂SS-Gal), 50.2 (SO₂CH₃), 62.8 (C-6'), 69.8, 70.5, 76.2 (C-3'), 81.1 (C-2',3',4',5'), 92.9 (C-1'), 129.4, 133.4, 138.6, 139.7 (Ar) ; MS (TOF ES+) *m/z* Found 747.0736 ([M+Na]⁺), C₂₅H₄₀O₁₂S₆ requires 747.0742.

galactopyranosyl)ethanoyl]amino}ethyl)(2-aminoethyl)amine (203 mg, 0.33 mmol) and 5-methanesulfonylsulfanylpentanoic acid 2,5-dioxopyrrolidin-1-yl ester (NHS-butyl-MTS) (101 mg, 0.33 mmol) were stirred in anhydrous DMF (10 mL) at room temperature for one hour. The solution was concentrated in vacuo and the residue was purified flash silica column chromatography, eluting with bv chloroform/methanol/water (60:30:5), to give the title compound (217 mg, 81%) as a pale yellow gum: $[\alpha]^{23}_{D} = -23.1$ (c=1.0 in H₂O); IR (KBr): v=3407 (OH), 1637 (amide I), 1543 (amide II) 1310, 1129 (S-SO₂) cm⁻¹; ¹H NMR (400 MHz, D₂O): δ=1.54-1.69 (m, 4H, (CH₂)₂CH₂S), 2.17 (t, J=7.1 Hz, 2H, COCH₂CH₂), 2.61 (br s, 6H, (CH₂)₃N), 3.12 (t, J=7.0 Hz, 2H, CH₂SSO₂), 3.16-3.24 (m, 6H, CH₂NHCO), 3.27 (d, J=15.3 Hz, 2H, COCHH'S), 3.37 (s, 3H, SO₂CH₃), 3.40 (d, J=15.3 Hz, 2H, COCHH'S), 3.44 (dd, J=9.6, 9.5 Hz, 2H, H2`), 3.50 (dd, J=9.5, 3.2 Hz, 2H, H3`), 3.53-3.63 (m, 6H, H5`, H6`), 3.83 (d, J=3.2 Hz, 2H, H4`), 4.34 (d, J=9.6 Hz, 2H, H1`); ¹³C NMR (125 MHz, D_2O): $\delta = 24.1$, 28.5 ((<u>CH</u>₂)₂CH₂S), 33.3 (NHCO<u>C</u>H₂S), 35.0 (CO<u>C</u>H₂CH₂), 35.7 (CH₂SSO₂), 36.5 (br, CH₂NHCO), 49.8 (SO₂CH₃), 52.9 (br, N(CH₂)₃), 61.3 (C6[`]), 68.9 (C4`), 69.7 (C2`), 74.0 (C3`), 79.3 (C5`), 85.7 (C1`), 173.3 (NHCOCH₂S), 177.0 $(NHCO(CH_2)_4)$; HRMS-ES: (m/z): $[M+H]^+$ calcd for C₂₈H₅₂N₄O₁₅S₄ 813.2390; found 813.2377.

3 : 4 - { 2 - [(2 - { [5 - (4 - { 2 - [B i s (2 - { [2 - (1 - t h i o - $\beta - D - galactopyranosyl)ethanoyl]amino}ethyl)amino]ethylcarbamoyl}propylsulfanyl)eth anoyl]amino}ethyl)(2-{[2-(1-thio-<math>\beta$ -D-

galactopyranosyl)ethanoyl]amino}ethyl)amino]ethylcarbamoyl}butyl

methanethiosulfonate: $(2 - \{ [5 - (4 - \{2 - [Bis(2 - \{[2 - (1 - thio - \beta - D - galactopyranosyl)ethanoyl]amino\}ethyl)amino]ethylcarbamoyl}propylsulfanyl)ethanoyl]amino}ethyl)(2 - {[2 - (1 - thio - \beta - D - galactopyranosyl)ethanoyl]amino}ethyl)(2 - tert-$

butoxycarbonylaminoethyl)amine (77 mg, 0.062 mmol) was stirred in dichloromethane (3 mL) and trifluoroacetic acid (3 mL) at room temperature for 2 hours. The resulting solution was concentrated in vacuo to give the trifluoroacetate salt, which was dissolved in DMF (5 mL). Diisopropylethylamine (108 μ L, 0.62 mmol) and 5-methanesulfonylsulfanylpentanoic acid 2,5-dioxopyrrolidin-1-yl ester (NHS-butyl-MTS) (19 mg, 0.062 mmol) were added and the resulting solution stirred for 1 hour and concentrated in vacuo. The residue was purified by flash silica column chromatography,

eluting with chloroform/methanol/water (60:45:10) to give a colourless foam (73 mg, 87% over 2 steps): $[\alpha]^{23}_{D} = -29$ (c=0.5 in H₂O); IR (KBr): v=3383 (OH), 1642 (amide I), 1549 (amide II), 1310, 1131 (S-SO₂) cm⁻¹; ¹H NMR (500 MHz, D₂O): ¹H NMR (500 MHz, D₂O): δ=1.55-1.68 (m, 4H, COCH₂(CH₂)₂CH₂SSO₂CH₃), 1.73 (tt, J=7.3, 7.2 Hz, 2H, COCH₂CH₂CH₂S), 2.19 (t, J=7.1 Hz, 2H, COCH₂(CH₂)₂CH₂SSO₂CH₃), 2.23 (t, J=7.1 Hz, 2H, COCH₂(CH₂)₂S), 2.47 (t, 2H, J=7.2 Hz, CO(CH₂)₂CH₂S), 2.60 (br s, 12H, (CH₂)₃N), 3.04 (t, J=6.1 Hz, 2H, CH₂NHBoc), 3.11 (t, 2H, J=7.0 Hz, COCH₂(CH₂)₂CH₂SSO₂CH₃), 3.16-3.22 (br m, 12H, CH₂NHCO), 3.28 (d, J=15.3 Hz, 3H, COCHH'S), 3.35 (s, 3H, SSO₂CH₃), 3.40 (d, J=15.3 Hz, 3H, COCHH'S), 3.44 (dd, J=9.5, 9.8 Hz, 3H, H2`), 3.51 (dd, J=9.5, 3.3 Hz, 3H, H3`), 3.52-3.62 (m, 9H, H5`, H6`), 3.84 (d, J=3.1 Hz, 3H, H4`), 4.38 (d, J=9.5 Hz, 3H, H1`); ¹³C NMR (125 MHz, D₂O): δ= 24.1, 28.3 (NHCOCH₂(<u>CH</u>₂)₂CH₂SSO₂CH₃), 25.0 (COCH₂<u>C</u>H₂CH₂S), 31.5 (CH₂SCH₂CO), 33.2 (NHCOCH₂(CH₂)₂CH₂SSO₂CH₃), 33.3, 33.4 (NHCOCH₂S), 34.7, 34.9 (COCH₂CH₂CH₂S), 35.6 (NHCOCH₂(CH₂)₂CH₂SSO₂CH₃), 37.0, 37.6, 37.7 (br, <u>CH</u>₂NHCO), 50.1 (SSO₂<u>C</u>H₃), 52.3, 52.4, 52.4, 52.9 (br, N(<u>C</u>H₂)₃), 61.2, 61.5 (C6[•]), 69.1 (C4'), 69.7 (C2'), 74.0 (C3'), 79.3 (C5'), 85.5, 85.7 (C1'), 172.6, 172.7 (NHCOCH₂SGal), 172.8, 176.0, 177.2 (NHCOCH₂); MS-ES (*m/z*) (%): 1358 (3) $[M+Na]^{+}$, 1336 (5) $[M+H]^{+}$, 691 (10) $[M+2Na]^{2+}$, 669 (60) $[M+2H]^{2+}$; $[M+2H]^{2+}$ Isotope ratios calcd for $C_{48}H_{88}N_8O_{23}S_6$ (0:+0.5:+1.0:+1.5:+2.0:+2.5) 100:63:51:24:11:4; found 100:83:62:33:15:7.

4: 4-{2-[Bis(2-{[5-(4-{2-[bis(2-{[2-(1-thio-β-D-

$galactopyranosyl)ethanoyl]amino\}ethyl)amino]ethylcarbamoyl\}butyldisulfanyl)pe ntanoyl]amino]ethylcarbamoyl}butyl methanethiosulfonate: Bis(2- {[5-(4-{2-[bis(2-{[2-(1-thio-<math>\beta-D-$

galactopyranosyl)ethanoyl]amino}ethyl)amino]ethylcarbamoyl}butyldisulfanyl)pentano yl]amino}ethyl)(2-*tert*-butoxycarbonylaminoethyl)amine (73 mg, 0.038 mmol) was stirred in dichloromethane (2 mL) and trifluoroacetic acid (2 mL) at room temperature for 1 hour. The resulting solution was concentrated in vacuo to give the trifluoroacetate salt. 0.1 M HCl (aq.) was added to give the hydrochloride salt. This solution was concentrated in vacuo, toluene added, again concentrated in vacuo and then the residue dissolved in DMF (3 mL). Triethylamine (106 μ L, 0.38 mmol) and 5methanesulfonylsulfanylpentanoic acid 2,5-dioxopyrrolidin-1-yl ester (NHS-butyl-

MTS) (12 mg, 0.038 mmol) were added and the resulting solution stirred for 1 hour and concentrated in vacuo. The residue was purified by flash silica column chromatography, eluting with chloroform/methanol/water (60:45:10) to give a colourless foam (51 mg, 67% over 2 steps): $[\alpha]_D = -36$ (c=0.4 in H₂O); IR (KBr): v=3415 (OH), 1658 (amide I), 1539 (amide II), 1310, 1130 (S-SO₂) cm⁻¹; ¹H NMR (500 MHz, D₂O): δ =1.60 (br s, 12H, SCH₂(CH₂)₂CH₂CO, COCH₂(CH₂)₂CH₂SSO₂CH₃), 1.89 (tt, J=7.3, 7.2 Hz, 4H, COCH₂CH₂CH₂S), 2.20 (br s, 4H, S(CH₂)₃CH₂CO), 2.23 (t, J=7.1 Hz, 2H, COCH₂(CH₂)₂CH₂SSO₂CH₃), 2.26 (t, J=7.2 Hz, 4H, COCH₂(CH₂)₂S), 2.57 (t, 4H, J=7.2 Hz, CO(CH₂)₂CH₂S), 2.62 (t, J=7.2 Hz, 4H, SCH₂(CH₂)₃CO), 2.67-2.90 (br m, 18H, (CH₂)₃N), 3.10 (t, 2H, J=7.0 Hz, COCH₂(CH₂)₂CH₂SSO₂CH₃), 3.12-3.30 (br m, 18H, CH₂NHCO), 3.31 (d, J=15.3 Hz, 4H, COCHH`S), 3.34 (s, 3H, SSO₂CH₃), 3.43 (d, J=15.3 Hz, 4H, COCHH S), 3.46 (dd, J=9.5, 9.8 Hz, 4H, H2), 3.52 (dd, J=9.5, 3.3 Hz, 4H, H3`), 3.57-3.63 (m, 12H, H5`, H6`), 3.86 (d, J=3.3 Hz, 4H, H4`), 4.36 (d, J=9.5 Hz, 4H, H1`); ¹³C NMR (125 MHz, D₂O): δ=24.1, 24.3 (S(CH₂)₂CH₂CH₂CO, NHCOCH₂CH₂CH₂CH₂SSO₂CH₃), 24.8 (COCH₂CH₂CH₂S), 28.0, 28.2 (SCH₂<u>C</u>H₂(CH₂)₂CO, NHCOCH₂CH₂CH₂CH₂SSO₂CH₃), 33.2 (NHCOCH₂(CH₂)₂CH₂SSO₂CH₃), 33.5 (NHCOCH₂S), 34.4 (COCH₂(CH₂)₂S), 35.4 (S(CH₂)₃CH₂CO), 35.5 (NHCOCH₂(CH₂)₂CH₂SSO₂CH₃), 36.4, 37.0 (br, CH₂NHCO), 37.2 (CO(CH₂)₂<u>C</u>H₂SS), 37.7 (SS<u>C</u>H₂(CH₂)₃), 50.0 (SSO₂<u>C</u>H₃), 52.7, 52.9, 53.9 (br, N(CH₂)₃), 61.3 (C6[`]), 68.8 (C4[`]), 69.7 (C2[`]), 74.0 (C3[`]), 79.2 (C5[`]), 85.7 (C1[`]), 172.9 (NHCOCH₂S), 176.2, 177.0, 177.1 (NHCO(CH₂)₃ (CH₂)₄CONH, NH<u>C</u>O(CH₂)₄SSO₂CH₃); MS-ES (*m*/*z*) (%): 1026 (92) [M+2Na]²⁺, 1015 (89) $[M+H+Na]^{2+}$, 692 (70) $[M+3Na]^{3+}$, 685 (100) $[M+H+Na]^{3+}$; Isotope ratios calcd for $C_{74}H_{136}N_{12}O_{31}S_{10}$ 100:98:98:64:39:20; $[M+3Na]^{3+}$ (0:+0.33:+0.67:+1.00:+1.33:+1.67) found 100:96:95:56:30:12; [M+2Na]²⁺ (0:+0.5:+1.0:+1.5:+2.0:+2.5) found 100:97:97:60:29:11.

Glycoprotein construction (Typical Procedure). 35 mg of SBL S156C protein was dissolved in a 2.4 mL of buffer containing 70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5. Glycosylating reagent was dissolved in 100 mL of distilled water to make a 0.26 M solution and mixed with the protein solution, vortexed for 5 min. The coupling mixture was sealed and placed on an end-over-end rotator at room temperature. After 30 min, 10 mL of the reaction mixture was withdrawn and tested for residual free thiol using

Rendle et al

page S9

Ellman's reagent; ²² negative colour change indicated completion of coupling reaction. For purification, the reaction mixture was purified by G-25 Sephadex (Pharmacia, PD10) size-exclusion chromatography eluting with a buffer containing 5 mM MES, 2 mM CaCl₂, pH 6.5. The glycoprotein extract was dialysed (3 × 60 min) at 4 °C against buffer containing 10 mM MES, 2 mM CaCl₂, pH 5.8, flash frozen in liquid nitrogen and stored at -18°C. The product was analysed by Ellman's titration,²² PMSF titration to determine concentration, ²³ mass spectrometry and SDS-PAGE. In a control reaction treatment of SBL WT, which contains no cysteine, did not result in glycosylation. Michaelis-Menten determination of amidase activity using suc-AAPF-*p*NA as a substrate was performed according to the initial rates method. The k_{cat}/K_M (s⁻¹mM⁻¹) values for suc-AAPF-*p*NA hydrolysis (pH 8.6, 0.1M Tris, 25° C) were: WT-SBL 209.0±15; S156C 147.1±11.2; S156C-1 89.7±4.4; S156C-2a 111.8±8.9; S156C-2b 94.5±7.0; S156C-2b(Glc) 106.3±9.7; S156C-3 100.2±7.4; S156C-4 108±9.1. PMSF-treated enzymes showed no measurable activity.

Electrospray mass spectroscopy of glycoproteins. Glycoprotein solution was diluted to 200 pmol/µL with water/acetonitrile/formic acid (95:5:0.5 v/v). A 10 µL aliquot was injected at 4 °C via a reverse phase chromatography column (Jupiter 5µ C5 300Å 150 x 2.00 mm) at 0.2 mL/min flow rate with solvent ramping from water/acetonitrile/formic acid 95:5:0.5 \rightarrow 0:100:0.5 and UV monitoring at 280 nm (Waters 2790 auto-sampler) into a Micromass LCT mass spectrometer (sample cone at 35 V). Measurements after MaxEntTM deconvolution (to a precision of ± 0.2 Da) were within ± 2.2 Da: SBL calcd 26698.5, found 26698.8; S156C calcd 26714.6, found 26715.0; S156C-**2a** calcd 27358.6, found 27360.0; S156C-**2b** calcd 27447.5, found 27448.0; S156C-**2b**(Glc) calcd 27447.5, found 27449.0; S156C-**3** calcd 27972.2, found 27970.0; S156C-**4** calcd 28645.1, found 28645.0.

Competitive Enzyme-Linked Lectin Assays (ELLA). All incubations were carried out at 37°C. Microtiter plates were coated overnight with peanut agglutinin (PNA) solution (100 μ L of a 0.1mg/mL solution in phosphate buffered saline pH 7.2 (PBS)). Excess protein was removed by washing (3 × 200 μ L of PBS containing 0.05% v/v Tween (PBST)). Coated wells were blocked with bovine serum albumin solution (150 μ L of a 1% w/v solution in PBS (PBST)) for 1h. During this period, serial (× 2)

dilutions of methyl β -D-galactopyranoside were mixed with glycodendriprotease (at a dilution which had been determined to give A₄₁₀ of 0.8-1.0 after 20 min). These solutions were then transferred to the blocked plate for a 1 h incubation. These plates were subsequently washed (2 × PBST, 1 × PBS) prior to development with suc-AAPF*p*NA (50 µL of a 10 mM solution in 0.1M Tris, pH 8.6). The inherent peptidase activity of SBL was therefore used as a readout. The absorbance at 410 nm was monitored after a 20 min period. Binding constants are based on A₄₁₀ in the absence of methyl β-D-galactopyranoside and referenced to the inhibitory potential of an methyl β-D-galactopyranoside determined for each plate and assuming a $K_D = 5.34 \times 10^{-4}.^{24}$ All assays were carried out in duplicate. Values are shown based on ligand concentration and it may be that the small variations in binding levels between S156C-1, **2a** and **2b** may be accounted for simply by differing concentrations of sugar epitope.

Co-aggregation analysis of pathogen adhesion inhibition. Streptococcus oralis 34, Actinomyces naeslundii strains T14V were purchased from the American Type Culture Collection, grown as static cultures in Todd-Hewitt broth (BBL Microbiology Systems, Cockeysville, MD) incubated at 37° C in 5% CO₂ and H₂:CO₂:N₂ (10:10:80) atmospheres, respectively and harvested in mid to late exponential growth phase by centrifugation (10, 000 × g, 10 min, 4 °C). A. naeslundii cells were washed twice and resuspended to $A_{540} = 1.0$ in coaggregation buffer (CB: 1 mM tris (hydroxymethyl)aminomethane, 10⁻⁴ M CaCl₂, 10⁻⁴ M MgCl₂, 0.15 M NaCl, pH 7.0,) and incubated for 30 min at 37 °C with inhibitor. Control experiments using no inhibitor, or S156C-2b + PMSF (~0.1 M) or PMSF alone (~0.1 M) were also conducted; in these experiments zero activity towards suc-AAPF-pNA was detected either before or after assay. After incubation an equal volume of S. oralis was added and vortexed for 10 s. Coaggregation was assessed by turbidity changes at 540 nm over 10 min. The possibility that enzymes also beneficially cleave factors on the surface of S. oralis, after initiation of coaggregation cannot be discounted; however the sugar dependency, the lack of incubation time for such a process and the observation of instant inhibitory effects suggest that action on A. naeslundii is the most likely source of inhibition.



Additional Figure 1 The bi-, tri- and tetra-antennary "glycodendrimer" structures constructed through chemical synthesis and used to glycosylate the protein were designed to mimic the branched display of carbohydrates found in naturally-occuring glycoproteins (illustrated top by a naturally-occurring Gal-terminated *N*-linked biantennary structure). All of these systems displayed Gal-terminated antennae designed to bind to the Gal-specific adhesin found in the target pathogen *A. naeslundii*.



Additional Figure 2 Mass spectrometric characterization of glycoproteins. ESMS analysis confirmed the formation, with correct mass, of the expected glycoproteins S156C-2a, 2b, 3, 4 as major products. It should be noted that S156C-3,4 show lower signal-to-noise ratios due to lower total ion counts (TIC) during mass spectrometry due to different ionization properties of these proteins.



Additional Figure 3 (a) Percent aggregation as compared with control of the pathogen *A. naeslundii* to *S. oralis* (after 10 min) after treatment with glycoprotein enzyme ($1.5\mu g/mL$, 30 min). Most potent inhibition of the binding of *A. naeslundii* was found following treatment with the biantennary synthetic glycoprotein S156C-2b (IC₅₀)

20nM) and is $>10^6$ more potent than that of small molecule inhibitor lactose (IC₅₀ 33mM) (b) Visual comparison of aggregation assays control (left) and treated with 15µg/mL S156C-2b (right). (c) Aggregation of *A. naeslundii* was assessed by turbidity changes at 540 nm over 10 min: decreased turbidity indicating increased aggregation. Increasing doses of inhibitor and analysis of inhibitory response reveal that S156C-2b (black, solid lines) is a more powerful inhibitor than SBL-WT (red, solid lines) or proteolytically inactive PMSF treated S156C-2b (S156C-2b + PMSF, black circles, broken line). Two key properties of hybrid glycoprotein S156C-2b are thus essential for optimal inhibition: its multiantennary carbohydrate display and its protein degrading activity.



Additional Figure 4 Representative Sigmoidal (Variable Slope) Dose Response Curve for Inhibition of Bacterial Aggregation

References for Supporting Information

(1) Nagahori, N.; Lee, R. T.; Nishimura, S.-I.; Page, D.; Roy, R.; Lee, Y. C. *ChemBioChem* **2002**, *3*, 836-844.

(2) Sears, P.; Wong, C.-H. *Science* **2001**, *291*, 2344-2350.

(3) Paulsen, H.; Heume, M.; Gyorgydeak, Z.; Lebuhn, R. *Carbohydr. Res.* **1985**, *144*, 57-70.

(4) Yamazaki, F.; Nukada, T.; Ito, Y.; Sato, S.; Ogawa, T. *Tetrahedron Lett.* **1989**, *30*, 4417-20.

(5) Prahl, I.; Unverzagt, C. *Tetrahedron Lett.* **2000**, *41*, 10189-10193.

(6) Ratner, D. M.; Plante, O. J.; Seeberger, P. H. *Abstracts of Papers,* 224th ACS National Meeting, Boston, MA, United States, August 18-22, 2002 **2002**, CARB-092.

(7) Marchese, A.; Schito, G. C.; Debbia, E. A. *J. Chemother.* **2000**, *12*, 459-462.

(8) Alksne Lefa, E. *Exp. Opin. Investigational Drugs* **2002**, *11*, 1149-59.

(9) Ruhl, S.; Cisar, J. O.; Sandberg, A. L. *Infect. Immun.* **2000**, *68*, 6346.

(10) Hallberg, K.; Holem, C.; Öhman, U.; N., S. *Infect. Immun.* **1998**, 66, 4403.

(11) Dimick, S. M.; Powell, S. C.; McMahon, S. A.; Moothoo, D. N.;

Naismith, J. H.; Toone, E. J. J. Am. Chem. Soc. 1999, 121, 10286.

(12) Dwek, R. A. Chem. Rev. **1996**, 96, 683-720.

(13) Page, D.; Roy, R. *Bioconj. Chem.* **1997**, *8*, 714-23.

(14) Horan, N.; Yan, L.; Isobe, H.; Whitesides, G. M.; Kahne, D. *Proc. Natl Acad. Sci, USA* **1999**, *96*, 11782–11786.

(15) Reddy, G. B.; Bharadwaj, S.; Surolia, A. *Biochemistry* **1999**, *38*, 4464-4470.

(16) MOPAC minimization using the maximization of C-4 to C-4 distances in 2b and 4 gave "tip-to-tip" C-4-to-C-4 distances for these biantennerary and tetraantennerary structures of 23.5 and 59.6 Å, respectively. Measurement in the X-ray structure of PNA cocrystallized with a Gal ligand (RCMB-code 1ciw) reveals corresponding C-4-to-C-4 distances of 48.1, 57.4 and 61.9 Å suggesting that S156C-4 is indeed capable of multi-point attachment mode binding.

(17) In this case KD values would not directly relate to the microscopic affinity of protein for ligand and might be more correctly described as global or apparent avidity.

(18) Corbell, J. B.; Lundquist, J. J.; Toone, E. J. *Tetrahedron: Asymm.* **2000**, *11*, 95.

(19) Lundquist, J. J.; Toone, E. J. Chem. Rev. 2002, 102, 555.

(20) As Toone, Naismith and coworkers have highlighted in Ref (5, main text) discrepancies in affinities might also arise from the nature of the assay used to determine them. Our results here further highlight the importance of assays employing a directly relevant functional readout.

(21) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614-9615.

(22) Ellman, G. L. Arch. Biochem. Biophys. 1958, 74, 443-50.

(23) Hsia, C. Y.; Ganshaw, G.; Paech, C.; Murray, C. J. *Anal. Biochem.* **1996**, *242*, 221-227.

(24) Reddy, G. B.; Srinivas, V. R.; Ahmad, N.; Surolia, A. *J. Biol. Chem.* **1999**, *274*, 4500-4503.