Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 3188

www.rsc.org/obc

PAPER

Stereoselective synthesis of light-activatable perfluorophenylazide-conjugated carbohydrates for glycoarray fabrication and evaluation of structural effects on protein binding by SPR imaging†

Lingquan Deng,[a] Oscar Norberg,[a] Suji Uppalapati,[b] Mingdi Yan*[a,b] and Olof Ramström*[a]

Received 7th January 2011, Accepted 14th February 2011
DOI: 10.1039/c1ob05040k

A series of light-activatable perfluorophenylazide (PFPA)-conjugated carbohydrate structures have been synthesized and applied to glycoarray fabrication. The glycoconjugates were structurally varied with respect to anomeric attachment, *S*-, and *O*-linked carbohydrates, respectively, as well as linker structure and length. Efficient stereoselective synthetic routes were developed, leading to the formation of the PFPA-conjugated structures in good yields over few steps. The use of glycosyl thiols as donors proved especially efficient and provided the final compounds in up to 70% total yield with high anomeric purities. PFPA-based photochemistry was subsequently used to generate carbohydrate arrays on a polymeric surface, and surface plasmon resonance imaging (SPRi) was applied for evaluation of carbohydrate-protein interactions using the plant lectin Concanavalin A (Con A) as a probe. The results indicate better performance and equal efficiency of *S*- and *O*-linked structures with intermediate linker length.

Introduction

Carbohydrates and carbohydrate-containing biomolecules are ubiquitous in nature. These compounds form a basis in the generation of a vast variety of biological complexities inherent in the development, growth, and function of intact organisms, diversifying the effects of the relatively small number of genes in a typical genome. Due to essential roles of glycans and glycoconjugates in biological processes, considerable efforts are being made to decipher the many mechanisms and effects of the glycocode.[1–3] Indeed, glycans modulate or mediate various fundamental processes in cell-cell-, cell-matrix-, and cell-molecule interactions, and also act as regulatory switches within the nucleus and cytoplasm. In most cases, involved glycans are attached to proteins or lipids, and these conjugates have proven to be major information carriers between cells and their surroundings. Thus, in order to further understand these biological processes, detailed elucidation of carbohydrate-protein interactions is important. New analytical methods and techniques have more recently been developed to meet this objective, among which a series of solid-supported techniques, including microarrays, quartz crystal microbalance (QCM), surface plasmon resonance (SPR), scanning probe microscopy (SPM) and nanoobject-based systems have been recognized.[4–7]

Many carbohydrate-mediated recognition events occur on cell surfaces. The presentation of glycans at the cell surfaces with respect to, *e.g.*, density, distribution, orientation, and linker length and flexibility strongly influences the recognition specificity. This feature underscores the importance of studying carbohydrate-protein interactions at surfaces rather than in solution. Furthermore, multivalent display and cooperative recognition, which greatly enhance the binding strength, can be achieved by surface supported techniques,[8,9] thus making them particularly well suited for studies of carbohydrate-protein interactions. A variety of methods have been developed to achieve the immobilization of carbohydrates on surfaces so that subsequent binding studies were possible. Of these, photochemical immobilization has proven to be a very robust and versatile method in this pursuit, either by using a pre-immobilized layer of photoreactive groups on the surface to further bind underivatized sugars, or by introducing a photoreactive tag onto carbohydrates to allow one-step functionalization under UV irradiation.[10–19] In the latter case, PFPA-based photochemistry,[20,21] has been utilized in our laboratories to enable rapid, covalent and nondestructive immobilization of carbohydrates onto solid supports, resulting in the development of new methodologies in microarray and QCM biosensor systems.[15–17]

Glycoarray technology, which permits screening of a number of biological interactions in parallel, is a very promising technique that has received much attention ever since its introduction into the scientific community.[22–29] In general, analysis of such arrays relies

[a]*Department of Chemistry, KTH - Royal Institute of Technology, Teknikringen 30, S-10044, Stockholm, Sweden. E-mail: ramstrom@kth.se; Fax: + 46 8 7912333*

[b]*Department of Chemistry, Portland State University, P.O. Box 751, Portland, Oregon, 97207-0751, USA. E-mail: yanm@pdx.edu; Fax: + 1 503 725-9525*

† Electronic supplementary information (ESI) available: NMR spectra of all new compounds, primary SPRi data and corresponding experimental description. See DOI: 10.1039/c1ob05040k

on various protein labeling formats in a step-wise batch mode. However, surface plasmon resonance imaging is a technique that can combine the advantages of glycoarray technology with the characteristic merits of the SPR technique to evaluate biological interactions on surfaces in a label-free manner.[30–35] This also enables the efficient analysis of multiple complex interactions in real-time. These features are attractive, leading to a recent upsurge of interest in exploring carbohydrate-based SPR.[27] On the other hand, since biomolecular interactions are delicately affected by the chemistry and topography of the generated surfaces.[36] Tailoring and evaluating the surface chemistry, *e.g.*, to enhance the specific binding and to reduce the nonspecific binding, is therefore also of fundamental importance in glycoarray and sensor development. However, systematic investigations of varied surface chemical structures and their effects on target biological interactions on such platforms are generally lacking.

In the present study, this has been addressed by evaluating the lectin binding performances of a number of light-activatable, PFPA-conjugated carbohydrate structures on surfaces. The glycoconjugates were varied with respect to glycosidic bond type (*S*-, and *O*-glycosides, respectively), linker length and linker structure, in order to probe the sensitivity and specificity of the photogenerated surface system, and to provide optimal surface structures for protein binding studies using such a platform. To take full advantage of the glycoarray technology for the study of carbohydrate-protein interactions, carbohydrate immobilization strategies are of vital importance. Since covalent immobilization methods offer more stable arrays than non-covalent alternatives, the former strategy has drawn greater attention. This strategy however requires efficient and straight-forward synthetic methods, resulting in robust functionalized surfaces. Herein, this challenge has also been addressed, where several synthetic methods were developed for efficient and stereo-controlled synthesis of the final PFPA-carbohydrate conjugates. Using these probes, photo-induced generation of glycoarrays and subsequent evaluation of biological interactions was subsequently addressed, and surface plasmon resonance imaging was used to combine microarray and imaging analysis for the multiplex examination of the binding capacities.

Results and discussion

Development of synthetic pathways

Six glycoconjugate structures (Fig. 1) were designed for the study of carbohydrate-protein interactions on the glycoarray. Two carbohydrate head-groups, D-mannose (Man) and D-galactose (Gal) were chosen as model carbohydrates, linked to the photo-activatable perfluorophenylazide moiety *via* different linkers and different glycosidic bonds. Structure **1** composed a mannose head-group linked by an *O*-glycosyl bond to a triethylene glycol (TEG) based linker, which was further coupled to the PFPA moiety. In comparison, compound **2** was linked through an *S*-glycosyl bond while all other features were the same as for compound **1**. Compounds **2–4** were all *S*-linked glycoconjugates, however differing in linker lengths from three ethylene glycol (EG) units to nine EG units. Compound **5** was designed to possess a triazole ring in the linker part with a linker length comparable to that of compound **3**. Finally, compound **6** was galactose-based, different from the other five structures, used as a reference compound in the study.

Four different synthetic pathways for the preparation of the target glycoconjugates were designed. Initially, the glycoconjugates **1** and **6** were synthesized following previously established synthetic routes.[17] Briefly, this involved anomeric coupling of the peracetylated parent donor to an azido-alcohol acceptor under Lewis acid activation. This product was subsequently reduced, subjected to ion exchange, and coupled to 4-azido-2,3,5,6-tetrafluorobenzoic acid using EDCI/triethylamine to give the acetyl-protected PFPA-conjugate. Deprotection under basic

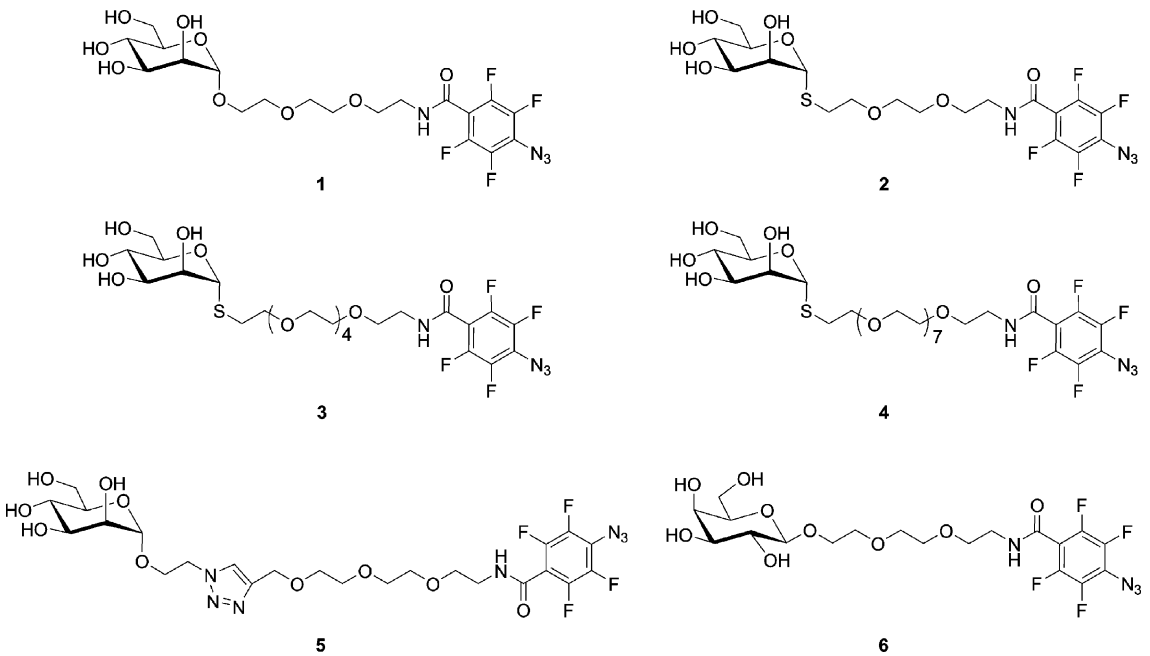


Fig. 1 PFPA-conjugated carbohydrate structures evaluated.

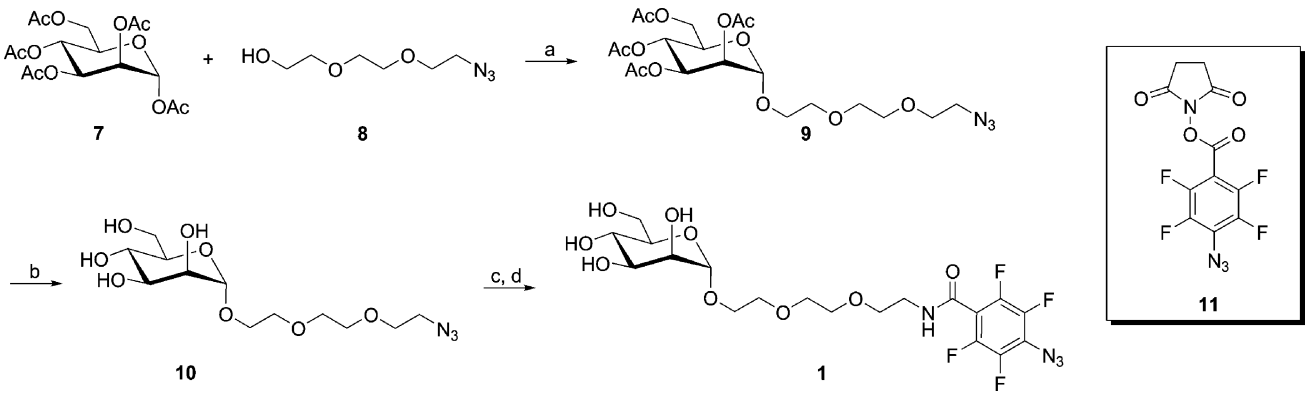
conditions yielded the final product. This synthetic pathway being straightforward, the total yield was, however, low. To address this, a second route was instead designed (Scheme 1). In this case, the acetylated carbohydrate-azide **9** was deprotected prior to amide formation. This manipulation enabled the subsequent reduction of azide **10** without formation of any side products. Final conjugation using NHS-PFPA **11**[15] instead of EDCI coupling resulted in good product yields.

This pathway, although considerably more efficient, still involved a protection-glycosylation-deprotection route. On the other hand, glycosylation using unprotected carbohydrates is more desirable as it reduces the number of synthetic steps and could enhance reactivity.[37] In the second pathway, free carbohydrate-azide **10** (Scheme 1) was employed for the final PFPA-coupling, which prompted attempts of using unprotected carbohydrate structures in the entire synthetic route. Taking advantage of the fact that sulfur is less basic and more nucleophilic than oxygen, glycosyl thiols are increasingly used in the synthesis of thiooligosaccharides, thioglycopeptides and thioglycoproteins.[38–45]

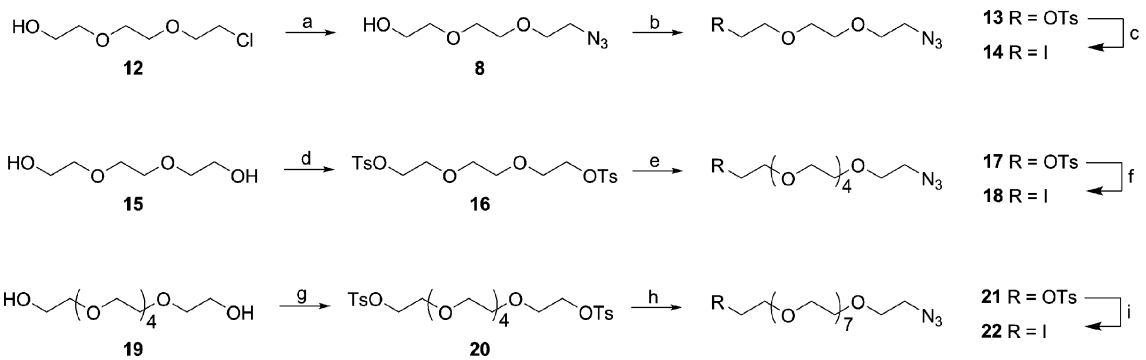
The use of glycosyl thiols was thus adopted in a more efficient, yet simpler, third synthetic route. In order to realize the third route, however, a series of different linker structures were first synthesized (Scheme 2). These were based on the oligoethylene glycol (OEG) motif, widely used in a variety of biotechnical and biomedical applications.[46] These were primarily chosen because of their water solubility, stability, and ability to reduce nonspecific binding of proteins and other biomolecules. However, the integration of ethylene glycols with defined length into other synthetic molecules requires a selective functionalization and prolongation of symmetrical diols/diol derivatives. In this study, an iterative method to produce different bifunctional and monodispersed OEGs with up to 9 units was developed (Scheme 2).

For the shortest TEG derivative **14**, the commercially available chloride **12** was chosen as starting material. The TEG azide **8** was prepared quantitatively from the chloride **12** by using sodium azide in DMF at 90 °C overnight. Tosylation at the other end of the azide **8** was achieved in good yield according to an established method.[47] S_N2 substitution of the tosylate **13** was subsequently achieved by use of sodium iodide in DMF at 50 °C for 5 h and the desired iodide **14** was obtained in a yield of 82%. To make longer OEG derivatives, OEG diols **15** and **19** were firstly transformed into OEG ditosylates **16** and **20**, respectively. The ditosylates were subsequently prolonged by nucleophilic substitution with the deprotonated azide **8** at only one end of the molecule to afford tosylates **17** and **21** in good yields. The final iodides **18** and **22** were obtained following similar conditions as for **14**.

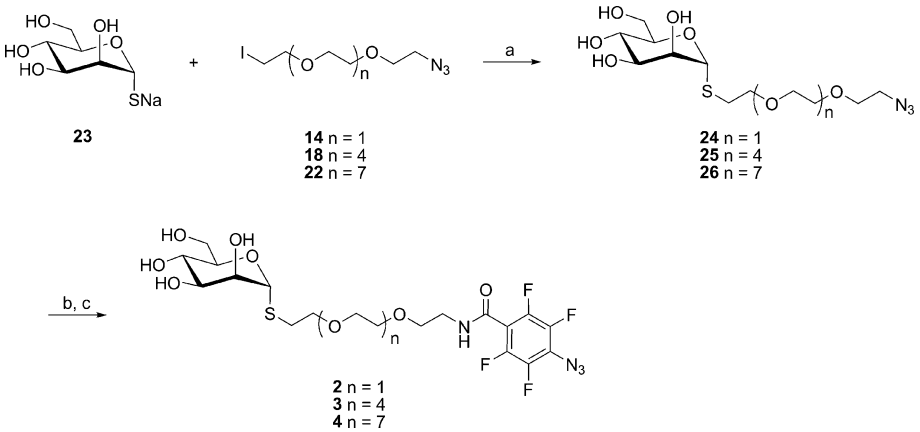
The glycosyl thiol pathway was next addressed (Scheme 3). Mannose thiolate **23**[48] was thus used as donor and the *S*-linked glycoconjugates **24–26** were conveniently obtained from S_N2 coupling of the free thiol sugar **23** and the iodide linker molecules **14**, **18** and **22**. Subsequent reduction and PFPA coupling using the methods mentioned above provided the final *S*-linked glycoconjugates **2**, **3** and **4** in total yields up to around 70%. This synthetic pathway proved highly efficient, requiring few steps



Scheme 1 a) $BF_3{\cdot}Et_2O$, DCM, 0 °C-RT, 24 h (60%); b) NaOMe, MeOH, 1 h (quant.); c) H_2, Pd/C, MeOH, 1 h (quant.); d) **11**, DMF, dark, 2.5 h (62%).



Scheme 2 a) NaN_3, DMF, 90 °C, overnight (quant.); b) TsCl, TEA, DMAP, THF, 0 °C-RT, 12 h (78%); c) NaI, DMF, 50 °C, 5 h (82%); d) TsCl, TEA, DMAP, THF, 0 °C–RT, overnight (81%); e) **8**, NaH, DMF, 1.5 h (59%); f) NaI, DMF, 50 °C, 5 h (81%); g) TsCl, TEA, DMAP, DCM, 0 °C-RT, 2 h (84%); h) **8**, NaH, DMF, 3 h (63%); i) NaI, acetone, reflux, 1 h (86%).



Scheme 3 a) DMF, 50 °C, 1 h (n = 1, 87%; n = 4, 84%; n = 7, 84%); b) H_2, Pd/C, MeOH, 1 h (n = 1, 4, 7, quant.); c) **11**, DMF, dark, 2.5 h (n = 1, 78%; n = 4, 81%; n = 7, 82%).

and minimal purification, potentiating its use for glycobiological studies. However, unexpected anomerization during the development of this synthetic route occurred, and efforts to improve the anomeric purities were thus undertaken (see details below).

Carbohydrate-protein interactions depend not only on the carbohydrate structure *per se*, but also on the mode of presentation with respect to linker length and linker structure. In addition to evaluating the OEG type of linker structures with three, six and nine ethylene glycol units, the fourth synthetic pathway was developed in which an alternative linker structure[15] was designed and synthesized (compound **5**, Fig. 1). This structure was based on Cu-catalyzed azide-alkyne cycloaddition,[49] yielding a linker-based triazole ring. This alternative method to functionalize carbohydrate structures is equally attractive, owing to the efficiency of the cycloaddition reaction, and therefore the binding properties of this structural element were evaluated in comparison with the OEG type of linkers.

Stereocontrol in the synthesis of *S*-linked glycoconjugates

The stereoselective formation of the glycosidic linkage is one of the most challenging aspects in glycoconjugate and oligosaccharide synthesis.[50] This remains a challenge not only for *O*-linked structures, but also for thioglycosides.

Sulfur-based glycosyl donors are increasingly used in synthesis,[38–45] partly because of their reactivities, and partly due to the high biocompatibility and improved resistance towards hydrolytic enzymes for the products.[40,51] However, the general understanding of thioglycosylation reactions, and especially of their stereocontrol, is limited.[52,53] This was further addressed in the the present study. Initially, 1-thio-α-D-mannopyranose **23** was used to displace a linker tosylate group (**13**), leading to the intermediate carbohydrate-azide. However, in contrast to the expectation, a mixture was produced under the reaction conditions used, and NMR-analysis revealed that both *S*-linked α- and β-mannose-azides were obtained (Fig. 2).[54]

These results prompted further investigation of the thioglycosylation process, and the origin of the observed anomerization. Generally, 1-thiosaccharides are regarded relatively stable due to the poor orbital overlap between the anomeric carbon and the sulfur atom, impeding ring opening and subsequent mutarotation. However, a recent in-depth study on the mutarotation of 1-thio-aldoses in aqueous media has shown that anomerization of certain 1-thiosaccharides readily occurs at different pH in aqueous media.[55] Ring opening and exocyclic cleavage were found to account for the conspicuous mutarotations.

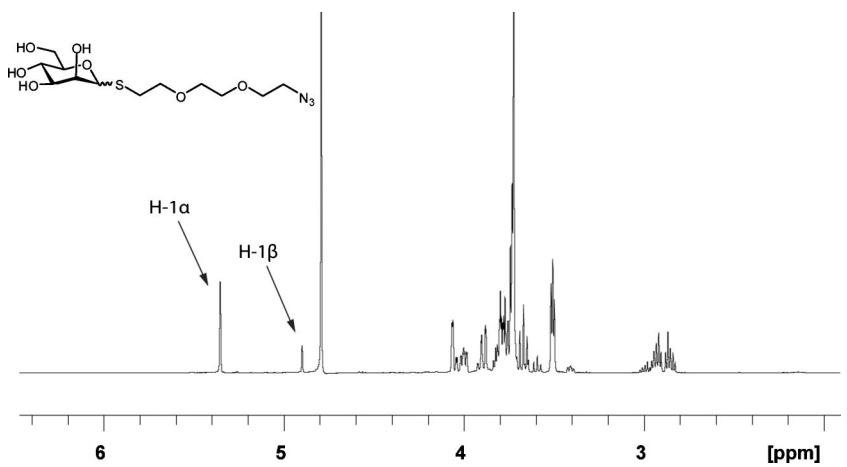


Fig. 2 ^{1}H-NMR spectrum of α- and β-anomers of compound **24** in D_2O.

The mechanistic nature of the S_N2 reaction combined with the mutarotation effect prompted the screening of the reaction conditions for stereocontrolled thioglycosylation.[54] Different reaction time, temperature and solvents were screened, and the effects of two electrophiles, tosylate and halide, were compared. Most significantly, when the linker iodide **14** and the 1-thiomannopyranose **23** were coupled in aprotic polar solvents like DMF at 50 °C for 1 h, complete stereoselectivity in favor of the α-anomer product was realized (Fig. 3).

Stereocontrol of the *O*-linked carbohydrate conjugates **1**, **5** and **6** was conveniently achieved using conventional neighboring group participation protocols (*cf*. Scheme 1).

Photo-induced glycoarray fabrication and SPRi analysis

Glycoarray fabrication was realized using the PFPA-conjugated carbohydrates using a double surface photoligation methodology (Scheme 4).[16] Gold substrates were first covalently activated to establish a PFPA-layer, and subsequently covered with a polymer film followed by UV irradiation. The polymer was covalently immobilized on the gold surface by a fast and efficient CH insertion reaction of the photogenerated perfluorophenyl nitrene from PFPA.[56] The resulting matrix then served as the substrate for the photoligation of the different PFPA-derivatized carbohydrates. A range of protein-resistant polymers, including poly(2-ethyl-2-oxazoline) (PEOX), poly(ethylene glycol) (PEG), and poly(ethylene oxide) (PEO), were initially evaluated.[57] Of these, PEOX was chosen due to a more stable performance in the instrumentation setup used. The different PFPA-carbohydrate conjugates (Fig. 1) were subsequently photo-immobilized on the surfaces, after which the sensors were ready for subsequent protein binding studies by SPRi.

Con A,[58] the plant lectin originally extracted from the jack-bean, *Canavalia ensiformis*, was used as a model lectin to probe the carbohydrate-protein interactions in this study. Firstly, the optimal concentration of the glycoconjugate needed for the lectin binding studies was determined. Printing solutions ranging from 0.1 μM to 100 mM, and subsequently in a more detailed study from 2 mM to 20 mM were evaluated, and the resulting sensors were interrogated with Con A (10 μM). As a result, 10 mM of glycoconjugate was chosen for following studies since it afforded the highest binding with relatively low concentration of glycoconjugates used, possibly due to the effect of surface carbohydrate ligand density on lectin binding (*cf.* ESI: Figure S21†).[18] The binding capacities of Con

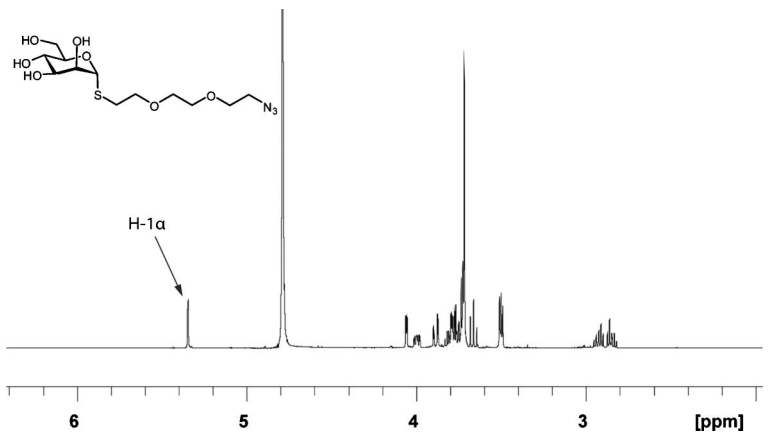
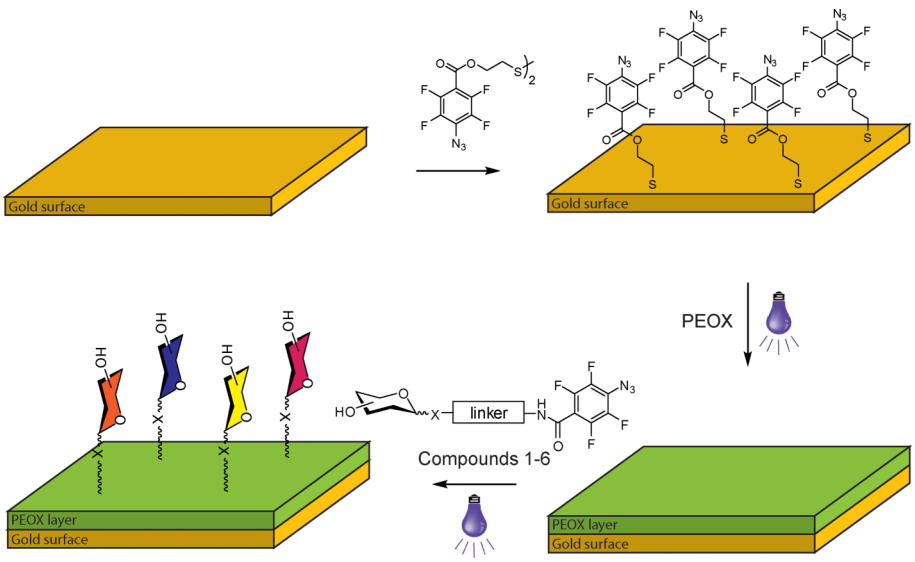


Fig. 3 ^{1}H-NMR spectrum of compound **24** in D_2O.



Scheme 4 Double surface photoligation of different glycoconjugate structures on gold surfaces for glycoarray fabrication.

A towards the six different glycoconjugates were next examined. SPRi data for a single sensor printed with six spots of each glycoconjugate and interrogated with Con A is shown in Fig. 4.

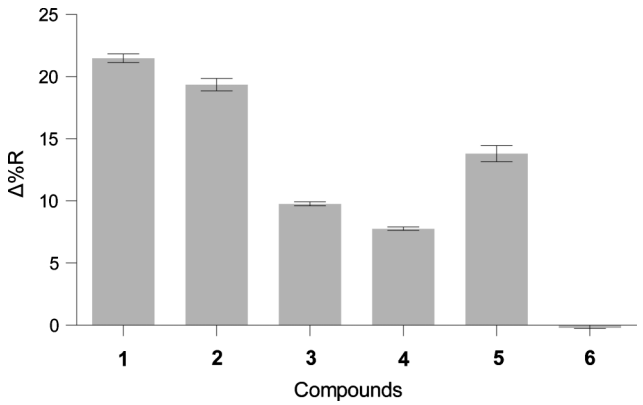


Fig. 4 SPRi responses of Compounds **1–6** toward Con A interrogation. Each data point (± SEM) represents the average response of six replicates for each compound. Δ%R in the Y-axis means the percentage of change in reflectivity.

The concentration of each compound was fixed at 10 mM, and 10 μM of interrogating Con A solution was then applied. First of all, the data clearly shows that Con A selectively binds to its cognate mannose ligands (compounds **1–5**), but not to the control galactose-derivative (compound **6**) on the sensor. This matches the established specificity of the lectin.[58,59] Good lectin selectivity and minimum nonspecific binding for the photogenerated biointerfaces were thus demonstrated. Moreover, the effects of structural variations among the mannose-derivatives (compounds **1–5**) on lectin binding performances were also shown.

The performances of the *S*-linked glycoconjugates and their *O*-linked analogues were compared by Con A interrogation towards glycoconjugates **1** and **2**. The results in Fig. 4 indicate that glycoconjugates **1** and **2** show relatively similar binding capacities toward Con A, with that of **1** being slightly higher. To the best of our knowledge, this is the first report on the biological evaluation of *S*-linked glycoconjugates in comparison with their *O*-linked analogs using the SPRi technique.

The effects of different linker lengths and linker structures on protein binding were further addressed. Three different linker lengths (3, 6 and 9 EG units) and two different linker structures (linear and triazole-containing) were evaluated for protein binding on the same sensor with the conditions developed. Previous studies showed that a reduced distance between the PFPA moiety and the carbohydrate group resulted in less efficient interactions, and the TEG-spacer linked PFPA-carbohydrates yielded the best binding results.[16] Thus, only linkers longer than TEG moiety were targeted in the present study. Glycoconjugate **2**, **3** and **4** contained spacers of three, six and nine ethylene glycol units, respectively. Glycoconjugate **5** had an equivalent chain length to **3** however possessing a triazole-ring within the molecule. As shown in Fig. 4, the lectin binding capacity decreases as the spacer length increases over the tested range under the system used. This observation is to some extent expected, likely due to a combination of entropic effects, potential entanglement, and increased intramolecular photoinsertion during UV irradiation. On the other hand, compared to the linear-linker glycoconjugate **3**, the triazole-ring-containing counterpart **5** shows slightly higher binding capacity when applied to lectin interrogation. This result is in agreement with previous results, showing that biomolecules retain their biological activities after the introduction of triazole rings.[60–62]

From these results, it is clear that all the synthesized glycoconjugates demonstrate good selectivity for the lectin, however varying in sensitivity. Especially, the *O*- and *S*-linked glycoconjugates **1** and **2** both show high performances, indicating that the replacement of the glycosidic oxygen for a sulfur atom results in no significant loss of binding. Furthermore, the results presented here show that this photogenerated glycoarray system is capable of distinguishing small variations in the glycoconjugate structures and conveying their effects on binding in a convenient manner.

Conclusions

The results reported in this study demonstrate that good selectivities and sensitivities can be achieved with photogenerated glycoarrays on SPRi sensors. Furthermore, the effects of surface chemical structural variations on protein binding were evaluated. The binding performances of *O*-linked glycoconjugates and its *S*-linked analogs were compared for the first time by the SPRi technique, indicating the effect of that element variation on protein binding is negligible in the present case. The intermediate linker length proved to be optimal, and the linker structure difference with regard to the incorporated triazole-ring and the linear feature showed little influence on lectin binding in the surface system used. The PFPA-based photochemistry proved versatile, and reliable for surface modifications and biosensor production. During the synthesis of the glycoconjugates, several different synthetic pathways were developed, resulting in good yields in few steps. The method utilizing glycosyl thiols proved especially efficient (up to 70% of total yield), highlighting its potential in various glycobiological studies, also owing to the fact that the products are as effective as their *O*-linked analogs while more resistant towards hydrolytic enzymes. Anomerization in the thioglycosylation process occurred, but excellent stereocontrol could be accomplished.

Experimental

General

Water used for analysis was obtained from a Millipore Milli-Q system with at least 18.2 MΩ resistivity. Concanavalin A (Con A) from jack bean *Canavalia ensiformis*, and poly(2-ethyl-2-oxazoline) (PEOX) (M_w *ca.* 500,000) were from Sigma–Aldrich (St Louis, MO). All other solvents and commercially available compounds were of HPLC, certified ACS, or reagent grade and used as received. PBS (pH 7.4) buffer was prepared with KH_2PO_4 (1.69 mM), Na_2HPO_4 (8.00 mM), and NaCl (154 mM) in Milli-Q water (1.00 L). ^{1}H, ^{13}C and ^{19}F NMR data were recorded on a Bruker Avance 400 instrument at 400 MHz (^{1}H) or a Bruker DMX 500 instrument at 500 MHz (^{1}H), 125 MHz (^{13}C) or 470 MHz (^{19}F). Chemical shifts are reported as δ values (ppm) with either $CDCl_3$ (^{1}H: δ = 7.26, ^{13}C = 77.16), DMSO-d_6 (^{1}H: δ = 2.50, ^{13}C = 39.52) or D_2O (^{1}H: δ = 4.79) as internal standard. *J* values are given in Hz. ^{1}H peak assignments were made by

first order analysis of the spectra supported by standard ^{1}H–^{1}H correlation spectroscopy (COSY). Thin layer chromatography (TLC) was performed on precoated Cromatofolios AL Silica gel 60 F_{254} plates (Merck, Darmstadt, Germany). Flash column chromatography was performed on silica gel 60, 0.040–0.063 mm (SDS, Val de Reuil, France). All experiments containing penta-/tetrafluorophenylazide groups were conducted in the absence of light with all reaction flasks covered with aluminum foil to prevent decomposition.

2-(2-(2-Azidoethoxy)ethoxy)ethanol (8)

2-(2-(2-Chloroethoxy)ethoxy)ethanol (5.0 g, 30 mmol) and sodium azide (3.9 g, 60 mmol) were dissolved in DMF (40 mL). The mixture was stirred at 90 °C overnight under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with ethyl acetate (EtOAc) for three times. The organic phase was combined, washed with brine, dried over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hexanes (Hex)/EtOAc (1 : 1 v/v), giving compound **8** as a colorless oil (quant.). ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.74 (t, 2 H, J = 4.4 Hz, CH_2CH$_2$OH), 3.70–3.66 (m, 6 H, 2 × OCH_2 and CH$_2$CH_2N$_3$), 3.62 (t, 2 H, J = 4.4 Hz, CH$_2$CH_2OH), 3.40 (t, 2 H, J = 5.1 Hz, CH_2N$_3$). ^{13}C NMR (125 MHz, $CDCl_3$): δ 72.61, 70.81, 70.55, 70.22, 61.93, 50.80. The spectroscopic data are in agreement with the literature.[14]

Compound 9

Penta-*O*-acetate-α-D-mannose **7** (3.0 g, 7.7 mmol) and compound **8** (2.0 g, 11 mmol) in dry dichloromethane (DCM, 30 mL) was cooled to 0 °C. BF_3·Et_2O, (5.5 g, 38 mmol) was added dropwise, and the solution was stirred overnight at RT. After completion of the reaction, the product solution was poured into ice water (30 mL), extracted with DCM, and the organic phase washed with saturated $NaHCO_3$, brine, dried over Na_2SO_4 and the solvent evaporated. Then the crude product was purified by flash column chromatography using solvent system Hex/EtOAc (2 : 1 v/v), giving compound **9** as a colorless oil (60%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.35 (dd, 1 H, J = 3.5 and 9.9 Hz, H-3), 5.28 (t, 1 H, J = 9.9 Hz, H-4), 5.26 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 4.86 (d, 1 H, J = 1.6 Hz, H-1), 4.28 (dd, 1 H, J = 5.0 and 12.3 Hz, H-6a), 4.09 (dd, 1 H, J = 2.5 and 12.3 Hz, H-6b), 4.05 (m, 1 H, H-5), 3.81 (m, 1 H, OCH_2), 3.71–3.64 (m, 9 H, OCH_2), 3.39 (t, 2 H, J = 5.4 Hz, CH_2N$_3$), 2.14 (s, 3 H, OAc), 2.09 (s, 3 H, OAc), 2.03 (s, 3 H, OAc), 1.98 (s, 3 H, OAc). ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.85, 170.21, 170.06, 169.88, 97.89, 70.96, 70.86, 70.26, 70.23, 69.74, 69.23, 68.56, 67.57, 66.33, 62.58, 50.84, 21.06, 20.92, 20.86. The spectroscopic data are in agreement with the literature.[63]

Compound 10

Compound **9** (202 mg, 0.4 mmol) was dissolved in dry MeOH (3 mL), and NaOMe (11 mg, 0.20 mmol) was added. The reaction mixture was stirred at RT for 1 h. Amberlite IR-120 H^+ resin was added to adjust the pH to 7, then the solution was filtered and the solvent evaporated, giving compound **10** as a colorless oil (quant.). ^{1}H NMR (500 MHz, D_2O): δ 4.88 (d, 1 H, J = 1.6 Hz, H-1), 3.96 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 3.91–3.85 (m, 2 H, OCH_2), 3.82 (dd, 1 H, J = 3.5 and 9.4 Hz, H-3), 3.78–3.71 (m, 9 H, 4 × OCH_2, H-5), 3.71–3.61 (m, 3 H, H-4, H-6a, H-6b), 3.50 (t, 2 H, J = 5.1 Hz, CH_2N$_3$). ^{13}C NMR (125 MHz, D_2O): δ 99.92, 72.71, 70.46, 69.92, 69.60, 69.52, 69.47, 69.23, 66.70, 66.37, 60.88, 50.12. The spectroscopic data are in agreement with the literature.[63]

Compound 1

Compound **10** (135 mg, 0.4 mmol) was dissolved in dry MeOH (4 mL), and Pd/C (40 mg) was added. Then the flask was purged with N_2 and filled with H_2. The solution was stirred vigorously at RT for 1 h. And then the reaction mixture was filtered through celite and the solvent evaporated to give the amino product. The crude amino product and compound **11** (133 mg, 0.4 mmol) were dissolved in dry DMF (4 mL), and the reaction mixture was stirred in dark at RT for 2.5 h under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the crude product was purified by flash column chromatography using solvent system DCM/MeOH (11 : 1 v/v), giving compound **1** as a colorless oil (62%). ^{1}H NMR (500 MHz, D_2O): δ 4.87 (d, 1 H, J = 1.5 Hz, H-1), 3.95–3.93 (m, 1 H, H-2), 3.89–3.59 (m, 17 H, 6 × OCH_2, H-3, H-4, H-5, H-6a, H-6b). ^{13}C NMR (125 MHz, D_2O): δ 160.73, 144.49, 142.56, 141.32, 139.39, 122.43, 110.11, 99.91, 72.71, 70.48, 69.94, 69.68, 69.49, 68.52, 66.69, 66.33, 60.88, 39.69. The spectroscopic data are in agreement with the literature.[17]

2-(2-(2-Azidoethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (13)

Compound **8** (1.0 g, 5.7 mmol) was dissolved in THF (30 mL), triethylamine (4.6 g, 45.6 mmol) and DMAP (77 mg, 0.6 mmol) were added. The mixture was stirred under N_2 and cooled down to 0 °C. Then a solution of tosyl chloride (4.4 g, 22.8 mmol) in THF (20 mL) was added portion by portion to the mixture. The reaction was allowed to warm up to room temperature. The mixture was washed by water and extracted by DCM three times. The organic phase was combined, washed with 1 M HCl, brine, dried by Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (2 : 1 v/v), giving compound **13** as a colorless oil (78%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.80 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 7.34 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 4.16 (t, 2 H, J = 4.8 Hz, CH_2OTs), 3.70 (t, 2 H, J = 4.8 Hz, CH_2CH$_2$OTs), 3.64 (t, 2H, J = 5.1 Hz, CH_2CH$_2$N$_3$), 3.60 (s, 4 H, 2 × OCH_2), 3.36 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.45 (s, 3 H, OC$_6$H$_4$CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.97, 133.15, 129.97, 128.12, 70.95, 70.77, 70.24, 69.38, 68.94, 50.81, 21.79. The spectroscopic data are in agreement with the literature.[64]

1-Azido-2-(2-(2-iodoethoxy)ethoxy)ethane (14)

Compound **13** (329 mg, 1 mmol) and sodium iodide (450 mg, 3 mmol) were dissolved in DMF (6 mL). The mixture was stirred at 50 °C for five hours under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with EtOAc three times. The organic phase was combined, washed with brine, dried by Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (7 : 1 v/v), giving compound **14** as a colorless oil (82%). R_f 0.15

(6 : 1 hexane/EtOAc). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 2924, 2867, 2104, 1444, 1350, 1284, 1125, 1034, 993, 931, 853. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.77 (t, 2 H, J = 6.6 Hz, CH_2CH$_2$I), 3.70 (t, 2H, J = 5.1 Hz, CH_2CH$_2$N$_3$), 3.68 (s, 4 H, 2 × OCH_2), 3.40 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 3.27 (t, 2 H, J = 6.6 Hz, CH_2I). ^{13}C NMR (125 MHz, $CDCl_3$): δ 72.17, 70.84, 70.43, 70.30, 50.87, 3.01. HRMS: calcd. for $C_6H_{12}IN_3O_2$ [M + Na]$^+$ 307.9866; observed 307.9868.

1,8-Ditosyl-3,6-dioxaoctane (16)

Triethylene glycol (**15**) (1.5 g, 10 mmol) was dissolved in THF (30 mL), triethylamine (4.0 g, 40 mmol) and DMAP (13 mg, 1.1 mmol) were added. The mixture was stirred under N_2 and cooled down to 0 °C. Then a solution of tosyl chloride (4.0 g, 21 mmol) in THF (20 mL) was added portion by portion to the mixture. The reaction was allowed to warm up to room temperature. The mixture was washed with water and extracted with DCM three times. The organic phase was combined, washed with 1 M HCl, brine, dried over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (3 : 1 v/v), giving compound **16** as a white powder (81%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.79 (d, 4 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 7.34 (d, 4 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 4.14 (t, 4 H, J = 4.8 Hz, 2 × CH_2OTs), 3.65 (t, 4 H, J = 4.8 Hz, 2 × CH_2CH$_2$OTs), 3.53 (s, 4 H, 2 × OCH_2), 2.45 (s, 6 H, 2 × OC$_6$H$_4$CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.99, 133.16, 120.99, 128.12, 70.86, 69.34, 68.91, 21.79. The spectroscopic data are in agreement with the literature.[65]

17-Azido-3,6,9,12,15-pentaoxaheptadecyl 4-methylbenzenesulfonate (17)

To a solution of compound **16** (3.7 g, 8 mmol) and compound **8** (700 mg, 4 mmol) was added 60% NaH in oil (640 mg, 16 mmol) under nitrogen portion by portion. The mixture was stirred at RT for 1.5 h. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with DCM three times. The organic phase was combined, washed with 1 M HCl, brine, dried over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (1 : 5 v/v), giving compound **17** as a colorless oil (59%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.79 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 7.34 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 4.15 (t, 2 H, J = 4.8 Hz, CH_2OTs), 3.72–3.55 (m, 20 H, 10 × OCH_2), 3.38 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.45 (s, 3 H, OC$_6$H$_4$CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.80, 132.98, 129.83, 127.95, 70.71, 70.64, 70.59, 70.55, 70.54, 70.49, 70.00, 69.28, 68.64, 50.66, 21.62. The spectroscopic data are in agreement with the literature.[66]

1-Azido-17-iodo-3,6,9,12,15-pentaoxaheptadecane (18)

Compound **17** (461 mg, 1 mmol) and sodium iodide (750 mg, 5 mmol) were dissolved in DMF (10 mL). The mixture was stirred at 50 °C for five hours under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with EtOAc three times. The organic phase was combined, washed with brine, dried over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (1 : 1 v/v), giving compound **18** as a colorless oil (81%). R_f 0.18 (1 : 1 hexane/EtOAc). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 2924, 2867, 2104, 1756, 1612, 1458, 1351, 1288, 1105, 1036, 998, 946, 854. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.76 (t, 2 H, J = 6.6 Hz, CH_2CH$_2$I), 3.71–3.63 (m, 18H, 9 × OCH_2), 3.39 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 3.26 (t, 2 H, J = 6.6 Hz, CH_2I). ^{13}C NMR (125 MHz, $CDCl_3$): δ 72.11, 70.84, 70.81, 70.76, 70.72, 70.36, 70.16, 50.82, 3.04. HRMS: calcd. for $C_{12}H_{24}IN_3O_5$ [M + Na]$^+$ 440.0653; observed 440.0645.

1,17-Ditosyl-3,6,9,12,15-pentaoxaheptadecane (20)

To a solution of hexaethyleneglycol (**19**) (2.82 g, 10 mmol) in DCM (35 mL) at 0 °C were added triethylamine (3.03 g, 30 mmol), tosyl chloride (4.78 g, 25 mmol) and DMAP (610 mg, 5 mmol). After being stirred under N_2 at RT for 2 h, the mixture was neutralized by saturated NH_4Cl solution and extracted with DCM three times. The organic phase was combined, washed with brine, dried by Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system Hex/EtOAc (1 : 4 v/v), giving compound **20** as a colorless oil (84%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.79 (d, 4 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 7.34 (d, 4 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 4.15 (t, 4 H, J = 4.8 Hz, 2 × CH_2OTs), 3.68 (t, 4 H, J = 4.8 Hz, 2 × CH_2CH$_2$OTs), 3.64–3.59 (m, 8 H, 4 × OCH_2), 3.58–3.56 (m, 8 H, 4 × OCH_2), 2.44 (s, 6 H, 2 × OC$_6$H$_4$CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.93, 133.13, 129.96, 128.11, 70.88, 70.74, 70.68, 70.64, 69.39, 68.81, 21.78. The spectroscopic data are in agreement with the literature.[65]

26-Azido-3,6,9,12,15,18,21,24-octaoxahexacosyl 4-methylbenzenesulfonate (21)

To a solution of compound **20** (1.89 g, 3.4 mmol) and compound **8** (368 mg, 2.1 mmol) was added 60% NaH in oil (337 mg, 8.4 mmol) under nitrogen portion by portion. The mixture was stirred at RT for 3 h. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with DCM three times. The organic phase was combined, washed with 1 M HCl, brine, dried over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system EtOAc/EtOH (10 : 1 v/v), giving compound **21** as a colorless oil (63%). R_f 0.38 (8 : 1 EtOAc/EtOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 2874, 2108, 1603, 1457, 1356, 1292, 1252, 1176, 1101, 1019, 925, 819, 781. ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.79 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 7.34 (d, 2 H, J = 8.3 Hz, OC$_6H_4$CH$_3$), 4.16 (t, 2 H, J = 4.8 Hz, CH_2OTs), 3.72–3.55 (m, 32 H, 16 × OCH_2), 3.39 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.45 (s, 3 H, OC$_6$H$_4$CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ 144.91, 133.14, 129.94, 128.11, 70.87, 70.82, 70.79, 70.75, 70.73, 70.69, 70.64, 70.15, 69.36, 68.80, 50.81, 21.76. HRMS: calcd. for $C_{25}H_{43}N_3O_{11}S$ [M + Na]$^+$ 616.2511; observed 616.2494.

1-Azido-26-iodo-3,6,9,12,15,18,21,24-octaoxahexacosane (22)

Compound **21** (593 mg, 1 mmol) and sodium iodide (750 mg, 5 mmol) were dissolved in acetone (5 mL). The mixture was stirred and refluxed for one hour under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the residue was washed with water and extracted with EtOAc three times. The organic phase was combined, washed with brine, dried

over Na_2SO_4 and evaporated. The crude product was purified by flash column chromatography using solvent system 100% of EtOAc, giving compound **22** as a colorless oil (86%). R_f 0.36 (8 : 1 EtOAc/EtOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 2867, 2106, 1458, 1351, 1298, 1256, 1108, 1043, 1000, 950, 854. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.76 (t, 2 H, J = 6.6 Hz, CH_2CH$_2$I), 3.71–3.62 (m, 30 H, 15 × OCH_2), 3.39 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 3.26 (t, 2 H, J = 6.6 Hz, CH_2I). ^{13}C NMR (125 MHz, $CDCl_3$): δ 72.11, 70.83, 70.80, 70.76, 70.72, 70.36, 70.16, 50.82, 3.05. HRMS: calcd. for $C_{18}H_{36}IN_3O_8$ [M + Na]$^+$ 572.1439; observed 572.1425.

Compound 24

Compound **14** (86 mg, 0.3 mmol) and compound **23** (78 mg, 0.36 mmol) were dissolved in dry DMF (2 mL). The mixture was stirred at 50 °C for 1 h under nitrogen atmosphere. The solvent was evaporated under reduced pressure. Then the crude product was purified by flash column chromatography using solvent system DCM/MeOH (10 : 1 v/v), giving compound **24** as a colorless oil (87%). R_f 0.50 (6 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 3383, 2931, 2879, 2114, 1657, 1604, 1448, 1422, 1351, 1304, 1076, 972, 920, 882, 853, 804. ^{1}H NMR (500 MHz, D_2O): δ 5.35 (d, 1 H, J = 1.6 Hz, H-1), 4.06 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 4.00 (m, 1 H, H-5), 3.89 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.84–3.69 (m, 10 H, 4 × OCH_2, H-3, H-6b), 3.67 (t, 1 H, J = 9.8 Hz, H-4), 3.50 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.96–2.81 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 84.90, 73.15, 71.66, 71.00, 69.57, 69.50, 69.37, 69.23, 67.09, 60.86, 50.17, 29.98. HRMS: calcd. for $C_{12}H_{23}N_3O_7S$ [M + Na]$^+$ 376.1149; observed 376.1143.

Compound 25

Compound **25** was prepared following the same procedure as for compound **24**. Colorless oil (84%). R_f 0.46 (6 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 3396, 2927, 2874, 2109, 1657, 1604, 1455, 1423, 1351, 1304, 1257, 1104, 1080, 953, 854, 802. ^{1}H NMR (500 MHz, D_2O): δ 5.35 (d, 1 H, J = 1.6 Hz, H-1), 4.06 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 4.00 (m, 1 H, H-5), 3.89 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.84–3.69 (m, 22 H, 10 × OCH_2, H-3, H-6b), 3.67 (t, 1 H, J = 9.8 Hz, H-4), 3.51 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.96–2.81 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 84.91, 73.17, 71.67, 71.01, 69.60, 69.55, 69.38, 69.23, 67.10, 60.87, 50.16, 29.97. HRMS: calcd. for $C_{18}H_{35}N_3O_{10}S$ [M + Na]$^+$ 508.1935; observed 508.1924.

Compound 26

Compound **26** was prepared following the same procedure as for compound **24**. Colorless oil (84%). R_f 0.45 (6 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 3402, 2879, 2114, 1657, 1604, 1457, 1351, 1306, 1257, 1106, 946, 853, 802. ^{1}H NMR (500 MHz, D_2O): δ 5.35 (d, 1 H, J = 1.6 Hz, H-1), 4.06 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 4.00 (m, 1 H, H-5), 3.89 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.84–3.69 (m, 34 H, 16 × OCH_2, H-3, H-6b), 3.67 (t, 1 H, J = 9.8 Hz, H-4), 3.51 (t, 2 H, J = 5.1 Hz, CH_2N$_3$), 2.96–2.81 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 84.91, 73.18, 71.67, 71.02, 69.59, 69.55, 69.38, 69.23, 67.10, 60.87, 50.15, 29.97. HRMS: calcd. for $C_{24}H_{47}N_3O_{13}S$ [M + Na]$^+$ 640.2722; observed 640.2709.

Compound 2

Compound **24** (69 mg, 0.19 mmol) was dissolved in dry MeOH (3 mL), and Pd/C (20 mg) was added. The flask was purged with N_2 and filled with H_2 after which the solution was stirred vigorously at RT for 1 h. The reaction mixture was subsequently filtered through celite and the solvent evaporated to give the amino product. To a solution of the crude amino product in dry DMF (3 mL) was added compound **11** (63 mg, 0.19 mmol). The reaction mixture was stirred in the dark at RT for 2.5 h under nitrogen atmosphere. The solvent was evaporated under reduced pressure, and the crude product was purified by flash column chromatography using solvent system DCM/EtOH (12 : 1 v/v), yielding compound **2** as a colorless oil (78%). R_f 0.55 (6 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 3300, 3099, 2933, 2884, 2129, 1657, 1562, 1489, 1423, 1354, 1334, 1272, 1231, 1078, 996, 953, 920, 882, 853, 802. ^{1}H NMR (500 MHz, D_2O): δ 5.32 (d, 1 H, J = 1.6 Hz, H-1), 4.03 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 3.96 (m, 1 H, H-5), 3.87 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.82–3.68 (m, 10 H, 4 × OCH_2, H-3, H-6b), 3.65 (t, 1 H, J = 9.8 Hz, H-4), 3.62 (t, 2 H, J = 5.4 Hz, CH_2NH), 2.93–2.78 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 160.65, 144.52, 142.56, 141.40, 139.42, 122.44, 110.17, 84.90, 73.16, 71.66, 71.03, 69.58, 69.49, 69.47, 68.52, 67.07, 60.85, 39.74, 29.98. HRMS: calcd. for $C_{19}H_{24}F_4N_4O_8S$ [M + Na]$^+$ 567.1149; observed 567.1135.

Compound 3

Compound **3** was prepared following the same procedure as for compound **2**. Colorless oil (81%). R_f 0.52 (6 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1}3328, 3095, 2929, 2882, 2129, 1657, 1562, 1491, 1423, 1330, 1272, 1231, 1101, 1099, 1082, 994, 953, 922, 882, 847, 802. ^{1}H NMR (500 MHz, D_2O): δ 5.34 (d, 1 H, J = 1.6 Hz, H-1), 4.05 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 3.98 (m, 1 H, H-5), 3.88 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.82–3.67 (m, 22 H, 10 × OCH_2, H-3, H-6b), 3.66 (t, 1 H, J = 9.8 Hz, H-4), 3.62 (t, 2 H, J = 5.4 Hz, CH_2NH), 2.95–2.80 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 160.67, 144.50, 142.53, 141.39, 139.40, 122.47, 110.18, 84.92, 73.17, 71.67, 71.02, 69.68, 69.61, 69.59, 69.54, 69.50, 69.36, 68.51, 67.08, 60.86, 39.74, 29.96. HRMS: calcd. for $C_{25}H_{36}F_4N_4O_{11}S$ [M + Na]$^+$ 699.1930; observed 699.1918.

Compound 4

Compound **4** was prepared following the same procedure as for compound **2**. Colorless oil (82%). R_f 0.25 (10 : 1 DCM/MeOH). IR (ReactIR iC10, MeOH): ν_{max}/cm^{-1} 3352, 3093, 2929, 2879, 2129, 1657, 1558, 1492, 1423, 1353, 1330, 1272, 1231, 1103, 994, 953, 925, 884, 849, 802. ^{1}H NMR (500 MHz, D_2O): δ 5.35 (d, 1 H, J = 1.6 Hz, H-1), 4.05 (dd, 1 H, J = 1.6 and 3.5 Hz, H-2), 3.99 (m, 1 H, H-5), 3.88 (dd, 1 H, J = 2.2 and 12.3 Hz, H-6a), 3.83–3.67 (m, 34 H, 16 × OCH_2, H-3, H-6b), 3.66 (t, 1 H, J = 9.8 Hz, H-4), 3.62 (t, 2 H, J = 5.4 Hz, CH_2NH), 2.96–2.81 (m, 2 H, SCH_2). ^{13}C NMR (125 MHz, D_2O): δ 160.61, 144.53, 142.54, 141.35, 139.37, 122.42, 110.22, 84.92, 73.19, 71.68, 71.04, 69.69, 69.59, 69.51, 69.39, 68.52, 67.09, 60.87, 39.74, 29.97. HRMS: calcd. for $C_{31}H_{48}F_4N_4O_{14}S$ [M + 2Na]$^{2+}$ 427.1304; observed 427.1297.

Glycoarray fabrication

SF 10 glass substrates (18 × 18 mm, Schott Glass Technology, Inc, Elmsford, NY) were cleaned with piranha solution (7 : 3 v/v 98% sulfuric acid/35% hydrogen peroxide; **Caution**: the piranha solution reacts violently with organic solvents—take extreme care when handling this solution), and titanium (2 nm) and gold films (45 nm) were subsequently deposited on the surfaces using electron beam (EB) evaporation (performed at Washington Technology Center (WTC), Seattle). The gold substrates were cleaned using piranha solution for 30 s, thoroughly washed using boiling water for 20 min (3 times) and dried under a stream of nitrogen. The cleaned substrates were immersed in a solution of PFPA-disulfide[16] (1.7 mM in CH_2Cl_2) at RT in the dark overnight. Then the substrates were rinsed several times with CH_2Cl_2 to remove excess PFPA-disulfide and dried under a stream of nitrogen. A solution of PEOX in $CHCl_3$ (10 mg mL^{-1}) was spin-coated onto the PFPA-activated substrates at 2000 rpm for 60 s using a model P6204 spin-coater (Specialty Coating Systems, Indianapolis, IN). The spin-coated substrates were then UV-irradiated for 5 min through a 280 nm optical filter using a 450 W medium pressure Hg lamp (Hanovia, ACE Glass Inc., Vineland, NJ). The intensity at the sample location was measured to be 4.5 mW cm^{-2} using a 254-nm sensor. The irradiation time includes a 2 min warm-up period of the lamp to reach its full intensity. The substrates were subsequently soaked in ethanol for 3 h to remove unbound PEOX and then dried under a stream of nitrogen.

Solutions of PFPA-derivatized carbohydrates (10 mM in water) were printed on the PEOX coated gold substrates using a BioOdyssey Calligrapher Miniarrayer (Bio-Rad Laboratories, Hercules, CA). An MCP 360 capillary pin was used to dispense the carbohydrate solutions of varying concentrations from a 96-well plate onto the gold substrate using a robotic array spotter. A 6 × 6 array pattern of the PFPA-derivatized carbohydrates was created either with 6 different concentrations of each compound or 6 spots of the same concentration of each compound. The spot size was 400 μm in diameter and the distance between the spots 800 μm. The humidity level in the arraying chamber was maintained at 60% and the temperature at 20 °C. The substrates were then UV-irradiated for 5 min through a 280 nm optical filter. The substrates were rinsed with water to remove any excess PFPA-derivatized carbohydrates and washed with PBS containing 0.1% Tween 20 for 30 min. The substrates were finally dried under a stream of nitrogen.

SPR imaging

SPR imaging was performed using an SPR imagerII (GWC technologies, Madison, WI), operating at ambient temperature. The flow rate of the solutions throughout the experiments was maintained at 100 μL/min. The glycoarrays were first equilibrated with PBS containing 0.1% (w/v) Tween 20 (running buffer), then treated with 0.2% (w/v) BSA in PBS containing 0.1% (w/v) Tween 20 in order to prevent nonspecific binding, and subsequently re-equilibrated with running buffer. All protein binding cycles were initiated by equilibrating the glycoarrays with running buffer, followed by the specific protein solution, and finally re-equilibration with running buffer. After protein binding, the glycoarrays were regenerated with urea (8 M), and then re-equilibrated with running buffer.

Acknowledgements

This study was supported in part by the Swedish Research Council (contract no. 622-2003-2343), the Royal Institute of Technology, the National Institutes of General Medical Science (NIGMS) under NIH Award Numbers R01GM080295 and 2R15GM066279, ARL-ONAMI Center for Nanoarchitectures for Enhanced Performance. LD thanks the China Scholarship Council for a special scholarship award.

Notes and references

1 A. Varki, *Glycobiology*, 1993, **3**, 97.
2 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
3 A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, *Essentials of glycobiology*, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009.
4 D. G. Castner and B. D. Ratner, *Surf. Sci.*, 2002, **500**, 28.
5 B. W. Daniel and P. H. Seeberger, *Chem.–Eur. J.*, 2005, **11**, 3194.
6 R. Jelinek and S. Kolusheva, *Chem. Rev.*, 2004, **104**, 5987.
7 K. Larsen, M. B. Thygesen, F. Guillaumie, W. G. T. Willats and K. J. Jensen, *Carbohydr. Res.*, 2006, **341**, 1209.
8 M. Mammen, S. K. Choi and G. M. Whitesides, *Angew. Chem., Int. Ed.*, 1998, **37**, 2754.
9 B. Houseman and M. Mrksich, *Top. Curr. Chem.*, 2002, **218**, 1.
10 S. Angeloni, J. L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard, S. Kochhar, H. Sigrist and N. Sprenger, *Glycobiology*, 2005, **15**, 31.
11 G. Carroll, D. Wang, N. Turro and J. Koberstein, *Glycoconjugate J.*, 2008, **25**, 5.
12 G. T. Carroll, D. Wang, N. J. Turro and J. T. Koberstein, *Langmuir*, 2006, **22**, 2899.
13 Y. Chevolot, O. Bucher, D. Léonard, H. J. r. Mathieu and H. Sigrist, *Bioconjugate Chem.*, 1999, **10**, 169.
14 L.-H. Liu, H. Dietsch, P. Schurtenberger and M. Yan, *Bioconjugate Chem.*, 2009, **20**, 1349.
15 O. Norberg, L. Deng, M. Yan and O. Ramström, *Bioconjugate Chem.*, 2009, **20**, 2364.
16 Y. Pei, H. Yu, Z. Pei, M. Theurer, C. Ammer, S. André, H.-J. Gabius, M. Yan and O. Ramström, *Anal. Chem.*, 2007, **79**, 6897.
17 Z. Pei, H. Yu, M. Theurer, A. Waldén, P. Nilsson, M. Yan and O. Ramström, *ChemBioChem*, 2007, **8**, 166.
18 A. Tyagi, X. Wang, L. Deng, O. Ramström and M. Yan, *Biosens. Bioelectron.*, 2010, **26**, 344.
19 X. Wang, O. Ramström and M. Yan, *J. Mater. Chem.*, 2009, **19**, 8944.
20 M. S. Platz, *Acc. Chem. Res.*, 1995, **28**, 487.
21 H. Bayley and J. V. Staros, *In: Azides and Nitrenes*, E.F. Scriven (Ed.), Academic Press, San Diego, 1984, pp 434.
22 F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson and C. H. Wong, *J. Am. Chem. Soc.*, 2002, **124**, 14397.
23 B. Houseman and M. Mrksich, *Chem. Biol.*, 2002, **9**, 443.
24 T. Feizi, F. Fazio, W. Chai and C.-H. Wong, *Curr. Opin. Struct. Biol.*, 2003, **13**, 637.
25 N. Laurent, J. Voglmeir and S. L. Flitsch, *Chem. Commun.*, 2008, 4400.
26 K. R. Love and P. H. Seeberger, *Angew. Chem., Int. Ed.*, 2002, **41**, 3583.
27 E.-H. Song and N. LB Pohl, *Curr. Opin. Chem. Biol.*, 2009, **13**, 626.
28 S. Park, M.-R. Lee and I. Shin, *Chem. Commun.*, 2008, 4389.
29 D. Wang, S. Liu, B. J. Trummer, C. Deng and A. Wang, *Nat. Biotechnol.*, 2002, **20**, 275.
30 R. D'Agata, G. Grasso, G. Iacono, G. Spoto and G. Vecchio, *Org. Biomol. Chem.*, 2006, **4**, 610.
31 R. Karamanska, J. Clarke, O. Blixt, J. MacRae, J. Zhang, P. Crocker, N. Laurent, A. Wright, S. Flitsch, D. Russell and R. Field, *Glycoconjugate J.*, 2008, **25**, 69.
32 M. J. Linman, H. Yu, X. Chen and Q. Cheng, *ACS Appl. Mater. Interfaces*, 2009, **1**, 1755.
33 W. Liu, Y. Chen and M. Yan, *Analyst*, 2008, **133**, 1268.
34 E. Mercey, R. Sadir, E. Maillart, A. Roget, F. Baleux, H. Lortat-Jacob and T. Livache, *Anal. Chem.*, 2008, **80**, 3476.
35 E. A. Smith, W. D. Thomas, L. L. Kiessling and R. M. Corn, *J. Am. Chem. Soc.*, 2003, **125**, 6140.

36 W. Senaratne, L. Andruzzi and C. K. Ober, *Biomacromolecules*, 2005, **6**, 2427.
37 S. Hanessian and B. Lou, *Chem. Rev.*, 2000, **100**, 4443.
38 D. R. Bundle, J. R. Rich, S. Jacques, H. N. Yu, M. Nitz and C. C. Ling, *Angew. Chem., Int. Ed.*, 2005, **44**, 7725.
39 H. Driguez, *Top. Curr. Chem.*, 1997, **187**, 85.
40 N. Floyd, B. Vijayakrishnan, J. Koeppe and B. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 7798.
41 D. P. Gamblin, E. M. Scanlan and B. G. Davis, *Chem. Rev.*, 2009, **109**, 131.
42 S. Knapp and D. S. Myers, *J. Org. Chem.*, 2002, **67**, 2995.
43 K. Pachamuthu and R. R. Schmidt, *Chem. Rev.*, 2006, **106**, 160.
44 J. R. Rich, W. W. Wakarchuk and D. R. Bundle, *Chem.–Eur. J.*, 2006, **12**, 845.
45 D. A. Thayer, H. N. Yu, M. C. Galan and C. H. Wong, *Angew. Chem., Int. Ed.*, 2005, **44**, 4596.
46 J. M. Harris, *In Poly(ethylene glycol) Chemistry: Biotechnical and Biomedical Applications*, J. M. Harris, Ed., Plenum. Press, New York, 1992.
47 Y. Lee, H. Koo, G.-W. Jin, H. Mo, M. Y. Cho, J.-Y. Park, J. S. Choi and J. S. Park, *Biomacromolecules*, 2005, **6**, 24.
48 Z. Pei, R. Larsson, T. Aastrup, H. Anderson, J.-M. Lehn and O. Ramström, *Biosens. Bioelectron.*, 2006, **22**, 42.
49 Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless and M. G. Finn, *J. Am. Chem. Soc.*, 2003, **125**, 3192.
50 A. V. Demchenko, *In: Handbook of Chemical Glycosylation: Advances in Stereoselectivity and Therapeutic Relevance*, A. V. Demchenko, Ed., Wiley-VCH, Weinheim, 2008.
51 H. Driguez, *ChemBioChem*, 2001, **2**, 311.
52 L. Encinas and J. L. Chiara, *Eur. J. Org. Chem.*, 2009, 2163.
53 H. N. Yu, C.-C. Ling and D. R. Bundle, *J. Chem. Soc., Perkin Trans. 1*, 2001, 832.
54 Details will be described in a subsequent report.
55 R. Caraballo, L. Deng, L. Amorim, T. Brinck and O. Ramström, *J. Org. Chem.*, 2010, **75**, 6115.
56 L.-H. Liu and M. Yan, *Acc. Chem. Res.*, 2010, **43**, 1434.
57 H. Wang, J. Ren, A. Hlaing and M. Yan, *J. Colloid Interface Sci.*, 2010, **354**, 160.
58 J. B. Sumner, *J. Biol. Chem.*, 1919, **37**, 137.
59 J. C. Manimala, T. A. Roach, Z. Li and J. C. Gildersleeve, *Angew. Chem., Int. Ed.*, 2006, **45**, 3607.
60 M. D. Best, *Biochemistry*, 2009, **48**, 6571.
61 N. Devaraj and J. Collman, *QSAR Comb. Sci.*, 2007, **26**, 1253.
62 S. Hanson, W. Greenberg and C. H. Wong, *QSAR Comb. Sci.*, 2007, **26**, 1243.
63 J. Li, S. Zacharek, X. Chen, J. Wang, W. Zhang, A. Janczuk and P. G. Wang, *Bioorg. Med. Chem.*, 1999, **7**, 1549.
64 E. Fernandez-Megia, J. Correa, I. Rodriguez-Meizoso and R. Riguera, *Macromolecules*, 2006, **39**, 2113.
65 D. L. Mohler and G. Shen, *Org. Biomol. Chem.*, 2006, **4**, 2082.
66 G. Lu and K. Burgess, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 3902.

