Monitoring glycan-protein interactions by NMR. A simple chemical tag that mimics natural CH/π interactions.

Luis P. Calle,[a] Begoña Echeverria,[b] Antonio Franconetti,[c] Sonia Serna,[b] M. Carmen Fernández-Alonso,[a] Tammo Diercks,[d] F. Javier Cañada,[a] Ana Ardá,[d] Niels-Christian Reichardt*,[b] Jesús Jiménez-Barbero*,[b,e]

Abstract: Detection of molecular recognition processes requires robust, specific and easily implementable sensing methods, especially for screening applications. Here we propose the difluoroacetamide moiety (an acetamide bioisoster) as a novel tag for detecting by NMR those glycan-protein interactions that involve N-acetylated sugars. While difluoroacetamide has been used previously as substituent in medicinal chemistry, we here employ it as a specific sensor to monitor interactions between GlcNAc-containing glycans and a model lectin (wheat germ agglutinin). Contrary to the widely employed trifluoroacetamide group, the difluoroacetamide tag contains geminal ^{1}H and ^{19}F atoms that allow both ^{1}H- and ^{19}F-observed NMR methods for easy and robust detection of molecular recognition processes involving GlcNAc- (or GalNAc-) moieties over a range of binding affinities. The CHF_2CONH- moiety behaves very similar to the natural acetamide fragment in the involved aromatic-sugar interactions, providing analogous binding energy and conformations, while the perfluorated CF_3CONH- analogue differs more significantly.

Introduction

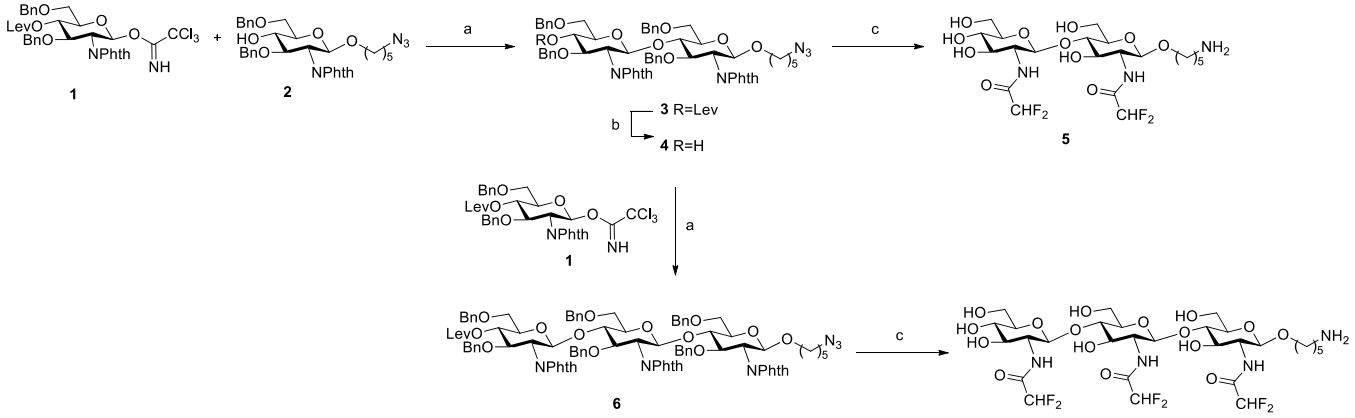
The study of molecular carbohydrate recognition processes is attracting enormous interest given their involvement in a variety of essential physiological events.[1] Advances in this field exploit new developments in glycan- and lectin-array technologies capable of detecting binding partners.[2] Their application would benefit from parallel structural studies to help understand the details of molecular interactions.[3] This combined approach, complemented by a profound knowledge of the conformational, dynamic and energetic aspects of the interaction, could guide the design of novel glycan-based therapeutics and/or probes able to modulate the interplay between the different actors.[4] At the chemical level, it is now well appreciated that sugar binding is mediated by hydrogen bonds, van der Waals interactions, and carbohydrate-aromatic CH-π stacking.[5] Depending on the particular system, different cations may also play an important role.[6] X-ray and NMR methods have demonstrated to be the methods of choice to analyse these forces and interactions.[7] NMR spectroscopy allows to follow the molecular recognition processes either from the ligand or the receptor's perspective, and via changes in various NMR parameters.[8] Particularly for large, poorly soluble or scarce target proteins the use of ligand-based NMR methods is indicated.[9] In this case, isotopes other than ^{1}H present in the ligand would be extremely helpful to provide for its observation with highest specificity and spectral resolution. For instance, recent examples have highlighted the application of ^{19}F[10] or ^{77}Se NMR[11] to detect molecular binding. Obviously, the synthesis of molecules containing ^{19}F or ^{77}Se may be very challenging, depending on the position to be labelled on the sugar moiety. For instance, introduction of ^{19}F in pyranose rings requires several synthetic steps and employs hazardous chemical reagents.[12] ^{19}F as a probe has also been applied to monitor structural features in proteins. In particular, the incorporation of difluoromethionine into proteins has been used to study the aminoacid packing or the existence of different conformational states.[13] Herein, we propose difluoroacetamide as a novel ^{19}F-containing tag that excellently mimics the common acetamide moiety present in a variery of saccharides. Unlike the well known trifluoroacetamide, difluoroacetamide has not been reported as an amine protecting group, but as a substituent in medicinal chemistry.[14] The difluoroacetamide group can be easily introduced by direct reaction of free amine groups with difluoroacetic anhydride, and offers unique NMR spectroscopic possibilities for both ^{1}H and ^{19}F detected methods that enormously facilitate the observation of ligand-protein interactions. We show that disaccharide and trisaccharide moieties containing this tag largely maintain the binding properties of their unmodified parent molecules in interactions with a model lectin, while their aromatic CH-π stacking is even more intense than for acetamide or trifluoroacetamide.

Results and Discussion

The difluoroacetamide-containing analogues of N,N´-diacetyl chitobiose and N,N',N¨-triacetyl chitotriose (see scheme 1,

[a] L. P. Calle, Dr. M. C. Fernández-Alonso, Prof. Dr. F. J. Cañada
Department of Chemical and Physical Biology
CIB-CSIC
Ramiro de Maeztu 9, 28040, Madrid (Spain)
[b] Dr. B. Echeverria, Dr. S. Serna, Dr. N.-C. Reichardt
Department of Glycotechnology
CICbiomaGUNE
Paseo Miramón 182, 20009, San Sebastián (Spain)
[c] A. Franconetti
Department of Organic Chemistry
University of Sevilla
Profesor García González 1, 41012, Sevilla (Spain)
[d] Dr. T. Diercks, Dr. A. Ardá, Prof. Dr. J. Jiménez-Barbero
Department of Chemical and Physical Biology
CICbiomaGUNE
Parque Tecnologico de Bizkaia, Building 801a, 48160-Derio (Spain)
[e] Ikerbasque. Basque Foundation for Science, 48013 Bilbao (Spain)

Supporting information for this article is given via a link at the end of the document.



Scheme 1. Synthesis of difluoroacetamide analogues **5** and **7** of N,N´-diacetyl chitobiose and N,N',N¨-triacetyl chitotriose. Reagents and conditions. a) 10% TMSOTf, dry CH_2Cl_2, 66% for **3**, 50% for **6**. b) N_2H_4AcOH, CH_2Cl_2, MeOH, 84%. c) i. $NH_2(CH_2)_2NH_2$, *n*BuOH, MW, 120ºC. ii. $(CHF_2O)_2CO$, pyridine. iii. NaOMe, MeOH. iv. H_2, 10% Pd/C, MeOH 1%TFA.

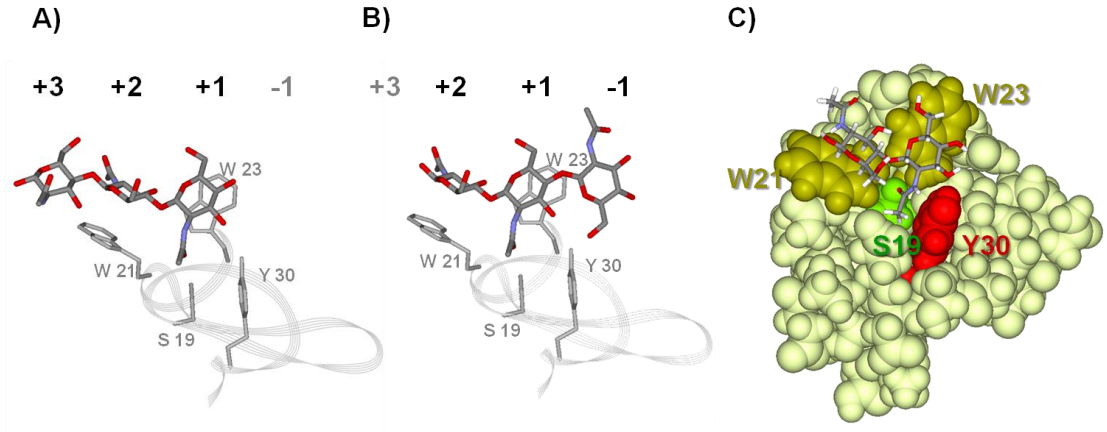


Figure 1. Representation of the two possible binding modes of $(GlcNAc)_3$ to hevein. A) Orientation +3, +2, +1. B) Orientation +2, +1, −1 (see text for discussion). C) Structure of the complex between hevein and N,N′-diacetyl chitobiose (amino acids in the binding site are highlighted).

compounds **5** and **7**) were synthesized as described in the experimental section. From the pool of known GlcNAc-binding lectins we chose wheat germ agglutinin (WGA) that is one of the best studied carbohydrate binding proteins.[15] It is composed of four hevein-like domains that preferentially bind GlcNAc, although sialic acid binding was also reported.[16] WGA recognizes GlcNAc moieties through a combination of hydrogen bonds and sugar-aromatic CH-π interactions (Figure 1) including one from the acetamide methyl group. [17]

In a glycan array experiment to confirm recognition of difluoroacetamide analogues by WGA we identified compounds **5** and **7** and their acetyl analogs (N,N´-diacetyl chitobiose and N,N',N¨-triacetyl chitotriose) as binding ligands on NHS activated glass slides. Incubation with the fluorescence labeled WGA-647 showed that the difluoroacetamide analogues were recognized as well as the natural N-acetyl ligands, suggesting only minimal impact (if any) of the difluoracetamide substitution on the carbohydrate-lectin interaction (Figure 2).

Difluoroacetamide as a dual NMR probe for molecular interaction.

^1H and ^{19}F NMR signals of the novel CHF_2-containing di- and tri-saccharides (**5** and **7**) were assigned by standard NMR techniques including homo- and heteronuclear ^{19}F-^{19}F and ^{19}F-^1H correlation experiments (Table S1 and Figures S1-S7 in the supporting information). The ^1H signal of each difluoroacetyl moiety appears as a triplet (due to heteronuclear coupling with $^2J_{HF} \approx 53.6$ Hz), and the strong deshielding by both geminal fluorine atoms favourably shifts it into a distinct spectral region around 6 ppm. The diastereotopic ^{19}F atoms in each sugar difluoroacetamide moiety give rise to two distinct NMR signals, analogous to diastereotopic methylene protons in saccharides (e.g., the hydroxymethyl protons at C6)[18] and differ between the connected monosaccharides in **5** and **7**, allowing their straightforward distinction. A spectral complication arises from strong homonuclear coupling ($^2J_{FF} \approx 305$ Hz) between the geminal ^{19}F spins that splits their signals into doublets, giving

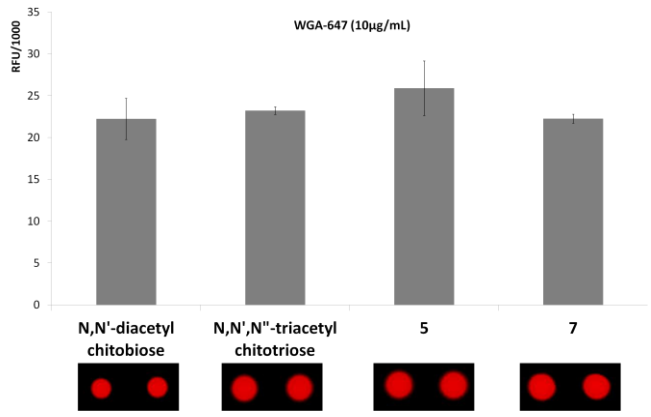


Figure 2. Fluorescence images showing the binding of wheat germ agglutinin-647 (10 µg/mL) to N,N'-diacetylchitobiose, N,N',N"-triacetylchitotriose, N,N'-di-(difluoroacetyl)chitobiose and N,N',N"-tri-(difluoroacetyl)chitotriose. Histograms show the RFU values averaged over 4 spots, and their standard deviation.

rise to a total of four ^{19}F signals per difluoroacetamide moiety. Yet, this homonuclear coupling is comparable to the frequency separation between both coupled ^{19}F signals (i.e., $^{2}J_{FF} \approx \Delta\nu_{FF}$) and produces a strong *roof effect* in the ^{19}F spectra, in which the inner (proximal) doublet lines remain intense, while the outer (distal) lines are much weaker (see Figures S1 and S7). Despite this fortunate simplifying consequence one must, however, consider the $^{2}J_{FF}$ coupling to extract the correct frequency of the pertaining ^{19}F signal (centered between both doublet lines). Of note, homonuclear ^{19}F decoupling during FID acquisition appears unfeasible given the small separation $\Delta\nu_{FF}$ between both geminal ^{19}F signals. An alternative would be the separation of both doublet lines via combinations of inphase and antiphase spectra (IP/AP). In an indirect ^{19}F dimension, true homonuclear decoupling can be achieved by constant time evolution.

After completing the NMR signal assignment we proceeded to study the interaction of the disaccharide **5** with WGA by recording both ^{1}H and ^{19}F spectra. A comparison of the ^{1}H NMR spectra of **5** alone and in the presence of WGA (Figure 3 and Fig S2 (with STD data) in the SI) immediately reveals a site-selective molecular recognition process through substantial signal broadening for the difluoroacetamide hydrogen at the non-reducing end. This indicates chemical exchange between free and bound ligand in the intermediate regime of the NMR timescale suggesting a dissociation constant K_D in the micromolar range, similar to that reported (0.19 mM) for the parent $(GlcNAc)_2$ saccharide.[19] In line with this, the ^{1}H signal of the difluoroacetamide hydrogen at the reducing end saccharide is essentially unaffected. Thus, the CHF_2 probe readily reveals preferred interaction of the non-reducing sugar moiety and, thus, immediately provides epitope information.

In the ^{19}F NMR spectrum of **5** (Figure 4), signal broadening due to reversible interaction is even more extensive than in the ^{1}H spectrum recorded at the same ligand-to-WGA ratio. Besides chemical exchange, chemical shift anisotropy of ^{19}F may also contribute to the observed line width. At low ligand excess (up to ca. 10:1) the ^{19}F signals are even broadened beyond observability (see supplementary Figure S4). Most notably, the difluoroacetamide ^{19}F signals are broadened not only at the non-reducing, but also (although more weakly) at the reducing end of the disaccharide. This observation underlines the known higher susceptibility of ^{19}F to changes in the local environment as compared to ^{1}H, making it the more sensitive nucleus for monitoring molecular interactions. Inversely, ^{1}H appears to provide higher selectivity for the binding epitope. Thus, the difluoroacetamide moiety provides a unique combination of complementary NMR probes to observe molecular recognition

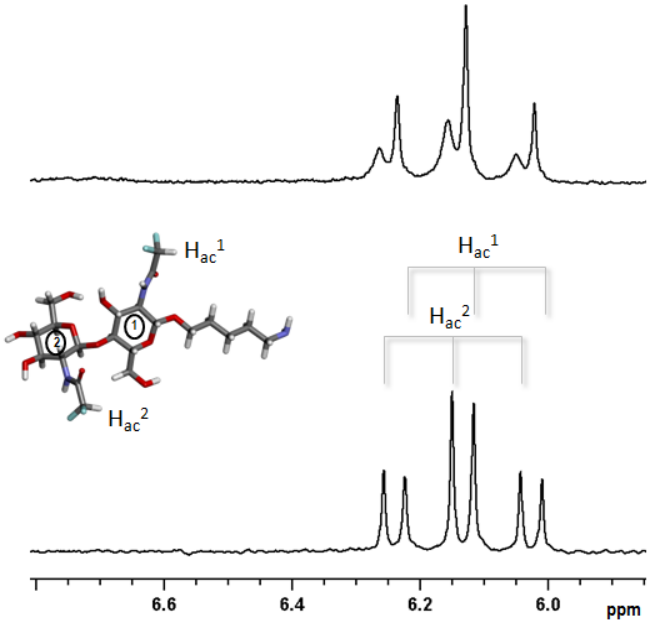


Figure 3. Interaction of **5** with WGA observed by ^{1}H NMR (at 500 MHz, 298 K). Only the well separated spectral region of the difluoromethyl signals is shown. Bottom: ^{1}H spectrum of the free ligand (1.35 mM in deuterated PBS, pH 6). Top: ^{1}H spectrum after addition of WGA at a 55:1 ligand:lectin molar ratio. Comparison of the spectra clearly reveals interaction between both molecules through selective line broadening for the triplet signal of the CHF_2 moiety at the non-reducing saccharide (in contrast, the CHF_2 at the reducing end is not affected at this molar ratio).

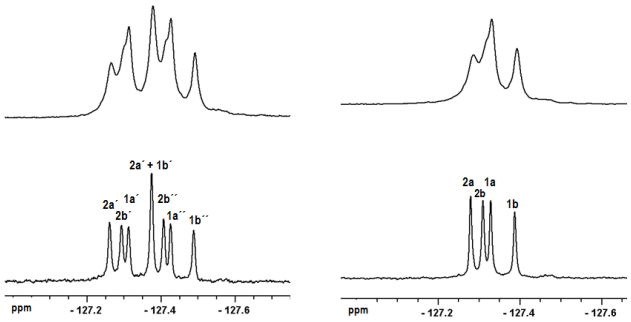


Figure 4. Interaction of **5** with WGA observed by ^{19}F NMR (at 500 MHz, 298 K), without (left) and with ^{1}H decoupling (right). Bottom: ^{19}F spectrum of the free ligand (1.35 mM in deuterated PBS, pH 6). Top: ^{19}F spectrum after addition of WGA at a 55:1 ligand:lectin molar ratio. The signal assignment for both diastereopotic ^{19}F signals (a) and (b) of the difluoroacetamide groups on the reducing (1) and non-reducing (2) saccharide is indicated.

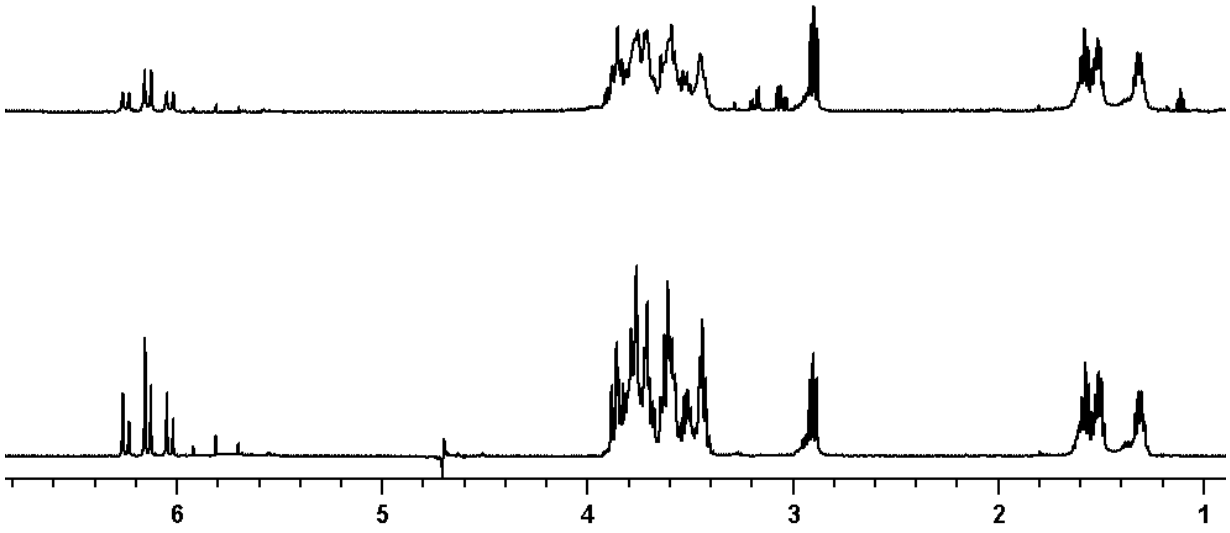


Figure 5. The interaction of **7** with WGA from the ^{1}H NMR perspective. The bottom spectrum corresponds to the free ligand (0.150 mM in deuterated PBS, pH 6, 298 K, 500 MHz). The top spectrum corresponds to a 44:1 ligand molar excess. The data clearly evidence the existence of interaction with WGA with the ligand. Paying attention to the behaviour of the hydrogens within the CHF_2 entities, it is evident that those at the non-reducing end and at the central unit are markedly decreased versus that at the reducing unit.

events with highest sensitivity and epitope selectivity, via ^{19}F and ^{1}H spectra, respectively. The same NMR protocol was applied to monitor the WGA interaction of trisaccharide **7**. The presence of three CHF_2 moieties causes increased spectral crowding, especially in the ^{1}H spectrum (Figure 5) where the signals of CHF_2 groups 2 and 3 cannot be resolved. Contrarily, all three CHF_2 groups can still be distinguished and assigned in the ^{19}F spectrum (Figure 6) using the cited homo- and heteronuclear experiments (supporting Figures S5-S7). In the presence of WGA we again observe a more significant attenuation of ^{1}H signals for the unresolvable CHF_2 moieties of the central (2) and non-reducing end (3) sugar residues (Figure 6), suggesting the same epitope and site-selectivity as with disaccharide **5**.

^{19}F signal broadening and attenuation in the presence of WGA (Figure 6) is also very substantial and conspicuous over a wide range of ligand-to-protein ratios, comprising the signals on all CHF_2 moieties as with disaccharide **5**. Again, while the site-selectivity is not as readily detectable as by ^{1}H NMR, both ^{19}F signals of the reducing-end saccharide (1) are the least affected (this discrimination is particularly evident between 25:1 and 44:1 molar ratios, see Figure 6). Chemical exchange in the intermediate regime suggests that the K_D is in the low micromolar range, similar to that reported for the parent $(GlcNAc)_3$ trisaccharide.[20] Indeed, contrary to the observations for **5** (Fig S4B in SI), STD experiments[21] for this system at any ligand:lectin molar ratio were unsuccessful (Fig. S4B). This fact represents a strong indication of the existence of a slower dissociation for **7,** strongly suggesting that **7** is recognized by WGA with higher affinity than disaccharide **5**, as occurs for the natural analogues (more than two-fold difference, 0.09 mM versus 0.19 mM).

The reversible binding of **7** to WGA is strong enough to also be detected in a ^{19}F-DOSY experiment,[22] where the diffusion coefficient of the free ligand drops by 50% (-0.3 log units) after addition of WGA at 10:1 ligand excess (Figure 7).

In summary, the exocyclic difluoroacetamide moiety is an excellent dual NMR probe to monitor protein-carbohydrate interactions for NAc-containing sugars. While ^{19}F NMR observation offers highest sensitivity with regard to affinity, complementary ^{1}H NMR observation is more site-selective for the binding epitope.

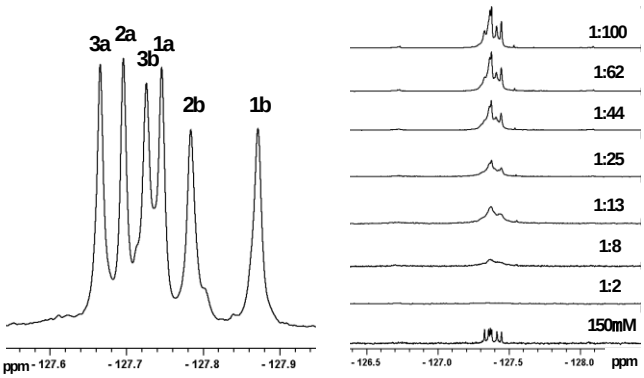


Figure 6. Interaction of **7** with WGA observed by ^{19}F NMR (without ^{1}H decoupling; 500 MHz, 298 K). Left: ^{19}F spectrum of the free trisaccharide (0.150 mM in deuterated PBS, pH 6) with assignment of the diastereotopic signal pairs from the CHF_2 moieties attached to the reducing (1), central (2) and non-reducing sugar (3). Right: evolution of the ^{19}F spectrum of **7** for increasing ligand-to-WGA ratios (from bottom to top), as indicated, with the free ligand spectrum shown below. Reversible binding to WGA in the intermediate exchange regime is indicated by ^{19}F signal broadening for all three CHF_2 moieties, with least perturbation at the reducing end (1).

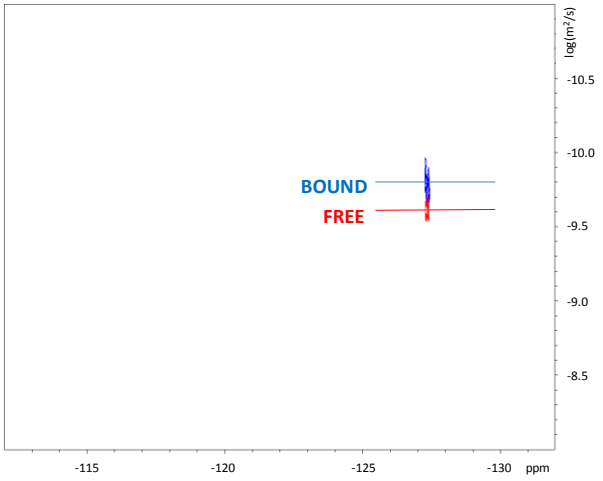


Figure 7. Interaction of **7** with WGA observed by ^{19}F-DOSY experiments (at 500 MHz, 298 K). Superposition of the spectra for the trisaccharide free (red, 0.5 mM in deuterated PBS, pH 6) and in the presence of WGA (blue, at 10:1 ligand:lectin ratio) reveals a 50% (-0.3 log units) decrease in diffusion coefficients, suggesting reversible binding in the micromolar range.

Impact of difluoroacetamide substitution on molecular interactions.

After successful use of the difluoroacetamide moiety as a dual NMR probe for molecular interactions, we analysed the possible impact that this chemical modification could introduce in the intermolecular forces responsible for the binding process. Since the interaction of WGA and hevein domains with GlcNAc-containing sugars has been widely investigated,[17, 23] we adopted a molecular modelling protocol to quantify the intermolecular forces between WGA and modified chitin analogues containing two CHF_2 groups.

First, we employed simple models to understand the interactions in the native sugar/lectin complex at an *ab initio* level, using the Gaussian 09 program package.[24] For this we composed energy optimised models of the complex formed between CH_3-NH-CO-CF_3, CH_3-NH-CO-CHF_2, CH_3-NH-CO-CH_2F, or CH_3-NH-CO-CH_3 with simple aromatic models, as benzene (the basic model) and p-methylphenol (as mimic of the aromatic moieties in Tyr), using the hybrid meta exchange-correlation functional M06-2X at 6-31G(d,p) level of theory,[25] with the counterpoise correction.[25]

The initial structures were prepared with the CF_xH_{3-x} moiety stacked on top of the aromatic ring to form CH/π[26] or CF/π interactions.[27, 28] For the CHF_2 and CH_2F groups, both possible CH/π or CF/π interactions were considered.

The results show the same trends for both aromatic moieties. For the benzene complexes, the lowest energy interaction occurs for -CHF_2, followed by -CH_2F (only 0.1 kcal/mol less stable), always with the C-H bond pointing towards the aromatic ring plane. The alternative orientation, with the C-F bond oriented towards the benzene moiety, is destabilized by more than 2.5 kcal/mol. The impact of C-H (electron acceptor) polarization[29] on the interaction with the aromatic ring (electron donor) is indicated by the calculations showing that the interaction of CHF_2 with benzene is 1 kcal/mol more favourable than for the regular methyl group. The π-interaction energy is the lowest for the –CF_3 group (1.6 kcal/mol less stable than for –CH_3) that can only interact via a disfavoured C-F instead of a C-H bond. Results for the π-interaction with p-methylphenol, used as a mimic of the tyrosine side chain, are very similar (Table 1).

The computed energies corroborate that the -CHF_2 moiety properly mimics the native methyl group in interactions with aromatic systems. The geometry of the formed CH/π complexes is also very similar, although the electron withdrawing fluorine atoms slightly shorten (by ca. 0.15 Å) the CH/π distance for the -CHF_2 and –CH_2F complexes, in line with their higher interaction energy. As also expected from the computed energies, CF/π are much larger than CH/π distances (by 0.4-0.5 Å). A depiction of the different π-interaction geometries is provided in the supporting Figure S8.

In conclusion, the energy calculations predict that the CH/π-interaction of benzene or p-methylphenol with the strongly polarized single C-H bond within the a -CHF_2 moiety is somewhat better than for a methyl group, and significantly better than for a –CF_3 group (that can only partake in a CF/π interaction). Thus, the computed π-interaction energy and geometry further emphasize the potential of the difluoroacetamide moiety as an excellent mimic of the acetamide group in molecular recognition studies.

Table 1. Interaction energies (uncorrected and corrected with BSSE, also including implicit solvation) for the complexes formed between benzene and the model compound N-methyl (fluoro)acetamide.

Molecule, π-Interaction type	E_{int} (kcal/mol)	E_{BSSE} (hartree)	E_{BSSE} (kcal/mol)	E_{corr} (kcal/mol)	E_{int} (kcal/mol) (in H_2O)
N-Me-acetamide, CH-π	–7.0	0,00339	2.1	–4.8	–5.2
N-Me-fluoroacetamide, CH/π	–8.9	0,00522	3.3	–5.6	–6.4
N-Me-difluoroacetamide, CH/π	–8.7	0.00467	2.9	–5.7	–6.7
N-Me-fluoroacetamide, CF/π	–6.2	0.00519	3.3	–3.0	–4.7
N-Me-difluoroacetamide, CF/π	–6.4	0.00552	3.5	–2.9	–4.9
N-Me-trifluoroacetamide, CF/π	–6.7	0.00576	3.6	–3.2	–5.3

The strongest π-interaction results for the N-Me-difluoroacetamide/benzene complex through CH/π stacking that is superior to the alternative CF/ π stacking (see also supporting Figure S8). Explicit consideration of the solvent (H_2O) does not change the general trends, except between fluoroacetamide and difluoroacetamide that always yield very similar results.

Finally, we composed 3D models of the interaction between the disaccharide or trisaccharide analogues with one of the WGA hevein domains,[29] based on our previously reported complex structures of N,N´-diacetyl chitobiose[30] and N,N',N''-triacetyl chitotriose with hevein and the B-domain of WGA,[31] and the X-ray structure of the WGA/N,N´-diacetyl chitobiose complex (PDB: 1K7U). These templates were adjusted to account for the presence of the difluoroacetamide instead of acetamide group.

The molecular recognition features of chitooligosaccharides by hevein domains are well established[32], and for N,N´-diacetyl chitobiose[33] involve interactions of the nonreducing GlcNAc at hevein subsite +1 with the aromatic amino acid at position 23 (via CH/π stacking), with Ser 19 (via hydrogen bond), and with Tyr 30 (via CH/π stacking and hydrogen bond). The reducing sugar interacts at hevein subsite +2, with the Trp 21 (CH/π stacking). On passing from chitobiose to chitotriose, however, additional dynamic features determine the molecular interaction, and the chitin trimer is recognized[34] in two different manners (Figure 1). In binding mode A (see Figure 1), the terminal nonreducing GlcNAc moiety is positioned at hevein subsite +1, and the adjacent central saccharide at subsite +2. The terminal reducing GlcNAc provides very few additional contacts with the lectin, at subsite +3. This binding mode resembles that of the simple disaccharide. In the alternative binding mode B (see Figure 1), however, the terminal reducing sugar ring is placed at subsite +2, interacting with Trp 21, while the central sugar residue localises to subsite +1, and interacts with the aromatic amino acid 23 (via CH/π stacking), Ser 19 (by hydrogen bond), and Tyr 30 (by CH_3/π stacking and hydrogen bond). Here, the terminal non-reducing sugar contacts the lectin at subsite −1.

For a 3D structure of the complexes we then subjected the models to energy minimization and molecular dynamics simulations with the AMBER[35] force field as implemented in the Maestro software suite.[36] The results are in full agreement with our NMR observations and *ab initio* calculations, showing that CH/π always prevail over alternative CF/π interactions.[26, 27] In the modeled complex with the difluoroacetamide containing disaccharide, the binding mode is identical to that described above for unmodified $(GlcNAc)_2$. The -CHF_2 moiety at the reducing end Glc cannot establish major contacts with the lectin since it is oriented towards the solvent. Contrarily, the -CHF_2 moiety at the non-reducing Glc establishes a CH/π interaction with Tyr 30. For the trisaccharide, both alternative binding modes A and B were also stable during the MD run. In mode A, interactions are analogous to those described above. The non-reducing saccharide is accommodated in the binding site in such a way that its -CHF_2 moiety makes CH/π- stacking contacts with Tyr 30; the other two -CHF_2 moieties on the central and reducing-end sugars show only weak or no interactions with lectin sidechains, respectively. Binding mode B contrarily shows CH/π-stacking between Tyr 30 and the -CHF_2 group on the central Glc. The -CHF_2 moieties at both terminal residues make minor contacts with the lectin. Both binding modes A and B are required to explain our ^{1}H as well as ^{19}F-based NMR results, where both ^{1}H and ^{19}F NMR signals of the -CHF_2 groups on the non-reducing and central sugars are primarily affected, while those from the reducing-end -CHF_2 group show much less signal broadening and attenuation.

Conclusions

In conclusion, we have presented a proof-of-concept on a novel chemical tag, difluoroacetamide, and pertaining detection method for glycan-protein interactions by NMR. While the tag had already been employed as a substituent in medicinal chemistry studies, we demonstrate its exceptional suitability also as a sensor for molecular recognition. The presence of both ^{1}H and ^{19}F atoms within the difluoroacetamide tag enables easy and robust ^{1}H and/or ^{19}F-based NMR detection of molecular recognition processes involving the relevant GlcNAc- moieties over a wide range of affinities. Contrary to the widely employed trifluoroacetamide group, the proposed difluoroacetamide tag still contains one methyl proton. This not only allows for ^{1}H NMR detection, but also for the critical CH/π interactions that are much more stable than analogous CF/π interactions. Moreover, the two geminal fluorine atoms intensify the CH/π interaction by polarising the C-H bond, and force it into a unique orientation for lack of alternative protons.

Thus, the CHF_2CONH- moiety partakes in the same interactions as the natural acetamide fragment, providing similar or even better binding energy and geometry than the native acetamide group, and much better than the perfluorinated CF_3CONH-analogue.

We believe that this tag will certainly be applicable to most GlcNAc-binding lectins, and especially to those that display aromatic/methyl acetamide stacking interactions. The possible applications to other fluoroacetamide-containing tags, other glycans (GalNAc, NeuNAc), and other lectin/glycan systems are currently under exploration.

Experimental Section

NMR. All NMR experiments were recorded at 298 K on a 500 MHz BRUKER Avance spectrometer equipped with a ^{19}F,^{1}H SEF dual probe optimised for direct ^{19}F detection. Complete ^{1}H signal assignment of the ligands was obtained from standard TOCSY (60 ms mixing time), NOESY (300 and 500 ms mixing time) and ^{1}H, ^{13}C-HSQC experiments.[37] Ligand concentrations typically varied between 1 and 2 mM. ^{19}F signals were then assigned from two-dimensional heteronuclear ^{1}H,^{19}F and homonuclear ^{19}F,^{19}F correlation experiments.

Computational Details. All DFT calculations were performed using the Gaussian 09 program package[24]. Geometry optimizations employed the hybrid meta exchange-correlation function M06-2X[25] at the 6-31G(d,p) level of theory. Benzene and 4-methylphenol were chosen as representative models for phenylalanine and tyrosine, respectively. In all cases, no symmetry constraints were enforced. The vibrational frequency analysis was carried out at the same level of theory to ensure that the geometry obtained corresponds to an energy minimum, and this was ensured by checking for the absence of negative eigenvalues (imaginary frequencies). Solvent effects (H_2O) were evaluated by applying the polarizable continuum model (PCM) with the integral equation formalism variant (IEFPCM).

Energies (kcal/mol) for CH/π and CF/π interactions, both in solution and in gas phase, were calculated as:

$$\Delta E_{int} = E_{complex} - E_{amide} - E_{Ar}$$

and for corrected energies:

$$\Delta E_{corr} = \Delta E_{int} + E_{BSSE}$$

where E_{amide} and E_{Ar} are the calculated monomer energies at the M06-2X/6-31G (d,p) level for amide and aromatic groups, respectively. Finally, single point energy calculations with counterpoise (CP) corrections for optimized amide-benzene complexes were performed to minimize the basis set superposition error (BSSE). These values are shown in Table 1.

Molecular modelling. Starting structures for the docking were composed from the published X-ray or NMR structures of hevein and WGA [PDB codes: 1K7U, 7WGA, 2UVO, 1T0W, 1Q9B] and manually modified by replacing the acetamide for difluoroacetamide groups. Their orientation was chosen to enable CH/π (not CF/π) interaction with Tyr 30 in the lectin. These initial structures were submitted to a short MD run (1 ns), followed by energy minimization with a low gradient convergence threshold (0.02) in 5000 steps. The AMBER force field integrated in the MAESTRO software suite[36] was employed.

Chemical Synthesis. General methods: Chemicals were purchased from Sigma-Aldrich or Acros Organics, and were used without further purification. WGA was purchased from Sigma-Aldrich. Difluoroacetic anhydride was purchased from Fluorochem (Derbyshire, United Kingdom). Thin layer chromatography was carried out with Merck aluminium sheets silica gel 60 F_{254}, and visualized by UV irradiation (254 nm) or vanillin staining. A Biotage Initiator monomode oven (Biotage AB, Uppsala, Sweden) was used for microwave irradiation. For hydrogenation we employed a H-Cube® continuous-flow hydrogenation reactor with 10%Pd/C CatCart® holder from ThalesNano NanoTechnology Inc., Budapest, Hungary. Compounds were purified by conventional flash chromatography on Merck 62Å 230-400 mesh silica gel, and size-exclusion chromatography on Biorad P2 gel, Biorad, Hercules, USA. Pooled glycan containing fractions were lyophilized on an ALPHA-2-4 LSC freeze-dryer from Christ, Osterode, Germany. All organic solvents were concentrated by rotary evaporation. ^{1}H and ^{13}C spectra were acquired on a Bruker AVANCE 500 MHz spectrometer, and chemical shifts (δ, in ppm) were referenced to the residual signal of the solvent. Signal multiplicities are designated as s (singlet), d (doublet), t (triplet) or m (multiplet). Scalar coupling constants (*J*) are reported in Hz. The mass spectrometric data was obtained on a Waters LCT Premier XE (Waters, Manchester, UK) equipped with a standard ESI source for direct injection, and operated with a capillary voltage of 1.0 kV and cone voltage of 200 V. Cone and desolvation gas flows were set to 50 and 500 L/h, respectively; the source and desolvation temperature was 100 ºC. The exact mass was determined using glycocholic acid (Sigma) as internal standard (2 M+Na$^+$, m/z = 953.6058).

3,6-di-O-benzyl-2-deoxy-4-O-levulinyl-2-phthalimido-1-thio-β-D-glucopyranosyl trichloroacetimidate (1). To a solution of ethyl-3,6-di-O-benzyl-2-deoxy-2-phthalimido-1-thio-β-D-glucopyranoside[38] (5.66 g, 10.61 mmol) in CH_2Cl_2 (25 mL), N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (4.05 g, 21.12 mmol, 2 eq), 4-dimethylaminopyridine (907 mg, 7.43 mmol, 0.7 eq) and levulinic acid (2.1 mL, 20.51 mmol, 1.93 eq) were added. The reaction mixture was stirred overnight at room temperature, quenched with water, and extracted three times with CH_2Cl_2. The combined organic phases were washed with saturated $NaHCO_3$, brine and dried over anhydrous $MgSO_4$. The solvent was removed by rotary evaporation, and the resulting crude was purified by flash column chromatography (hexane:EtOAc, 2:1) to give ethyl-3,6-di-O-benzyl-2-deoxy-4-O-levulinyl-2-phthalimido-1-thio-β-D-glucopyranoside as a white solid (6.22 g, 93%). CD: $[\alpha]_D^{20}$ = +37.4 (c=0.5, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 – 7.60 (m, 4H, Phth), 7.43 – 7.27 (m, 5H, Ph), 7.05 – 6.83 (m, 5H: Ph), 5.27 (d, *J* = 10.5 Hz, H-1), 5.15 (dd, *J* = 10.1 Hz, 9.0 Hz, H-4), 4.66 (d, *J* = 12.3 Hz, 1H, CH_2 Bn), 4.59 – 4.51 (m, 2H, 2 × C*H*$_2$ Bn), 4.46 (dd, *J* = 10.3 Hz, 8.9 Hz, H-3), 4.36 – 4.27 (m, 2H, H-2, C*H*$_2$ Bn), 3.84 – 3.76 (m, H-5), 3.63 (m, *J* = 4.7 Hz, 1.9 Hz, 2H, H-6a, H-6b), 2.74 – 2.56 (m, 4H, CH_2 Lev, CH_2 EtS), 2.45 (t, *J* = 6.3 Hz, 2H, CH_2 Lev), 2.14 (s, *J* = 1.9 Hz, 3H, CH_3 Lev), 1.18 (t, *J* = 7.5 Hz, 3H, CH_3 EtS). ^{13}C NMR (126 MHz, $CDCl_3$) δ 206.2, 171.6, 168.1, 167.2, 138.1, 137.8, 133.9, 133.1, 131.6, 128.3, 128.1, 127.8, 127.6, 127.4, 123.6, 123.3, 81.2 (C-1), 78.0 (C-3), 77.7 (C-5), 74.2 (CH_2 Bn), 73.5 (CH_2 Bn), 72.8 (C-4), 69.8 (C-6), 54.6 (C-2), 37.7, 29.8, 27.9, 24.1, 14.9. HRMS (ESI): *m/z* calcd for $C_{35}H_{37}NO_8SNa$ [M+Na]$^+$ 654.2137, found 654.2105.To a solution of ethyl-3,6-di-O-benzyl-2-deoxy-4-O-levulinyl-2-phthalimido-1-thio-β-D-glucopyranoside (217 mg, 0.34 mmol) in CH_2Cl_2 (3 mL) at 0°C, N-iodosuccinimide (154 mg, 0.69 mmol, 2 eq) and trifluoroacetic acid (90 µL, 0.69 mmol, 2 eq) were added. After 2h of stirring at room temperature, $Na_2S_2O_3$ and saturated solution of $NaHCO_3$ were added. The organic layer was washed with saturated $NaHCO_3$, brine, and dried over anhydrous $MgSO_4$. The crude was concentrated and purified over a short silica plug (hexane: EtOAc = 3:2) to yield the corresponding hemiacetal as a white solid. This compound was dissolved in CH_2Cl_2, cooled to 0 ºC, and DBU (5 µL, 0.03 mmol, 0.1 eq) and trichloroacetonitrile (490 µL 4.85 mmol, 20 eq) were added. The reaction was stirred at room temperature for 1h. The crude was purified by flash column chromatography (Hexane: EtOAc, 3:2 containing 1% Et_3N) to give compound **1** as a white solid (202 mg 80% in 2 steps). CD: $[\alpha]_D^{20}$ = +44.3 (c=0.5, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.57 (s, NH), 7.73 – 7.64 (m, 4H, Phth), 7.40 – 7.27 (m, 5H, Ph), 7.07 – 6.83 (m, 5H, Ph), 6.43 (d, J = 8.4 Hz, H-1), 5.28 (dd, J = 10.0 Hz, 8.1 Hz, H-4), 4.69 (d, J = 12.2 Hz, CH_2 Bn), 4.60 – 4.50 (m, 4H, H-2, H-3, 2 × CH_2 Bn), 4.36 (d, J = 12.3 Hz, CH_2 Bn), 3.96 (ddd, J = 10.0 Hz, 5.0 Hz, 3.3 Hz, H-5), 3.73 – 3.61 (m, 2H, H-6a, H-6b), 2.72 – 2.57 (m, 2H, CH_2 Lev), 2.52 – 2.39 (m, 2H, CH_2, Lev), 2.14 (s, 3H, CH_3 Lev). ^{13}C NMR (126 MHz, $CDCl_3$) δ 206.3, 171.6, 160.9, 138.1, 137.9, 134.1, 131.6, 128.5, 128.2, 128.1, 128.0, 127.8, 127.5, 123.5, 94.1 (C-1), 74.6 (C-5), 74.2, 73.6, 72.3 (C-4), 69.1 (C-6), 54.6 (C-2), 37.8, 29.9, 28.0. HRMS (ESI): m/z calcd for $C_{35}H_{33}Cl_3N_2O_9Na$ [M+Na]$^+$ 755.1129, found 755.1056.

5-azidopentyl 3,6-di-O-benzyl-2-deoxy-4-O-levulinyl-2-phthalimido-β-D-glucopyranosyl-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (3). A mixture of acceptor **2**[39] (150 mg, 0.25 mmol), donor **1** (219 mg, 0.30 mmol, 1.2 eq) and 3Å molecular sieves in dry CH_2Cl_2 (2.5 mL) was stirred under argon for 45 min at room temperature. The solution was cooled to 0°C, and trimethylsilyl trifluoromethanesulfonate (4.5 µL, 0.025 mmol, 0.1 eq) was added dropwise. After 2 hours stirring at room temperature, the reaction mixture was quenched with triethylamine, filtered through a celite plug, and concentrated. The crude was purified by flash colum chromatography (hexane: acetone, 3:1) to yield the title compound as a white solid (193 mg, 66% yield). CD: $[\alpha]_D^{20}$= +17.2(c=0.5, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.51 (m, 8H, Phth), 7.38 – 7.26 (m, 10H, 2 × Ph), 7.03 – 6.79 (m, 10H, 2 × Ph), 5.32 (d, *J* = 8.4 Hz, H-1'), 5.17 (dd, *J* = 10.4 Hz, 8.3 Hz, H-4'), 4.93 (d, *J* = 8.3 Hz, H-1), 4.83 (d, *J* = 12.6 Hz, C*H*$_2$ Bn), 4.65 (d, *J* = 12.2 Hz, C*H*$_2$ Bn), 4.55 – 4.39 (m, 6H, 5 × C*H*$_2$ Bn, H-3'), 4.32 (d, *J* = 12.2 Hz, C*H*$_2$ Bn), 4.26 (dd, *J* = 10.8 Hz, 8.3 Hz, H-2'), 4.18 – 4.06 (m, 3H, H-2, H-3, H-4), 3.65 (dt, *J* = 9.9 Hz, 6.0 Hz, C*H*$_2$O), 3.59 – 3.52 (m, 2H, H-6a', H-5'), 3.53 – 3.43 (m, 2H, H-6a, H-6b'), 3.39 (dd, *J* = 11.0 Hz, 4.1 Hz, H-6b), 3.31 (m, *J* = 8.4 Hz, 4.0 Hz, H-5), 3.24 (dt, *J* = 9.9 Hz, 6.7 Hz, C*H*$_2$O), 2.93 – 2.80 (m, 2H, CH_2N_3), 2.69 – 2.50 (m, 2H, CH_2 Lev), 2.50 – 2.34 (m, 2H, CH_2 Lev), 2.11 (s, 3H, CH_3 Lev), 1.45 – 1.26 (m, 4H, 2 × CH_2 linker), 1.06 (ddt, J = 15.7 Hz, 8.8 Hz, 6.3 Hz, 2H, CH_2 linker). ^{13}C NMR (126 MHz, $CDCl_3$) δ 206.3, 171.6, 138.6, 138.5, 138.4, 138.0, 134.1, 133.8, 131.8, 131.8, 128.4, 128.36, 128.2, 128.1, 128.0, 127.8, 127.8, 127.5, 127.5, 127.4, 127.0, 123.3, 98.2 (C-1), 97.3 (C-1'),

77.0 (C-3'), 76.4 (2C, C-3, C-4), 74.7 (C-5), 74.5 (CH_2 Bn), 74.1 (CH_2 Bn), 73.6 (CH_2 Bn), 73.4 (C-5'), 73.1 (C-4'), 72.8 (CH_2 Bn), 69.5 (C-6'), 68.9 (CH_2O), 68.3 (C-6), 56.4 (C-2'), 55.8 (C-2), 51.2 (CH_2N_3), 37.8 (CH_2 Lev), 29.9 (CH_3 Lev), 28.8 (CH_2), 28.4 (CH_2), 28.1 (CH_2 Lev), 23.1 (CH_2). HRMS (ESI): m/z calcd for $C_{66}H_{67}N_5O_{15}Na$: 1192.4531 $[M+Na]^+$, found 1192.4456.

5-azidopentyl 3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (4). To a solution of 3 (141 mg, 0.12 mmol) in CH_2Cl_2 (1.2 mL), hydrazine acetate (17 mg, 0.18 mmol, 1.5 eq) dissolved in MeOH (0.12 mL) was added. After 1h stirring at room temperature, the solvents were evaporated. The crude product was purified by flash column chromatography (hexane: EtOAc, 3:2), to yield the title compound as a white solid (108 mg, 84%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.94 – 7.58 (m, 8H, 2 × Phth), 7.42 – 7.27 (m, 10H, 2 × Ph), 7.06 – 6.80 (m, 10H, 2 × Ph), 5.30 (d, *J* = 8.2 Hz, H-1'), 4.93 (d, *J* = 7.8 Hz, H-1), 4.78 (d, *J* = 12.4 Hz, 2H, 2 × C*H*$_2$ Bn), 4.55 – 4.44 (m, 6H, 6 × C*H*$_2$ Bn), 4.25 (dd, *J* = 10.8 Hz, 8.3 Hz, H-3'), 4.19 – 4.05 (m, 4H, H-2', H-4, H-3, H-2), 3.81 (dd, *J* = 9.6 Hz & 8.2 Hz, H-4'), 3.70 (dd, *J* = 10.0 Hz, 4.4 Hz, H-6a'), 3.68 – 3.63 (m, C*H*$_2$O), 3.57 – 3.51 (m, 2H, H-6b', H-6a), 3.45 – 3.35 (m, 2H, H-6b, H-5'), 3.34 – 3.28 (m, H-5), 3.24 (m, C*H*$_2$O), 2.94 – 2.80 (m, 2H, CH_2N_3), 1.44 – 1.23 (m, 4H, 2 × CH_2 linker), 1.14 – 1.00 (m, 2H, CH_2 linker). ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6, 167.8, 138.8, 138.5, 138.4, 137.6, 133.8, 128.7, 128.3, 128.2, 128.04, 127.97, 127.9, 127.9, 127.5, 127.4, 127.0, 123.8, 123.3, 98.2 (C-1), 97.1 (C-1'), 78.4 (C-3'), 75.9 (C-3), 75.6 (C-4), 74.7 (C-4'), 74.5 (C-5), 74.3 (CH_2 Bn), 73.8 (CH_2 Bn), 72.9 (C-5'), 72.8 (CH_2 Bn), 71.1 (C-6'), 69.0 (CH_2O), 68.4 (C-6), 56.2 (C-2'), 55.8 (C-2), 51.2 (CH_2N_3), 28.8, 28.4, 23.1.

5-azidopentyl (3,6-di-O-benzyl-2-deoxy-4-O-levulinyl-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (6). A mixture of acceptor 4 (120 mg, 0.112 mmol), donor 1 (138 mg, 0.189 mmol, 1.7 eq) and 3Å molecular sieves in dry CH_2Cl_2 was stirred for 45 minutes at room temperature. The mixture was cooled to 0°C, and TMSOTf (2 µL, 0.001 mmol, 0.10 eq) was added. After 3 hours stirring at room temperature, the reaction mixture was quenched with triethylamine, filtered through a celite plug, and concentrated. The crude was purified by flash column chromatography (hexane: EtOAc, 3:2) to yield the title compound as a white solid (95 mg 50%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.96 – 7.48 (m, 12H, 3 × Phth), 7.37 – 7.09 (m, 15H, 3 × Ph), 7.10 – 6.80 (m, 15H, 3 × Ph), 6.74 – 6.62 (m, 3H, Ph), 5.31 (d, *J* = 8.3 Hz, H-1''), 5.16 (t, *J* = 9.1 Hz, H-4''), 5.08 (d, *J* = 7.5 Hz, H-1'), 4.91 – 4.83 (m, 2H, H-1, 2 × C*H*$_2$ Bn), 4.73 (d, *J* = 12.9 Hz, C*H*$_2$ Bn), 4.65 (d, *J* = 12.2 Hz, C*H*$_2$ Bn), 4.54 (d, *J* = 12.2 Hz, C*H*$_2$ Bn), 4.51 – 4.40 (m, 4H, 3 × C*H*$_2$ Bn, H-3''), 4.40 – 4.28 (m, 5H, 5 × C*H*$_2$ Bn), 4.25 (dd, *J* = 10.7 Hz, 8.4 Hz, H-2''), 4.21 – 4.09 (m, 3H, H-2', H-3', H-4'), 4.08 – 3.98 (m, 3H, H-2, H-3, H-4), 3.63 (dt, *J* = 9.9 Hz, 6.1 Hz, C*H*$_2$O), 3.56 – 3.49 (m, 2H, H-6a'', H-5''), 3.49 – 3.39 (m, 2H, H-6b'', H-6a), 3.39 – 3.28 (m, 2H, H-6b, H-6a'), 3.27 – 3.17 (m, 2H, H-5, C*H*$_2$O), 3.12 (dd, *J* = 11.3 Hz, 3.2 Hz, H-6b'), 2.85 (m, 3H, CH_2N_3, H-5'), 2.67 – 2.48 (m, 2H, CH_2 Lev), 2.48 – 2.31 (m, 2H, CH_2 Lev), 2.09 (s, 3H, CH_3 Lev), 1.40 – 1.20 (m, 4H, 2 × CH_2 linker), 1.12 – 0.98 (m, 2H, CH_2 linker). LRMS (ESI): m/z calcd for $C_{94}H_{92}N_6O_{21}Na$, 1663.6213 $[M+Na]^+$, found, 1663.5509.

General Procedure. Preparation of difluoroacetamide derivatives. A solution of protected saccharides in 1,2-ethylenediamine/*n*BuOH (1:4) was heated to 120 ºC under microwawe irradiation (3 cycles, 30 minutes each). The solvent was evaporated and the crude product was dissolved in pyridine, cooled to 0 ºC, and 3 equivalents of difluoroacetic anhydride were added dropwise. The reaction was stirred overnight at room temperature. The reaction mixture was evaporated to dryness, and filtered through a short pad of silica gel (hexane:EtOAc = 3:2). The crude product was dissolved in MeOH (2 mL) and treated with 0.2 mL NaOMe in MeOH. The reaction mixture was stirred for one hour at room temperature, neutralized with Amberlite IR-120 (H), filtered and evaporated. The residue was dissolved in MeOH:water (9:1) containing 1% trifluoroacetic acid, and the solution was hydrogenated by passing twice through a H-Cube reactor at 0.5 mL/minute, 50 ºC and full hydrogen. The reaction mixture was evaporated to dryness. The crude was purified by size exclusion chromatography on Biogel P2, and product containing fractions were pooled and freeze-dried to yield the difluoroacetamides.

5-aminopentyl 2-deoxy-2-difluoroacetamido-β-D-glucopyranosyl-(1→4)-2-deoxy-2-difluoroacetamido-β-D-glucopyranoside (5). Following the general procedure starting from compound **3**, the title compound was obtained as a white solid (29 mg, 56% over four steps). ^{1}H NMR (500 MHz, D_2O) δ 6.24, 6.21 (2 × t, *J* = 53.5 Hz, 2H, $COCHF_2$), 4.70 (d, *J* = 8.4 Hz, H-1'), 4.59 (d, *J* = 8.33 Hz, H-1), 3.98 – 3.89 (m, 2H, H-6a', CH_2O), 3.89 – 3.82 (m, 2H, H-2', H-6a), 3.72 – 3.64 (m, 3H, H-2, H-3, H-6b'), 3.82-3.64 (m, 3H, H-6b, H-4, H-3'), 3.61 (d, *J* = 10.2 Hz, CH_2O), 3.56 – 3.46 (m, 3H, H-5, H-5', H-4'), 2.98 (t, *J* = 7.7 Hz, 2H, CH_2NH_2), 1.67 (m, *J* = 7.6 Hz, 2H, CH_2 linker), 1.63 – 1.55 (m, 2H, CH_2 linker), 1.44 – 1.35 (m, 2H, CH_2 linker). ^{13}C NMR (126 MHz, D_2O) δ 165.6, 108.4 (CHF_2), 100.7 (C-1), 100.5 (C-1'), 79.0 (C-4), 76.0 (C-5'), 74.5 (C-5), 73.0 (C-3'), 71.9 (C-3), 70.2 (CH_2O), 69.7 (C-4'), 60.5 (C-6'), 60.2 (C-6), 55.7 (C-2''), 55.1 (C-2), 39.3 (CH_2NH_2), 28.0 (CH_2), 26.4 (CH_2), 22.1 (CH_2). HRMS (ESI) m/z calcd for $C_{21}H_{35}F_4N_3O_{11}H$: 604.2106 $[M+H]^+$, found 604.2070.

5-aminopentyl 2-deoxy-2-difluoroacetamido-β-D-glucopyranosyl-(1→4)-2-deoxy-2-difluoroacetamido-β-D-glucopyranosyl-(1→4)-2-deoxy-2-difluoroacetamido-β-D-glucopyranoside (7). Following the general procedure starting from compound 6, the title compound was obtained as a white solid (13 mg, 81%).^{1}H NMR (500 MHz, D_2O) 6.39 – 6.04 (m, 3H, 3 × CHF_2), 4.69 (t, *J* = 8.2 Hz, 2H, H-1', H-1''), 4.63 – 4.54 (m, H-1), 4.02 – 3.45 (m, 20H), 2.88 – 2.78 (m, 2H, CH_2NH_2), 1.67 – 1.50 (m, 4H, 2 × CH_2 linker), 1.36 (m, *J* = 8.7 Hz, 7.8 Hz, 2.2 Hz, 2H, CH_2 linker).^{13}C NMR (126 MHz, D_2O) δ 165.9, 165.6, 165.45, 108.5, 108.4, 100.7 (C-1), 100.5 (2C, C-1', C-1''), 78.8, 78.8, 76.1, 74.6, 73.1, 72.0, 71.8, 70.3, 69.8, 60.6, 60.1, 60.0, 55.8, 55.3, 55.2, 39.8, 28.2, 22.3. HRMS (ESI): m/z calcd for $C_{29}H_{46}F_6N_4O_{16}H$: 821.2891 $[M+H]^+$, found 821.2855.

Acknowledgements

We thank the Ministry of Economy and Competitiveness of Spain for financial support (Grants CTQ2012-32025 and CTQ2011-27874). We also thank EU for ITN/ETN Marie Curie Actions Glycopharm, Dynano, Tollerant, Immunoshape, and Euroglycoarrays.

Keywords: Molecular Recognition • glycan • NMR • lectins • chemical tag

[1] D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Jr Smetana, H. J. Gabius, *Biochim Biophys Acta*., **2015**, 1850, 186-235.

[2] a) J. P. Ribeiro, L. K. Mahal, *Curr Opin Chem Biol*., **2013**, *17*, 827–831; b) S. Park, J. C. Gildersleeve, O. Blixt, O. I. Shin, *Chem. Soc. Rev*., **2013**, *42*, 4310-4326.

[3] Ardá, P. Blasco, D. Varón Silva, V. Schubert, S. André, m. Bruix, F. J. Cañada, H. J. Gabius, C. Unverzagt, J. Jiménez-Barbero, *J. Am. Chem. Soc*., 2013; 135, 2667-2675.

[4] H. J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís, *Trends Biochem. Sci*., **2011**, *36*, 298-313.

[5] S. Perez, I. Tvaroška, *Adv. Carbohydr. Chem. Biochem.*, **2014**, *71*, 9-136.

[6] M. Pokorná, G. Cioci, S. Perret, E. Rebuffet, N. Kostlánová, J. Adam, N. Gilboa-Garber, E. P. Mitchell, A. Imberty, M. Wimmerová, *Biochemistry*, **2006**, *45*, 7501-7510.

[7] J. L. Asensio, A. Ardá, F. J. Cañada, J. Jiménez-Barbero, *Acc Chem. Res*., **2013**, *46*, 946-954.

[8] B. Meyer, T. Peters, *Angew. Chem. Int. Ed. Engl.*, 2003, 42, 864-890.

[9] L. Unione, S. Galante, D. Díaz, F. J. Cañada, J. Jiménez-Barbero, *MedChemComm.*, **2014**, *5*, 1280-1289.

[10] a) M. Garavís, B. López-Méndez, A. Somoza, J. Oyarzabal, C. Dalvit, A. Villasante, R. Campos-Olivas, C. González, *ACS Chem. Biol*., **2014**, *9*, 1559-1566; b) A. Vulpetti, Dalvit C., *ChemMedChem*., **2013**, *8*, 2057-2069.

[11] C. Hamark, J. Landström, G. Widmalm, *Chem. Eur. J*., **2014**, *20*, 13905-13908.

[12] Y. Cheng, A.-L. Guo, D.-S. Guo, *Curr. Org. Chem.,* **2010**, *23*, 977-999.

[13] a) M. D. Vaughan, P. Cleve, V. Robinson, H. S. Duewel, J. F. Honek, *J. Am. Chem. Soc.*, **1999**, 121,8475-8478; B. Salopek-Sondi, M. D. Vaughan, M. C. Skeels, J. F. Honek, L. A. Luck, *J. Biomol. Struct. Dyn.* **2003**, 21, 235.

[14] a) H. Suzuki, I. Utsunomiya, K. Shudo, N. Fukuhara, T. Iwaki, T. Yasukata, *Eur. J. Med. Chem*., **2013**, *69*, 262-277; b) A. J. Humphrey, C. Fremann, P. Critchley, Y. Malykh, R. Schauer, T. D. H. Bugg, *Bioorg. Med. Chem*., **2002**, *10*, 3175-3185; c) C. Isanbor, D. O'Hagan, *J. Fluor. Chem*., 2006, 127, 303-319.

[15] C. S. Wright, S. Olafsdottir, *J. Biol. Chem*., **1986**, *261*, 7191-7195.

[16] C. S. Wright, *J. Mol. Biol.,* **1990**, *215*, 635-651.

[17] J. Jiménez-Barbero, F. J. Cañada, J. L. Asensio, N. Aboitiz, P. Vidal, A. Canales, P. Groves, H. J. Gabius, H. C. Siebert, *Adv Carbohydr. Chem. Biochem*., **2006**, *60*, 303-354.

[18] K. Bock, J. O. Duus, *J. Carbohydr. Chem*. **1994**, *13*, 513-546.

[19] a) G. Bains, R. T. Lee, Y. C. Lee, E. Freire, *Biochemistry*, **1992**, 31, 12624-12628. b) M. Muraki, M. Ishimura, K. Harata, *Biochim Biophys Acta*. **2002**, *1569*, 10-20.

[20] A. M, Wu, J. H. Wu, S.-C. Song, M.-S. Tsai, A. Herp A., *Febs Lett*. **1998**, *440*, 315-319.

[21] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed. Engl*. **1999**, *35*, 1784-1786.

[22] P. Groves, M. O. Rasmussen, M. D. Molero, E. Samain, F. J. Cañada, H. Driguez, J. Jiménez-Barbero, *Glycobiology*, **2004**, *14*, 451-456

[23] H. S. G. Beckmann, H. H. Möller, V. Wittmann, J. Beilstein, *Org. Chem.,* **2012**, *8*, 819–826.

[24] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, J., M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Jr. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, M. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian, Inc*., Wallingford CT, **2009**.

[25] a) Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.,* **2008**, *41*, 157-167; b) Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput*., **2006**, *2*, 364-382; c) S. F. Boys, F. Bernardi, *Mol. Phys.,* 1970, 19, 553-566.

[26] M. I. Chávez, C. Andreu, P. Vidal, N. Aboitiz, F. Freire, P. Groves, J. L. Asensio, G. Asensio, M. Muraki, F. J. Cañada, J. Jiménez-Barbero, *Chem. Eur. J.,* **2005**, *11*, 7060-7074.

[27] 26. Xu X, Pooi B, Hirao H, Hong SH., *Angew. Chem. Int. Ed. Engl*., **2014**, *53*, 1283-1287.

[28] L. L. Wu, C. L. Yang, F. C. Lo, C. H. Chiang, C. W. Chang, K. Y. Ng, H. H. Chou, H. Y. Hung, S. I. Chan, S. S. Yu, *Chem. Eur. J*., **2011**, *17*, 4774-4787.

[29] A. G. Santana, E. Jiménez-Moreno, A. M. Gómez, F. Corzana, C. González, G. Jiménez-Oses, J. Jiménez-Barbero, J. L. Asensio, *J. Am. Chem. Soc*., **2013**, *135*, 3347-3350.

[30] J. L. Asensio, F. J. Cañada, M. Bruix, C. González, N. Khiar, A. Rodríguez-Romero, J. Jiménez-Barbero, *Glycobiology*, **1998**, *8*, 569-577.

[31] J. F. Espinosa, J. L. Asensio, J. L. García, J. Laynez, M. Bruix, C. Wright, H. C. Siebert, H. J. Gabius, F. J. Cañada, J. Jiménez-Barbero, *Eur. J. Biochem*., **2000**, *267*, 3965-3978.

[32] J. L. Asensio, H. C. Siebert, C. W. von Der Lieth, J. Laynez, M. Bruix, U. M. Soedjanaamadja, J. J. Beintema, F. J. Cañada, H. J. Gabius, J. Jiménez-Barbero, *Proteins*, **2000**, *40*, 218-236.

[33] J. L. Asensio, F. J. Canada, M. Bruix, A. Rodriguez-Romero, J. Jimenez-Barbero, *Eur J Biochem.*, **1995**, *230*, 621-633.

[34] J. L. Asensio, F. J. Cañada, H. C. Siebert, J. Laynez, A. Poveda, P. M. Nieto, U. M. Soedjanaamadja, H. J. Gabius, J. Jiménez-Barbero, *Chem. Biol*., **2000**, *7*, 529-543.

[35] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Jr Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.*, **1995**, *117*, 5179-5197.

[36] Maestro, version 9.4, Schrödinger, LLC, New York, NY, **2013**.

[37] *NMR spectroscopy of glycoconjugates*, **2002**, J. Jiménez-Barbero, T. Peters, (Eds) Wiley-VCH Verlag GmbH & Co. KGaA.

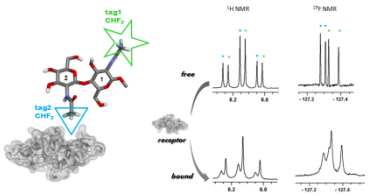
[38] J. Kerckgydrt, J. G. M. van der Ven, J. P. Kamerling, A. Lipt, A; J. F. G. Vliegenthart, *Carbohydr. Res*., **1993**, *238*, 135-145.

[39] S. Serna, B. Kardak, N.-C. Reichardt, M. Martin-Lomas, *Tetrahedron Asymm*., **2009**, *20*, 851-856.

Monitoring glycan/protein interactions

Layout 1:

A simple chemical tag to monitor glycan/protein molecular recognition by NMR has been devised. The chemical nature of the tag permits the easy and straightforward detection of interactions between the partners and allows deducing the glycan binding epitope.



Luis P. Calle, Dr. Begoña Echeverria, Antonio Franconetti, Dr. Sonia Serna, Dr. M. Carmen Fernández-Alonso, Dr. Tammo Diercks,Prof. Dr. F. Javier Cañada, Dr. Ana Ardá, Dr. Niels-Christian Reichardt, Prof. Dr. Jesús Jiménez-Barbero**

Page No. – Page No.

Title