Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202202208
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Check for updates

NMR Investigation of Protein-Carbohydrate Interactions: The Recognition of Glycans by Galectins Engineered with Fluorotryptophan Residues

Marta G. Lete[+],[a] Antonio Franconetti[+],[a] Sara Bertuzzi,[a] Sandra Delgado,[a] Mikel Azkargorta,[a] Félix Elortza,[a] Oscar Millet,[a] Gonzalo Jiménez-Osés,[a, b] Ana Arda,*[a, b] and Jesús Jiménez-Barbero*[a, b, c, d]

Dedicated to Professor Miquel Pericàs for his 71st birthday.

Abstract: Fluorine (^{19}F) incorporation into glycan-binding proteins (lectins) has been achieved and exploited to monitor the binding to carbohydrate ligands by nuclear magnetic resonance (NMR) spectroscopy. Galectins are a family of lectins that bind carbohydrates, generally with weak affinities, through a combination of intermolecular interactions including a key CH-π stacking involving a conserved tryptophan residue. Herein, Galectin-3 (Gal3) and Galectin-8 (Gal8) with one and two carbohydrate recognition domains (CRDs), respectively, were selected. Gal3 contains one Trp, whereas Gal8 contains three, one at each binding site and a third one not involved in sugar binding; these were substituted by the corresponding F-Trp analogues. The presence of fluorine did not significantly modify the affinity for glycan binding, which was in slow exchange on the ^{19}F NMR chemical-shift timescale, even for weak ligands, and allowed binding events taking place at two different binding sites within the same lectin to be individualized.

Introduction

Molecular recognition of carbohydrates (saccharides, sugars, glycans) is at the heart of essential events of biological and biomedical interest.[1] Glycans act as contact points for cell–matrix and cell–cell interactions, strongly related to health and disease.[2] Most of these interactions involve lectins, which are protein receptors exquisitely specific for certain sugars.[3] Among lectins, galectins display diverse biological functions. It has been described that, as part of our innate immune system, they are involved in the regulation of immune activities, microbial recognition, and also display roles in development.[4–6] Although there are different types of galectins (prototype, chimera-type, and tandem repeat), they all bind β-galactose-containing glycans.[7] Moreover, they share a large structural homology along their so-called carbohydrate-recognition domains (CRDs).

For all galectins, glycan recognition involves the synergistic combination of intermolecular lectin-sugar hydrogen bonding (a key His moiety is always hydrogen bonded to the axial 4-OH hydroxy of the canonical β-Gal moiety), electrostatic interactions, and CH-π stacking interactions between the less polar face of the β-Gal pyranose ring (especially H4, H5, one H6 and H3 to a minor extent) and a highly conserved Trp residue, ubiquitous in all galectins (Figure 1).[4,8–10]

The key role of sugar-aromatic CH–π stacking interactions for stabilizing glycan/lectin complexes has been extensively demonstrated.[11,12] Theoretical and experimental studies have pointed out that the dispersive component of the CH–π interaction is the central one, with a minor contribution of the electrostatic factor. Interestingly, the predominant role of the dispersion implies that the sugar-aromatic interaction does not strictly depends on a very well-defined orientation between both partners and that these complexes display a highly dynamic nature.[11] Moreover, solvation–desolvation effects are

[a] Dr. M. G. Lete,[+] Dr. A. Franconetti,[+] S. Bertuzzi, S. Delgado, Dr. M. Azkargorta, Dr. F. Elortza, Dr. O. Millet, Dr. G. Jiménez-Osés, Dr. A. Arda, Prof. Dr. J. Jiménez-Barbero
CICbioGUNE, Basque Research & Technology Alliance (BRTA)
Bizkaia Technology Park, Building 800
48160 Derio, Bizkaia (Spain)
E-mail: aarda@cicbiogune.es
jjbarbero@cicbiogune.es

[b] Dr. G. Jiménez-Osés, Dr. A. Arda, Prof. Dr. J. Jiménez-Barbero
Ikerbasque, Basque Foundation for Science
Plaza Euskadi 5, 48009 Bilbao, Bizkaia (Spain)

[c] Prof. Dr. J. Jiménez-Barbero
Department of Organic Chemistry
II Faculty of Science and Technology
University of the Basque Country, EHU-UPV
48940 Leioa (Spain)

[d] Prof. Dr. J. Jiménez-Barbero
Centro de Investigación Biomédica
En Red de Enfermedades Respiratorias
Madrid (Spain)

[+] These authors contributed equally.

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202202208

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

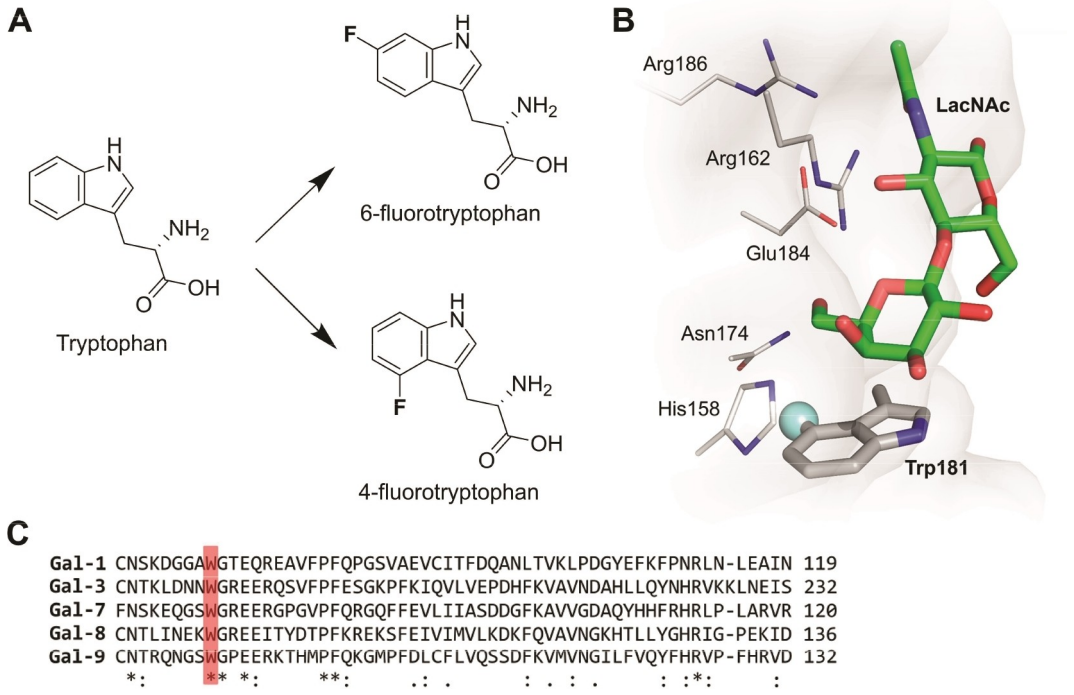


Figure 1. A) Tryptophan and fluorinated tryptophan analogue structures used for protein fluorination. B) Gal3 complexed with LacNAc (PDB ID: 1KJL) showing the fluorination of Trp181 at position 4; the residues of the binding site providing key interactions with the sugar are presented as sticks. C) Alignment of five human galectins' CRDs showing the conserved tryptophan residue in the binding site (in red). Uniprot codes used for the sequence alignment: galectin-1 (P09382), galectin-3 (P17931), galectin-7 (P47929), galectin-8 (O00214-1), and galectin-9 (O00182-1). Below the sequence "*" denotes conserved residues, ":" conservative substitutions, and "." semi-conservative substitutions.

of key importance, especially when charged species are involved or when highly polar or nonpolar groups are present.[11]

Different techniques, methods, and protocols are employed to monitor different aspects of carbohydrate-ectin interactions, from assessing their mere existence (yes/no) for a given system to their structural analyses at high resolution by employing NMR, X-ray crystallography or cryo-electron microscopy (cryoEM).[13]

Additionally, the thermodynamic and/or kinetic parameters of the recognition event may be deduced by using ITC, SPR, or BLI. Obviously, different protocols are more useful depending on the context: in vivo, in cell, in vitro, isolated entities, etc. In any case it is well recognized that molecular probes are required to obtain the experimental response and that, in principle, the smaller the chemical modification on any of the partners (either the glycan or the lectin), the closer the system under study will be to that taking place in the "natural" environment.

NMR has been extensively used to study protein-glycan interactions, especially given the intrinsic dynamic character of glycans and their interactions.[14] Both ligand-based and receptor-based NMR methods have provided tremendous insights into a large variety of systems.[15] Receptor-based NMR methods are grounded on the use of ^{15}N-labeled lectins and require the specific assignment of all the ^{1}H,^{15}N signals for all the amino acids that constitute the lectin, a process that is laborious and highly time consuming.

It has been recently reported that different chemical modifications at the Trp rings of different lectins (including galectin-1) do not substantially modify the binding affinity towards neutral sugars. In particular, ^{19}F substitution at different positions of the indole aromatic ring only affected the binding towards sulfated sugars,[16] due to changes in solvation/desolvation effects around the interacting partners. In contrast, the binding to the neutral sugars was basically unaffected. ^{19}F-protein labeling has been pursed in the past for diverse NMR applications,[17] including in-cell studies, ligand discovery or protein structural/dynamic studies, and different methods have been described for the incorporation of ^{19}F-containing amino acids into proteins.[18,19]

On this basis, we describe herein the NMR analysis of the interaction of different F-Trp-containing versions of two human galectins (Gal3 and Gal8), including the use of ^{19}F NMR spectroscopy to easily detect diverse features of the interaction of these lectins with a variety of ligands.

Results and Discussion

We have focused on two different galectins. First, the carbohydrate recognition domain of human galectin-3 (Gal3 CRD) was employed. Gal3 CRD is a monomer lectin that holds a unique

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tryptophan moiety (Trp181), specifically located at the canonical binding site and that is directly involved in the interaction with carbohydrates. As second target, we have employed human galectin-8 (Gal8), a heterodimer with two different lectin domains. Each domain displays a key tryptophan residue (Trp86 at the N-terminal domain and Trp249 at the C-terminal domain) that provides key CH-π stacking interactions with the β-galactose. Additionally, Gal8 has one further Trp moiety (Trp317) that is not involved in sugar binding.

For Gal3 CRD, two different tryptophan analogs, 4-F-Trp and 6-F-Trp (Figure 1), were used (see experimental section) to generate two fluorinated lectins (Gal3-4FW and Gal3-6FW). To facilitate the NMR analysis, the lectins were expressed and purified in their ^{15}N-labeled versions. Indeed, their corresponding ^{1}H,^{15}N HSQC spectra are shown in Figure S2 in the Supporting Information. The cross-peak pattern for each fluorinated galectin spectrum is fairly similar and points out that they display a similar folded structure. The average (^{1}H, ^{15}N) chemical shift perturbations (CSP) for both Gal3-4FW and Gal3-6FW with respect to the non-fluorinated (WT) Gal3 CRD are also gathered in Figure S2. Interestingly, the observed CSP for the Gal3-4FW where more notorious (several in the range of 0.1 ppm or larger) than those for the Gal3-6FW (always smaller than 0.06 ppm). This is not surprising since according to X-ray crystallographic structures the fluorine atom in Gal3-4FW is pointing towards several amino acid residues of the lectin (Figure 1), while that in the Gal3-6FW version is essentially solvent exposed (Figure S1).[20] In fact, the CSP involve residues 156–160 and 174–184 (especially for the 4-F-Trp analogue), located at strands S4 and S5/6, respectively (Figure S2).

The binding affinities for the canonical galectin ligand, LacNAc, were then estimated. By using standard ^{1}H,^{15}N HSQC titration experiments (see the Experimental Section for details and Figures S3, S6 and S7), the estimated dissociation constants (K_D) were 65 and 63 μM for the Gal3-4FW and Gal3-6FW variants, respectively, strikingly very close to that reported for Gal3 CRD (53 μM).[21] ITC measurements confirmed the similar affinities between the non-fluorinated Gal3 and the Gal3-6FW versions (Table S1, Figure S9) for LacNAc. Thus, the presence of a single fluorine atom in the key Trp ring does not significantly affect the binding affinity versus LacNAc.[20] Moreover, the observed trend of CSP for the ^{1}H,^{15}N crosspeaks of the two fluorinated variants in the presence of LacNAc were basically identical to those observed for the WT lectin (Figure S4).

Once the structure and the binding affinity of the two fluorinated-Trp lectins had been assessed, the possible added value of employing ^{19}F NMR spectroscopy was evaluated. Titration experiments using LacNAc as a ligand were again carried out, but in this case by monitoring the behavior of the ^{19}F NMR resonance signals of the fluorine atoms at the Trp rings of the Gal3-4FW and Gal3-6FW variants (Figure 2). Interestingly, for these binding affinities (see above), the free-to-bound equilibrium is in the slow exchange regime in the ^{19}F NMR chemical shift timescale. For instance, for the Gal3-6FW lectin, when 1 equivalent of LacNAc was added, a new ^{19}F NMR signal appeared at lower field ($\delta = -123.09$ ppm), while the intensity of the initial signal ($\delta = -124.40$ ppm), belonging to the free

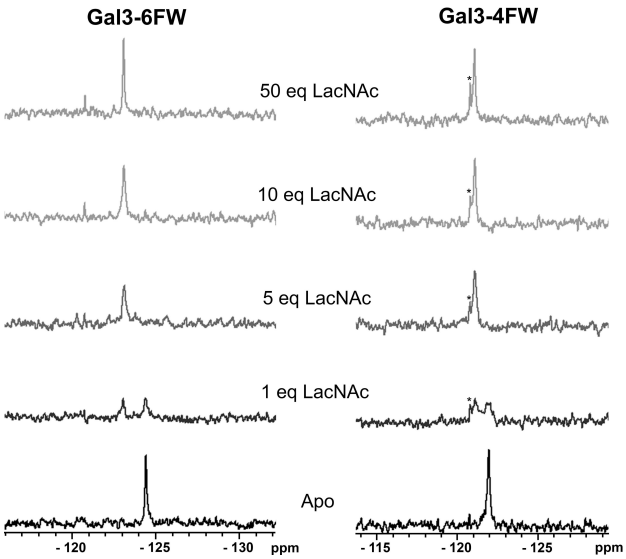


Figure 2. 1D ^{19}F NMR spectra recorded for the titration of fluorinated Gal3 variants with LacNAc. Bottom spectra correspond to the apo form of the proteins and the subsequent spectra show the addition of indicated LacNAc equivalents. Asterisks indicate an impurity, identified as the fluoride anion present in water.

lectin, clearly decreased (Figure 2). When 5 equivalents of LacNAc were added, the initial signal basically disappeared and only the signal at $\delta = -123.09$ ppm, corresponding to the bound lectin was observed. A similar behavior was observed for the Gal3-4FW analogue, for which the initial signal ($\delta = -121.94$ ppm) for the free lectin was gradually transformed into a new one at $\delta = -121.06$ ppm upon addition of LacNAc. Fittingly, the existence of slow exchange on the chemical shift timescale facilitates monitoring of binding events, while the presence of a single ^{19}F NMR signal further simplifies the analysis.

The Gal3-6FW variant was used to monitor the binding to two additional ligands: lactose and the histo-blood group B tetrasaccharide antigen type-II (B-II), which differ in their binding affinity for WT Gal3 by about 25-fold.[21] When lactose, the weaker binder (K_D 112 μM with WT Gal3),[21] was added to Gal3-6FW the initial peak ($\delta = -124.40$ ppm) shifted and broadened at the same time until saturation was reached, at high excess of the ligand, where the same intensity and shape as at the beginning were recovered at a different chemical shift ($\delta = -123.14$ ppm; Figure 3, left). For the stronger binder blood group antigen B-II (K_D 4 μM with Gal3 WT[21] and K_D 2 μM with Gal3-6FW as measured by ITC; Table S1 and Figure S9), a different situation was observed, with a slow exchange regime between the free and bound forms of Gal3-6FW, with two clear isolated and defined peaks at $\delta = -124.40$ and -122.69 ppm respectively at protein/ligand ratios below saturation (Figure 3, right).

The methodology was then extended to Gal8, which is a heterodimer with distinct glycan selectivity for the two domains. While both domains recognize LacNAc and Lac with relatively low affinity and with a slight preference for the

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

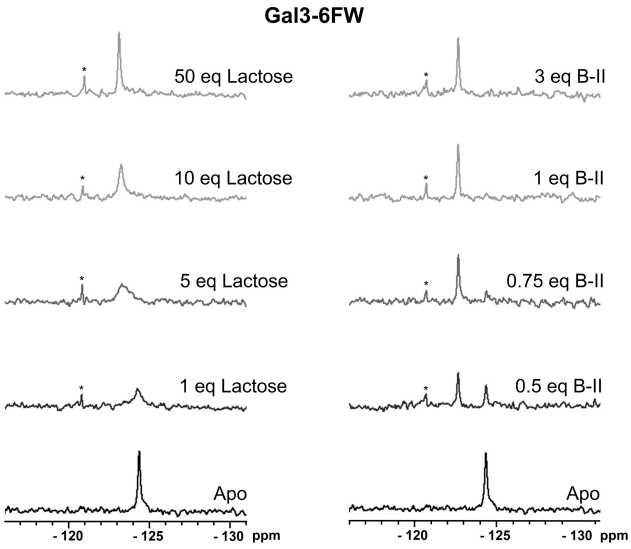


Figure 3. 1D ^{19}F NMR spectra for the titration of Gal3-6FW with lactose (left) and HBGA B type II (right). Bottom spectra correspond to the apo form of the protein and the subsequent spectra show the addition of ligand as indicated to the right. Asterisks indicate an impurity, identified as the fluoride anion present in water.

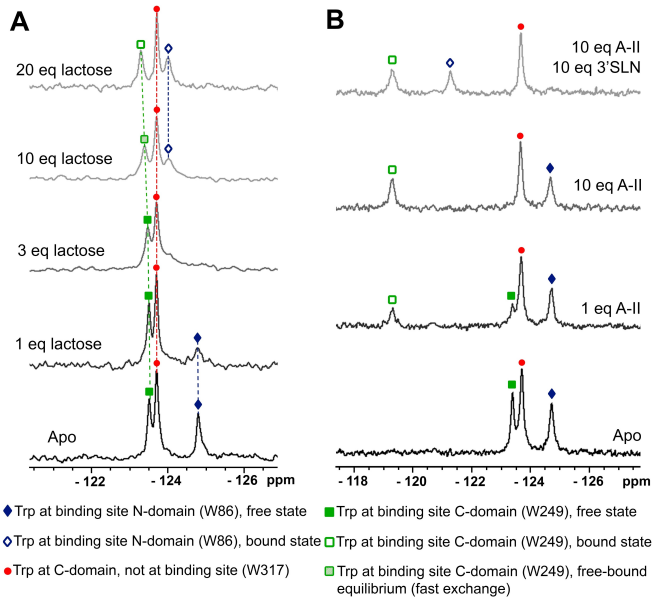


Figure 4. 1D ^{19}F NMR spectra recorded for the titration of Gal8-6FW with A) lactose and B) the specific ligands, Neu5Acα2-3Galβ1-3GlcNAc (3'SLN) and GalNAcα1-3(Fucα1–2)Galβ1-4GlcNAc (A-II), of the N- and C-terminal domains, respectively. In all cases, the NMR spectrum at the bottom corresponds to the apo form of Gal8-6FW. The molar equivalents of the added ligand are given on the right of the corresponding spectrum. Dashed lines at each ^{19}F signal follow chemical shift perturbations.

N terminus,[22,23] only the N-terminal domain binds sialylα2-3Gal-containing glycans (K_D ca. 4 μM for Neu5Acα2-3Galβ1-3GalNAc),[24] whereas the C terminus prefers the histo-blood group antigens (K_D ca. 12 μM for the A type II antigen: GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAc).[24] Moreover, it has been demonstrated that the two domains bind these sialyl- and fucose-containing oligosaccharides in an independent manner.[24] As mentioned above, Gal8 contains three Trp residues: one at each sugar binding site and a third one located at the C terminus of the C-domain, which is not involved in carbohydrate binding. In this case, only the 6-F-Trp analogue was employed to produced Gal8-6FW.

As for Gal3, ^{19}F incorporation produced very small perturbations in the ^{1}H,^{15}N HSQC spectrum of ^{15}N-Gal8-6FW compared to WT ^{15}N-Gal8 (Figure S5). As expected, the ^{19}F NMR spectrum of Gal8-6FW shows three different peaks (Figure 4, bottom spectra). The weak binder, lactose, was chosen as a first ligand for the titration experiments (Figure 4A). When 1 equiv. of lactose was employed, only the highest field peak ($\delta = -124.80$ ppm, blue filled-diamond) was affected. This peak broadened and completely disappeared in the presence of 10 equivalents of ligand, moment in which a new peak appeared at lower field ($\delta = -124.00$ ppm, blue empty-diamond). The peak at the lowest field ($\delta = -123.51$ ppm, green filled-square) was hardly affected at protein/ligand ratios below 1:3, but at higher molar ratios, it progressively shifted downfield to reach $\delta = -123.30$ ppm in the presence of 20 lactose equivalents (green empty-square). The third peak ($\delta = -123.70$ ppm, red circle) did not shift during the titration, indicating that it is not involved in the sugar binding event. As it is known that lactose binds slightly better to the N- than to the C-domain of Gal8,[22,23] it was anticipated that the ^{19}F-signal affected at lower protein/lactose ratios ($\delta = -124.80$ ppm, blue filled-diamond) belonged to the N-domain, while the peak at lower field ($\delta = -123.51$ ppm, green filled square), only affected at high protein/lactose ratios, corresponded to the C-domain. This fact is very remarkable. Lactose binds to both domains of Gal8 with very similar and weak affinities, which make that both events cannot be individualized and are detected as undistinguishable binding events by most standard binding techniques, including NMR. Here, due to the ^{19}F NMR properties,[25] both recognition processes can be clearly differentiated, thus allowing the identification of a slightly higher affinity binding event and a lower affinity one, despite their rather small differences affinities. Even more, while the ^{19}F-signal at the N-domain is in slow-intermediate exchange upon lactose binding, the ^{19}F signal at the C-domain is in a fast-exchange regime.

As described above, it has been demonstrated for Gal8[24] that the addition of the specific ligand for either the N- or C-domains, Neu5Acα2-3Galβ1- derivatives and fucosylated blood group antigens respectively, do not produce any effect on the other domain. On this basis, titration experiments with these ligands were carried out to confirm the assignment of each of the three ^{19}F signals to the corresponding Trp residue. The analysis of the ^{19}F NMR spectra recorded after systematic additions of the blood group antigen A type II (A-II) to the NMR tube containing Gal8-6FW showed (Figure 4B) specific chemical shift perturbations only for the lowest field peak ($\delta = -123.51$ ppm, green filled-square). Gal8 WT has been reported to bind this glycan with relatively high affinity (K_D 12 μM),[24] and indeed the binding event was slow in the ^{19}F NMR

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

chemical-shift timescale, with a new peak appearing at $\delta = -119.32$ ppm (green empty square). Again, ITC measurements corroborated that the affinity of Gal8 for this ligand was not altered by the presence of ^{19}F at the Trp residues (Table S1, Figure S9). After addition of 10 equivalents of the A type II antigen, the peak at $\delta = -123.51$ ppm disappeared. Therefore, we can safely conclude that this signal corresponds to the 6-FW residue at the binding site of the C-terminal domain (W249). Subsequently, Neu5Acα2–3Galβ1–4GlcNAc (3'SLN), which is only recognized by the N-domain of Gal8[22,23] was added to the sample. Upon addition of 10 equivalents of the glycan, only the highest field peak at $\delta = -124.80$ ppm (blue filled diamond) vanished, and a new peak appeared at $\delta = -121.29$ ppm (blue empty diamond). Therefore, this resonance signal corresponds to the 6-FW residue at the N terminus (W86). Again, no chemical shift perturbation or broadening was observed for the intermediate peak ($\delta = -122.6$ ppm, red circle), corroborating that it corresponds to W317, far from either binding site.

A schematic view of the possible complexes formed by the two ligands at the N- and C-domains is given in Figure 5. The Trp indole ring has been simply exchanged by the corresponding F-analogue in the deposited X-ray coordinates for both domains and the Neu5Acα2-3Galβ1-4GlcNAc and GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAc ligands have been docked at their specific sites.

Therefore, the data presented herein show that the substitution of Trp residues for F-Trp moieties in the binding site of two galectins does not substantially modify the binding affinity towards their ligands in a range of dissociation constants (low-mM to low-μM). Given the key role of pyranose-aromatic stacking interactions to stabilize sugar-lectin complexes,[4] this result seems particularly striking. Indeed, it has been reported that the presence of 4-F-Phe rings replacing Trp or Tyr rings at the binding site in hevein domains strongly decrease the binding energy and enthalpy versus GlcNAc-containing oligosaccharides with respect to that measured for the WT lectin,[26] as the fluorine atom deactivates the π cloud of the aromatic ring essential to provide the stabilizing interaction.[27] However, Tobola et al.[16] have recently reported that the affinity for neutral sugars is kept for fluorinated-Trp versions of Gal1, in line with our observations herein for Gal3 and Gal8. Those scientists demonstrated that the affinity is kept due to new modulated interactions and solvation patterns in the fluorinated analogues.[16] Therefore, quantum mechanics calculations were carried out to explain these experimental observations. Quantum mechanical calculations on a cluster model of the LacNAc:Gal3 CRD complex (see Computational Details) proved that binding affinity is not altered (ca. ± 0.1 kcal mol^{-1}) by the presence of a single fluorine atom in Trp181. Indeed, no major decreases of the electron densities at the center of the benzene rings of the 4F/6F-fluorinated indol moieties were deduced from the calculations (Figure 6A). The calculated geometries of the laterally stacked Gal pyranose fragment were virtually identical, with equal or slightly longer CH–π distances in the presence of 4F/6F-fluorinated Trp181 (Figure 6B and C).

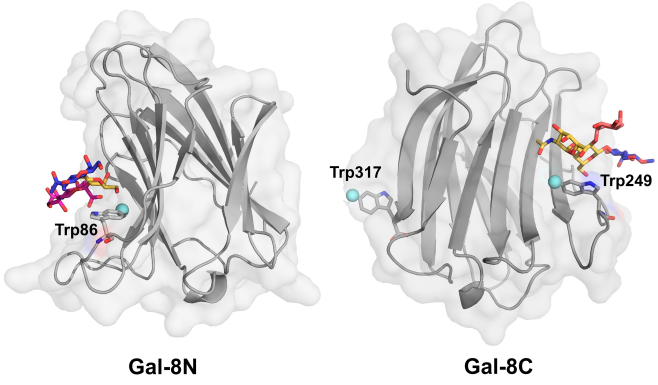
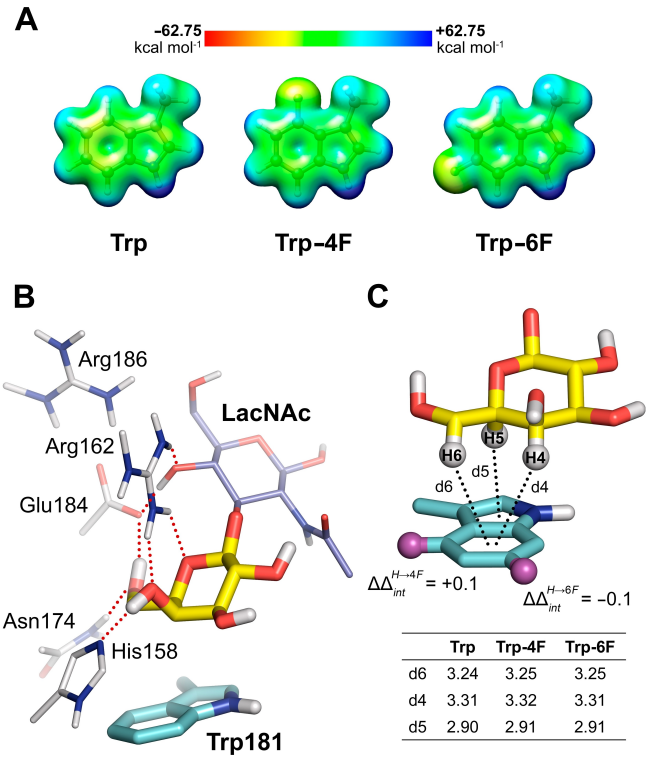


Figure 5. 3D models for the complexes between the N- and C-domains of Gal8 with 6F-Trp, with 3'SLN and blood group A-II tetrasaccharide, respectively, based on available X-ray crystallographic structures (PDB IDs: 3AP4, 3VKL).



	Trp	Trp-4F	Trp-6F
d6	3.24	3.25	3.25
d4	3.31	3.32	3.31
d5	2.90	2.91	2.91

Figure 6. A) Electrostatic potential surfaces calculated with PCM(H2O)/ωB97X-D/6-311++G(2d,p)//ωB97X-D/6-31G(d,p) of the indole rings of non-fluorinated and 4F/6F-fluorinated Trp. There is minimal decrease in the electron density at the center of the benzene rings of the fluorinated indoles. B) Model of LacNAc bound to Gal3 CRD computed from PDB structure 4XBN; hydrogen bonds are shown as red dashed lines. C) Detail of the stacked binding mode between Trp181 and Gal; CH/π interactions are shown as black dashed lines. The Trp181 positions for which H has been substituted by F are represented by purple spheres. The relative interaction energies between the Gal3 CRD binding site and LacNAc upon Trp181 fluorination ($\Delta\Delta_{int}^{H\rightarrow 4F/6F}$), calculated with PCM(H2O)/ωB97X-D/6-311++G-(2d,p)//ωB97X-D/6-31G(d,p), are given in kcal mol^{-1}; positive and negative values indicate a decrease or increase in binding affinity, respectively. Distances in the table are given in ångströms.

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conclusions

The use of ^{19}F NMR for monitoring glycan recognition in F-Trp-containing lectins has been shown to be very efficient. Obviously, the corresponding Trp residue needs to be involved in sugar binding, as described above, and they should have been previously identified. As a strong point, very few NMR signals appear in the ^{19}F NMR spectrum, thus facilitating the analysis. For relatively high affinity events (in the low-μM range), the large ^{19}F chemical shift dispersion allows entry to the slow-exchange chemical shift timescale, which is also extremely useful for discriminating and assessing binding. According to the measured binding affinities, the chemical perturbation introduced by the fluorine atom in the Trp ring is minimal. Obviously, it is higher than the introduction of an isotope (i.e., ^{13}C instead of ^{12}C,[28,29]), but it still produces a receptor whose chemical nature remains similar to that of the natural lectin.

The protocol has been shown to be very useful for distinguishing between independent recognition events that take place at two different binding sites within the same lectin. Indeed, there are very few (bio)chemical methods that are able to detect binding at very high resolution with relatively small manipulation and by using fairly simple experimental protocols (^{19}F NMR spectroscopy). Thus, as also pointed out by others in a different context,[16,30] the use of F-Trp moieties could be a useful strategy to monitor the selective or competitive binding of different glycans by lectins, provided that those Trp residues are involved in sugar binding.

Experimental Section

Materials: *N*-(Phosphonomethyl)glycine (glyphosate), 4-Fluoro-DL-tryptophan were purchase from Sigma-Aldrich and 6-fluoro-DL-tryptophan from Acros Organics. Oligosaccharides were from Elycityl (references: GLY008 (LacNAc), GLY038-2 (blood group antigen B type II), and GLY035-2 (blood group antigen A type II)).

Protein expression and purification: The expression of fluorinated galectins was performed following described protocols, using *Escherichia coli* BL21(DE3) strain.[31,32] The incorporation of fluorinated tryptophan was achieved by the addition of glyphosate and exogenous addition of aromatic amino acids. Briefly, *E. coli* BL21 were transformed, and a 5 ml saturated culture was added in M9 medium and grown overnight. Next, a culture was started in M9 media at an OD_{600} of 0.1 and grown up to OD_{600} 0.4. At that point, glyphosate (*N*-(phosphonomethyl)glycine) (1 g L^{-1}) that was first dissolved in 1 M sodium hydroxide,[33] phenylalanine (60 mg L^{-1}) and tyrosine (60 mg L^{-1}) were added to the culture. After 1 h the fluorinated tryptophan analog (120 mg L^{-1}) was added, and protein production was induced with 1 mM isopropyl β-D-1-thio-galactopyranoside (IPTG) and the culture continued to grow for 3 h at 37 °C. For ^{15}N-labeled protein ^{15}N-NH_4Cl (1 g) was added as the nitrogen source. All proteins were purified with the same protocol. The cultures were centrifuged and proteins were purified by lactose affinity (α-lactose-agarose resin (Sigma-Aldrich)) as described elsewhere.[21,24,34] Briefly, the cell pellet was resuspended in lysis buffer (22 mM Tris·HCl, 5 mM EDTA, 1 mM PMSF, and 1 mM DTT at pH 7.5). Later, lysis by sonication on ice with 60% amplitude, 12 repetitions of 20 s, and 59 s of an interval between each pulse. In order to clarify the sample ultracentrifugation at 35000 rpm 141316 g for 1 h at 4 °C was performed. Lactose-agarose resin was loaded with the soluble fraction pre equilibrated in buffer (50 mM Tris pH 7.2, 150 mM NaCl). Protein elution was achieved by addition of 150 mM lactose containing PBS buffer. Protein purity was checked by 4–12% SDS-PAGE. The incorporation of fluorine at the desired Trp residue was confirmed by LC–MS. Fluorine incorporation levels by following this protocol have been reported to occur in more than 95%.[29] We have performed a rough estimation ourselves by analyzing the ^{1}H,^{15}N HSQC spectra of ^{15}N-Gal3-4FW and ^{15}N-Gal3-6FW (see the Supporting Information) that fully agrees with this estimation. Galectins were thoroughly dialyzed against PBS, pH 7.4 until no lactose was present, before use.

NMR Experiments. All spectra were performed at 298 K on a Bruker Avance 2 600 MHz spectrometer. The ^{19}F NMR spectra of proteins were recorded using a selective SEF probe with ^{1}H decoupling. Acquisition parameters varied depending on the protein concentration (usually between 60 and 90 μM). For a 60 μM of fluorinated protein, for instance ^{19}F NMR spectra were recorded with SW = 60 Hz, ns = 80, TD = 8 K and processed with exponential apodization (lb between 15 and 30 depending on the desired result). Spectra were referenced to internal TFA (10 μM, δ $^{19}F = -76.5$ ppm). ^{1}H,^{15}N HSQC experiments were recorded using a standard triple-channel probe. All the protein samples (500 μL total in 5 mm standard NMR tubes) were prepared at a concentration between 100 and 40 μM. The buffer used was 90% phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) and 10% deuterated water (D_2O). Ligands were titrated to the protein sample, and a ^{1}H,^{15}N HSQC experiment or a ^{1}H-decoupled ^{19}F spectra were recorded at each point. Chemical shift perturbation (CSP) from the HSQC was analyzed using CcpNmr Analysis 2.4.2 software.[35] The CSP were obtained by applying the formula $\Delta\delta$ (ppm) $= [(\Delta\delta H^2 + (0.14 \cdot \Delta\delta N)^2)/2]$.[36] The obtained data were plotted for each NH residue of the protein's backbone. The K_D values were obtained by fitting the CSP at different ligand/protein ratios to the corresponding equation[35,36] as implemented in the corresponding module in CcpNmr Analysis 2.4.2 program.

Quantum mechanical calculations: An abbreviated model of LacNAc bound to Gal3 CRD binding site was constructed from crystallographic structure 4XBN; only the co-crystallized ligand (LacNAc) and the protein side chains directly interacting with it were kept in the model. Restrained geometry optimizations were carried out with Gaussian 16[37] using the ωB97X–D hybrid functional[38] and 6–31 + G(d,p) basis set with ultrafine integration grids. The whole ligand and all added hydrogens were allowed to move while the protein heavy atoms were fixed. Bulk solvent effects in water were considered implicitly through the IEF-PCM polarizable continuum model.[39] Electronic energies (ΔE) were used for the discussion on the relative stabilities of the calculated structures.

Acknowledgements

This research was funded by the European Research Council (ERC-2017-AdG, project number 788143-RECGLYCA NMR to J.J.-B.), Agencia Estatal de Investigación (Spain), though projects RTI2018-094751-B-C21 to J.J.-B. and A.A. and PID2021-125946OB-I00 to G.J.-O., and CIBER -Consorcio Centro de Investigación Biomédica en Red- (CIBERES), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación. A.F. thanks MICINN/AEI from Spain for a "Juan de la Cierva" contract (IJC2019-042061-I). We also acknowledge the NMR resources and the technical support provided by the Euskadi NMR lab

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(LRE) of the ICTS "Red de Laboratorios de RMN de biomoléculas (R-LRB)" of Spain.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: engineered galectins · fluorotryptophan · glycans · molecular recognition · ^{19}F NMR spectroscopy

[1] H. J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís, *Trends Biochem. Sci.* **2011**, *36*, 298–313.
[2] A. Varki, *Glycobiology* **2017**, *27*, 3–49.
[3] I. J. Goldstein, C. E. Hayes, *Adv. Carbohydr. Chem. Biochem.* **1978**, *35*, 127–340.
[4] R. D. Cummings, F.-T. Liu, G. A. Rabinovich, S. R. Stowell, G. R. Vasta in *Essentials of Glycobiology*, 4th ed. (Eds: A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, P. H. Seeberger), Cold Spring Harbor Laboratory Press, New York, **2022**, Chapter 36.
[5] D. Laaf, P. Bojarová, L. Elling, V. Křen, *Trends Biotechnol.* **2019**, *37*, 402–415.
[6] L. Johannes, R. Jacob, H. Leffler, *J. Cell Sci.* **2018**, *131*.
[7] G. A. Rabinovich, L. G. Baum, N. Tinari, R. Paganelli, C. Natoli, F. T. Liu, S. Iacobelli, *Trends Immunol.* **2002**, *23*, 313–320.
[8] S. Di Lella, V. Sundblad, J. P. Cerliani, C. M. Guardia, D. A. Estrin, G. R. Vasta, G. A. Rabinovich, *Biochemistry* **2011**, *50*, 7842–7857.
[9] G. R. Vasta, *Adv. Exp. Med. Biol.* **2020**, *1204*, 169–196.
[10] S. Bertuzzi, J. I. Quintana, A. Ardá, A. Gimeno, J. Jiménez-Barbero, *Front. Chem.* **2020**, *8*, 593.
[11] J. L. Asensio, A. Ardá, F. J, Cañada, J. Jiménez-Barbero, *Acc. Chem. Res.* **2013**, *46*, 946–954.
[12] V. Spiwok, *Molecules* **2017**, *22*(7), 1038.
[13] A. Gimeno, P. Valverde, A. Ardá, J. Jiménez-Barbero, *J. Curr. Opin. Struct. Biol.* **2020**, *62*, 22–30.
[14] A. Ardá, J. Jiménez-Barbero, *Chem. Commun.* **2018**, *54*, 4761–4769.
[15] P. Valverde, J. I, Quintana, J.I Santos, A. Ardá, J. Jiménez-Barbero, *ACS Omega* **2019**, *4*, 13618–13630.
[16] F. Tobola, M. Lepšík, S. R. Zia, H. Leffler, U. J. Nilsson, O. Blixt, A. Imberty, B. Wiltschi, *ChemBioChem* **2022**, *23*, e202100593.
[17] C. R. Buchholz, W. C. K. Pomerantz, *RSC Chem. Biol.* **2021**, *2*, 1312.
[18] Y. Ye, Q. Wu, W. Zheng, B. Jiang, G. J, Pielak, M. Liu, C. Li, *Anal. Bioanal. Chem.* **2018**, *410*, 869–874.
[19] C. Li, G. F. Wang, Y. Wang, R. Creager-Allen, E. A. Lutz, H. Scronce, K. M. Slade, R. A. S. Ruf, R. A. Mehl, G. J. Pielak, *J. Am. Chem. Soc.* **2010**, *132*, 321.
[20] P. Sörme, P. Arnoux, B. Kahl-Knutsson, H. Leffler, J. M. Rini, U. J. Nilsson, *J. Am. Chem. Soc.* **2005**, *127*, 1737–1743.
[21] A. Gimeno, S. Delgado, P. Valverde, S. Bertuzzi, M. A. Berbís, J. Echavarren, A. Lacetera, S. Martín-Santamaría, A. Surolia, F. J. Cañada, *Angew. Chem. Int. Ed.* **2019**, *58*, 7268–7272; *Angew. Chem.* **2019**, *131*, 7346–7350.
[22] S. Carlsson, C. T. Öberg, M. C. Carlsson, A. Sundin, U. J. Nilsson, D. Smith, R. D. Cummings, J. Almkvist, A. Karlsson, H. Leffler, *Glycobiology* **2007**, *17*, 663–676.
[23] J. Hirabayashi, T. Hashidate, Y. Arata, N. Nishi, T. Nakamura, M. Hirashima, T. Urashima, T. Oka, M. Futai, W. E. G. Muller, F. Yagi, K. Kasai, *Biochim. Biophys. Acta Gen. Subj.* **2002**, *1572*, 232–254.
[24] M. Gómez-Redondo, S. Delgado, R. Núñez-Franco, G. Jiménez-Osés, A. Ardá, J. Jiménez-Barbero, A. Gimeno, *RSC Chem. Biol.* **2021**, *2*, 932–941.
[25] "^{19}F in Biomolecular NMR", A. M. Gronenborn, *Structure* **2022**, *30*, 6–14.
[26] J. Jiménez-Barbero, F. J. Javier 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.
[27] 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, *Chemistry* **2005**, *11*, 7060–7074.
[28] L. Unione, M. J. Moure, M. P. Lenza, I. Oyenarte, J. Ereño-Orbea, A. Ardá, J. Jiménez-Barbero, *Angew. Chem. Int. Ed.* **2022**, *61*, e202201432.
[29] M. J. Moure, A. Gimeno, S. Delgado, T. Diercks, G. J. Boons, J. Jiménez-Barbero, A. Ardá, *Angew. Chem. Int. Ed.* **2021**, *60*, 18777–18782.
[30] F. Tobola, M. Lelimousin, A. Varrot, E. Gillon, B. Darnhofer, O. Blixt, R. Birner-Gruenberger, A. Imberty, B. Wiltschi, *ACS Chem. Biol.* **2018**, *13*, 2211–2219.
[31] P. B. Crowley, C. Kyne, W. B. Monteith, *Chem. Commun.* **2012**, *48*, 10681–10683.
[32] "Development and Application of ^{19}F NMR of Proteins", J. Kitevski-LeBlanc, *Doctoral Thesis*, University of Toronto (Canada), **2010**.
[33] H. W. Kim, J. A. Perez, S. J. Ferguson, I. D. Campbell, *FEBS Lett.* **1990**, *272*, 34–36.
[34] S. Bertuzzi, A. Gimeno, R. Núñez-Franco, G. Bernardo-Seisdedos, S. Delgado, G. Jiménez-Osés, O. Millet, J. Jiménez-Barbero, A. Ardá, *Chem. Eur. J.* **2020**, *26*, 15643–15653.
[35] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, *59*, 687–696.
[36] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.
[37] *Gaussian 16, Revision C.01*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, J. E. Jr.; Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.
[38] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
[39] G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110.

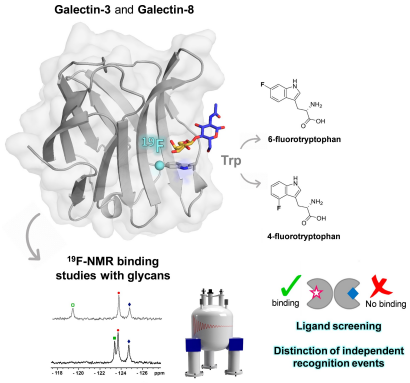
Manuscript received: July 14, 2022
Accepted manuscript online: November 7, 2022
Version of record online: ■■, ■■■

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RESEARCH ARTICLE

The incorporation of single-^{19}F-tryptophan residues into two glycan binding proteins, Galectin-3 and Galectin-8, was exploited to monitor their binding to carbohydrate ligands by using NMR spectroscopy. Although the affinity of the molecular recognition events was not substantially altered, ^{19}F NMR spectroscopy provided interesting added values with respect to other NMR binding strategies.



Dr. M. G. Lete, Dr. A. Franconetti, S. Bertuzzi, S. Delgado, Dr. M. Azkargorta, Dr. F. Elortza, Dr. O. Millet, Dr. G. Jiménez-Osés, Dr. A. Arda, Prof. Dr. J. Jiménez-Barbero**

1 – 8

NMR Investigation of Protein-Carbohydrate Interactions: The Recognition of Glycans by Galectins Engineered with Fluorotryptophan Residues

15213765, 0, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/chem.202202208 by Universidad de la Rioja, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License