Breaking the limits in glycan recognition by NMR

Jesus Jimenez-Barbero
CIC bioGUNE

Glycans are everywhere. Every cell, from the simplest bacteria and viruses to those of complex organisms such as human beings, is covered by a layer of glycans.They influence a myriad of biological events of biomedical importance, including the modulation of adaptive and innate immune responses. Therefore, these molecules hold tremendous potential in the quest for new therapeutics.

Adobe Stock © huenstructurebio.com

RECGLYCANMR uses nuclear magnetic resonance (NMR) for studying, at atomic resolution, molecular recognition processes in which glycans are involved. Till now, sugar recognition NMR studies have been exclusively limited to in vitro protocols. However, the exquisite glycan selectivity of lectins and modulation of promiscuity of glycans is likely strongly influenced by the environment, and the cell is a very special compartment. RECGLYCANMR also studies these interactions in the biological context: at the cell surface and within it, a crowded ambient with many components and where microviscosity is doubled with respect to water. Thus, RECGLYCANMR represents a multidisciplinary approach that combines state-of-the-art chemistry and chemical biology methods, in-cell structural chemistry, and biophysics protocols under crowding conditions to ask and answer pivotal questions related to sugar molecular recognition in Nature and their implications for finding solutions for numerous diseases.

Advances in synthetic chemistry, including chemo-enzymatic methods with glycosyltransferases and solid-phase oligosaccharide synthesis, allow us to access pure glycan structures. Moreover, the advances in eukaryotic expression systems also enable access to well-defined protein structures with isotope labelling for NMR studies. Therefore, the concomitant employment of methods that permit us to detect binding with atomic resolution (as NMR), to fully characterise the dynamic features of the players (NMR and molecular dynamics) and to deduce the energy of the interaction (ITC, SPR) open new horizons for glycan recognition.

A novel method has been presented to the scientific community to unambiguously assess the specific binding epitope, towards galectins, of glycans displaying repeating disaccharide entities, such as polyLacNAc. The combination of chemical synthesis, using a chemo-enzymatic approach, with NMR methods (STD-HSQC) allowed the description of the specific contribution of each LacNAc moiety, within a hexasaccharide containing three identical LacNAc disaccharides, to its interaction with human galectin-1, -3, -7, -8, and 9. The concept is based on the selective incorporation of ^{13}C-labelled atoms only at designed Gal fragments in the disaccharide (Moure *et al*., 2021). No additional chemical tags that may modify the natural hexasaccharide were employed. Just the existence of the ^{13}C-labelled galactose cracks the degeneracy of the chemical shifts of the NMR signals that take place in the regular ^{12}C-containing molecule and permit the identification of the specific interacting Gal moiety (Figure 1). This breakthrough opens new avenues for investigating biomedically relevant glycan-galectin interactions at high resolution.

A second breakthrough partially based on this concept has unambiguously demonstrated that the spike glycoprotein of SARS CoV2, initially responsible for the viral infection, directly binds external sialic acids (Unione *et al*., 2022). The challenge here was to distinguish the exogenous sialyl-containing glycans from those located at viral glycoprotein. Therefore, two different ^{13}C-labelled sialyl glycans were chemically synthesised, and NMR spectroscopy (STD-HSQC) was employed to experimentally assess their interactions with the fully glycosylated SARS CoV2 spike glycoprotein. Additional NMR experiments carried out with isolated regions of the spike (RBD, NTD) and others performed with blocking antibodies concluded that the spike efficiently binds external sialic acids through its N-terminal domain. These investigations may open new opportunities to develop glycan-based inhibitors to interfere with the infection and thus combat COVID-19.

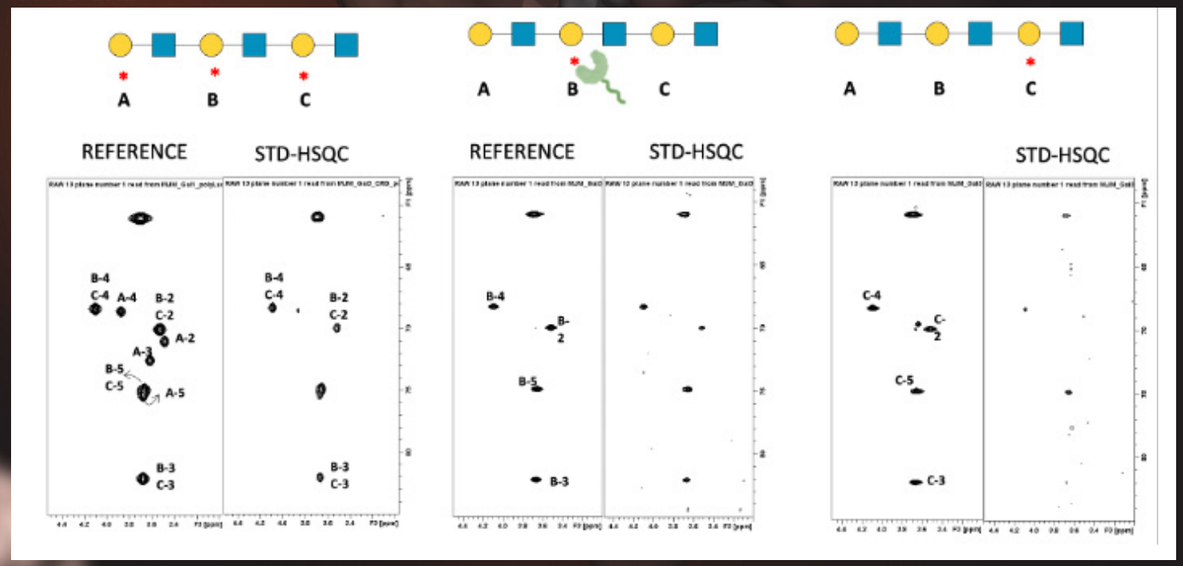

Figure 1: The specific interaction of the three distinct Gal moieties within a hexasaccharide with 3 LacNAc repeating units with the carbohydrate recognition domain unambiguously detected using three specifically 13C-labelled molecules at the Gal positions. NMR allow assessing that the best interaction resides at the intermediate moiety.

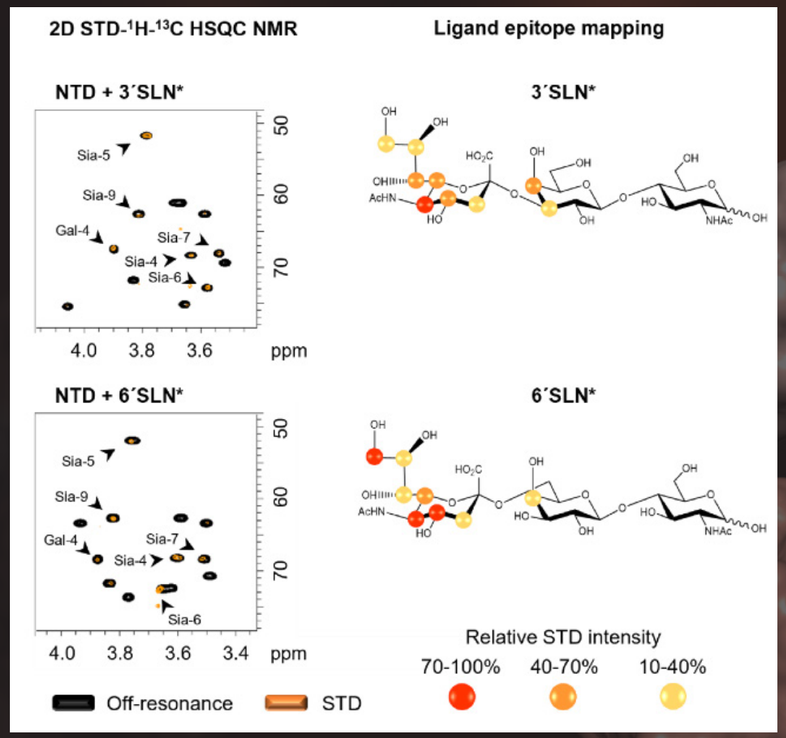

Figure 2: 2D STD-¹H-¹³C HSQC NMR experiments demonstrate the direct interaction of the N-terminal domain of the spike glycoprotein of SARS-CoV-2 with two different sialyl containing trisaccharides with ¹³C-labels.

Understanding the role of presentation and multivalency on glycan recognition in Nature is one of the key issues within RECGLYCANMR. Indeed, the presentation of the glycan molecules on the biological membranes and the intrinsic structural and chemical characteristics of these actors probably display a major impact on the recognition events. In this context, we have employed liposomes to mimic the natural membrane (Lete *et al*., 2022). The liposomes present diverse glycosphingolipids (GSLs), exposing diverse glycans whose interaction with two different human galectins, galectin-3 and -8, has been scrutinised by NMR. Strikingly, the two galectins show different tendencies in their recognition preferences towards GM1, GM3, and their non-sialylated analogues, which depend on the presence or not of the bilayer. This fact highlights the key relevance of glycan presentation to be efficiently recognised by protein receptors and thus provide the concomitant biological response.

Regarding the role of multivalency, the interaction of galectins with glycans presented in a multivalent manner in N-(2-hydroxypropyl) methacrylamide (HPMA) copolymers have been evaluated using a combination of NMR, cryo-electron-microscopy and dynamic light scattering (DLS) methods, in collaboration with the group of P. Borajova at Prague. Interestingly, it was demonstrated that, for galectin monomers, such as galectin-3, the interaction systematically increased with the number of lactose units present in the polymer scaffold, probably due to statistical effects. In this case, strong multivalent features were not observed. However, for galectin homodimers, as human galectin-1, large cross-linked supramolecules were generated in the presence of the polymers, especially for a particular decoration of the polymer, again evidencing the key role of proper presentation of the glycan to provide efficient binding (Bertuzzi *et al*., 2021).

One of the most relevant families of glycan receptors in Nature is galectins. These proteins display a myriad of biological functions related to health and disease issues. In RECGLYCANMR, we are focused on exploring the interaction of galectins with different glycans, especially the histo-blood group antigens, since it has been shown that human galectin-4 can bind to and destroy some bacteria (*Escherichia coli*) that express the blood group B antigen. We have carried out a systematic multidisciplinary study (Quintana *et al*., 2021) combining NMR with isothermal titration calorimetry (ITC). These experimental data have been assisted by computational methods to disentangle the fine structural, conformational and dynamic features of the interaction of the N-terminal domain of Gal-4 with the histo-blood group antigens. The B-antigens were demonstrated to be better binders for

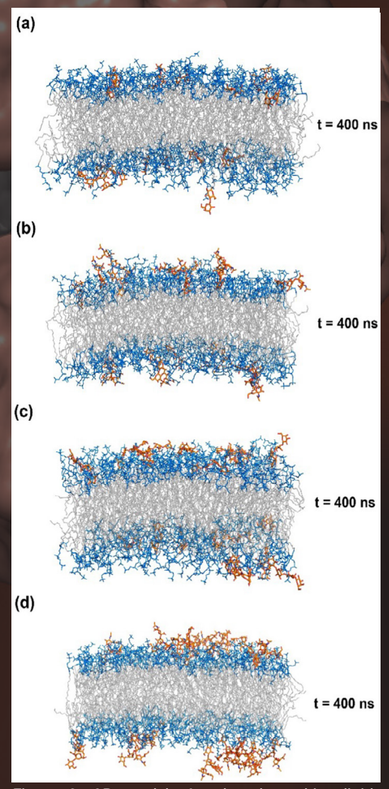

Figure 3: 3D models for the glycosphingolipids embedded into liposomes, showing the presentation of the glycan moieties ready to interact with their biological receptors.

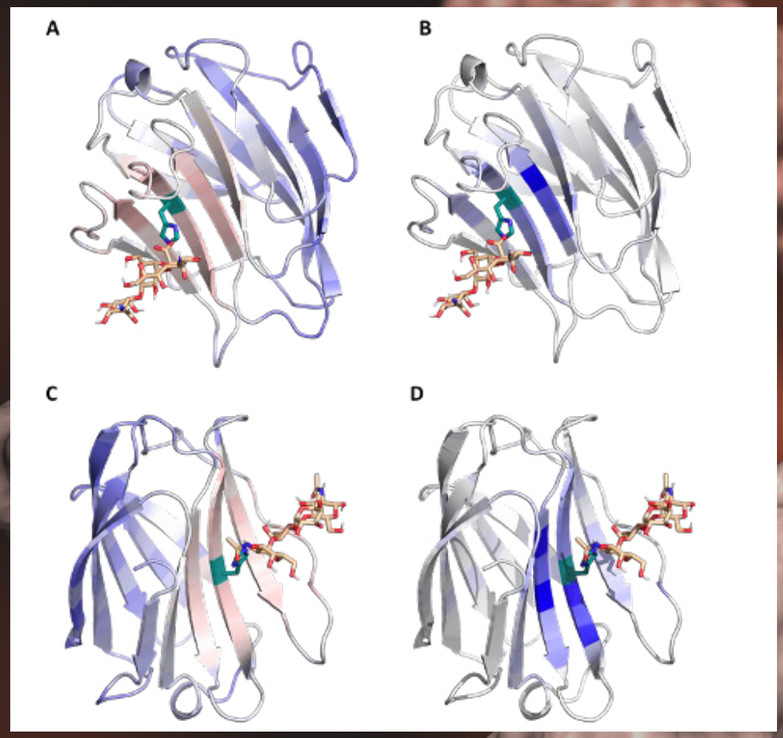

Figure 4: Representation of the interactions of hGal-8 N-terminal domain (A and B) and hGal-8 C-terminal domain (C and D) with their key glycan ligands.

this galectin. Moreover, we have also studied the recognition of these and other glycans by other human galectins. In this case, the full protein and its C-terminal and N-terminal domains (CRD) were analysed using a similar methodology to that described previously. Strikingly, both separate domains, the C-term and the N-term, show a drastic opposing specificity to recognise fucose or sialic-acid-containing ligands, respectively (Gomez-Redondo *et al*., 2021). Therefore, the two domains can target diverse partners, thus building up supramolecular entities and generating cluster effects to provide the biological signal.

References

Bertuzzi, S., Gimeno, A., Martínez-Castillo, A., Lete, M.G., Delgado, S., Airoldi, C., Rodrigues Tavares, M., Bláhová, M., Chytil, P., Křen, V., Abrescia, N.G.A., Ardá, A., Bojarová, P. and Jiménez-Barbero, J. (2021) 'Cross-Linking Effects Dictate the Preference of Galectins to Bind LacNAc-Decorated HPMA Copolymers', *International Journal of Molecular Sciences*, 22(11), 6000. doi: 10.3390/ijms22116000.

Gómez-Redondo, M., Delgado, S., Núñez-Franco, R., Jiménez-Osés, G., Ardá, A., Jiménez-Barbero, J. and Gimeno A. (2021) 'The two domains of human galectin-8 bind sialyl- and fucose-containing oligosaccharides in an independent manner. A 3D view by using NMR.', *RSC Chemical Biology*, 2(3), pp. 932–941. doi: 10.1039/d1cb00051a.

Lete, M.G., Franconetti, A., Delgado, S., Jiménez-Barbero, J. and Ardá, A. (2022) 'Oligosaccharide Presentation Modulates the Molecular Recognition of Glycolipids by Galectins on Membrane Surfaces', *Pharmaceuticals*, 15(2), p. 145. doi: 10.3390/ph15020145.

Moure, M.J., Gimeno, A., Delgado, S., Diercks, T., Boons, G.J., Jiménez-Barbero, J. and Ardá, A. (2021) 'Selective 13C-Labels on Repeating Glycan Oligomers to Reveal Protein Binding Epitopes through NMR: Polylactosamine Binding to Galectins', *Angewandte Chemie International Edition*, 60(34), pp. 18777–18782. doi: 10.1002/anie.202106056.

Quintana, J.I., Delgado, S., Núñez-Franco, R., Cañada, F.J., Jiménez-Osés, G., Jiménez-Barbero, J. and Ardá, A. (2021) 'Galectin-4 N-Terminal Domain: Binding Preferences Toward A and B Antigens With Different Peripheral Core Presentations', *Frontiers in Chemistry*, 9, 664097. doi: 10.3389/fchem.2021.664097.

Unione, L., Moure, M.J., Lenza, M.P., Oyenarte, I., Ereño-Orbea, J., Ardá, A. and Jiménez-Barbero, J. (2022) 'The SARS-CoV-2 Spike Glycoprotein Directly Binds Exogeneous Sialic Acids: A NMR View', *Angewandte Chemie International Edition*, 61. doi: 10.1002/anie.202201432.

PROJECT NAME
RECGLYCANMR

PROJECT SUMMARY
Breaking the limits in glycan recognition by NMR (RECGLYCANMR) is a multidisciplinary project that combines state-of-the-art chemistry and chemical biology methods, including in vitro and in-cell NMR, and biophysics protocols under crowding conditions to ask and answer pivotal questions related to sugar molecular recognition in Nature that are intimately related to infection and inflammation diseases.

PROJECT PARTNERS
RECGLYCANMR is based at the Center for Cooperative Research in Biosciences (CIC bioGUNE, Bilbao), taking advantage of the state-of-the-art NMR facility and technology platforms (crystallography, electron microscopy, biophysics, computing cluster, oligosaccharide solid-phase synthesis) at this multidisciplinary centre.

PROJECT LEAD PROFILE
Jesús Jiménez-Barbero has been Ikerbasque Professor and Scientific Director of CIC bioGUNE since 2014. He received his PhD in Madrid (1987). Earlier in his career, he worked at Grenoble, Zurich, Mill Hill and Pittsburgh. His investigations on glycan-receptor interactions have been disseminated through more than 600 scientific publications and more than 250 invited and plenary lectures.

PROJECT CONTACTS
Bizkaia Technology Park, Building 801A;
48160 Derio, Spain
+34944061300
www.cicbiogune.es
@cicbiogune
@GlycoNmr_Bilbao

Jesus Jimenez-Barbero
jjbarbero@cicbiogune.es

Itziar Gil de la Pisa
ipisa@cicbiogune.es

FUNDING
This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 788143.