Full Paper
doi.org/10.1002/chem.202003212

Chemistry Europe
European Chemical Societies Publishing

NMR Spectroscopy

Unravelling the Time Scale of Conformational Plasticity and Allostery in Glycan Recognition by Human Galectin-1

Sara Bertuzzi,[a] Ana Gimeno,[a] Reyes Núñez-Franco,[a] Ganeko Bernardo-Seisdedos,[a] Sandra Delgado,[a] Gonzalo Jiménez-Osés,[a] Oscar Millet,[a] Jesús Jiménez-Barbero,*[a, b, c] and Ana Ardá*[a]

Dedicated to the memory of Doctor Kilian Muñiz

Abstract: The interaction of human galectin-1 with a variety of oligosaccharides, from di-(*N*-acetyllactosamine) to tetrasaccharides (blood B type-II antigen) has been scrutinized by using a combined approach of different NMR experiments, molecular dynamics (MD) simulations, and isothermal titration calorimetry. Ligand- and receptor-based NMR experiments assisted by computational methods allowed proposing three-dimensional structures for the different complexes, which explained the lack of enthalpy gain when increasing the chemical complexity of the glycan. Interestingly, and independently of the glycan ligand, the entropy term does not oppose the binding event, a rather unusual feature for protein-sugar interactions. CLEANEX-PM and relaxation dispersion experiments revealed that sugar binding affected residues far from the binding site and described significant changes in the dynamics of the protein. In particular, motions in the microsecond-millisecond timescale in residues at the protein dimer interface were identified in the presence of high affinity ligands. The dynamic process was further explored by extensive MD simulations, which provided additional support for the existence of allostery in glycan recognition by human galectin-1.

Introduction

Human galectins are β-galactoside (βGal) binding lectins that participate in the regulation of an extraordinary variety of biological phenomena most of them related, but not only, to immunity.[1] At the same time, their connection with several diseases, such as cancer[2] or diabetes has been established, increasing the interests in exploiting them in different therapeutic strategies, as well in the development of disease biomarkers. These carbohydrate binding lectins are broadly distributed throughout the body, and while some of them are restricted to certain tissues or cells, others such as human galectin-1 (Gal-1) and human galectin-3 (Gal-3) are ubiquitous.[3] Gal-1 in particular, has been proven to participate in B-cell development and signalling,[4] T-cell immunity,[5] and the regulation of different inflammatory responses.[6] Gal-1 has been recently shown to promote bacterial infections,[7] and to have a prominent relationship with certain types of cancers,[8] where its increased expression has been related to different processes in the disease progression. In fact, it has been pointed out as a key player for cancer immunotherapy resistance.[9]

Galectins perform their biological functions through the recognition of specific βGal-containing epitopes present on glycoproteins and glycolipids. Their multimeric nature endows galectins with the ability to cross-link these glycoconjugates, which is at the heart of their regulatory mechanisms. These oligomerization phenomena strongly depend on the organization of their carbohydrate-recognition-domains (CRD), according to which galectins are in fact classified. Thus, prototype galectins, as Gal-1, display two identical CRDs that dimerize in a non-covalent manner, while tandem-repeat contain two distinct CRDs covalently linked through a peptide fragment. Finally, chimera-type, the only member being Gal-3, displays a single CRD connected to a long tail domain at the N-terminus through which

[a] S. Bertuzzi, Dr. A. Gimeno, R. Núñez-Franco, Dr. G. Bernardo-Seisdedos, S. Delgado, Dr. G. Jiménez-Osés, Dr. O. Millet, Prof. Dr. J. Jiménez-Barbero, Dr. A. Ardá
Molecular Recognition and Host-Pathogen Interactions, CIC bioGUNE
Basque Research and Technology Alliance, BRTA
Bizkaia Technology Park, Building 800, 48162 Derio, Bizkaia (Spain)
E-mail: jjbarbero@cicbiogune.es
aarda@cicbiogune.es

[b] Prof. Dr. J. Jiménez-Barbero
Ikerbasque—Basque Foundation for Science
48013 Bilbao, Bizkaia (Spain)

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

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.202003212.

Part of a Special Issue celebrating the 1000th Issue of Chemistry—A European Journal.

it oligomerizes. Although the quaternary organization of galectins is fundamental for their biological functions, in most cases it is not clear how it does influence ligand binding. A large part of our current knowledge about how galectins bind to their carbohydrate ligands has been obtained through X-ray crystallography[10] although, for these particular systems however, this cannot account for dynamic effects that could have an impact in ligand binding, including conformational plasticity or allostery.[11]

Gal-1 is a homodimer with a dimerization equilibrium constant in the low micromolar range.[12] This oligomeric architecture may be relevant for its biological activity,[13] and in fact Gal-1 mutants with altered dimerization properties have shown to have altered biological functions.[14] Early studies[15] postulated that lactose binding to Gal-1 occurs with a negative cooperativity between the two lectin binding sites, which was related to a global increase of protein dynamics in the low frequency motion range (picoseconds timescale), with a concomitant increase in conformational entropy. More recently,[16] based on hydrogen-deuterium exchange experiments, lactose binding was found to increase the exchange rates of Gal-1 residues located on the opposite side of the ligand-binding site, strongly suggesting the existence of protein allostery, which seems difficult to reconcile with the very fast picoseconds timescale of motion. These studies used lactose (Lac) as a ligand, which binds Gal-1 ca. twofold weaker than lactosamine (Galβ1-4GlcNAc, LacNAc, **1**, Figure 1). Glycan binding preferences of Gal-1, in fact, point to extended glycan chains terminating in LacNAc, both on N- and O-glycans.[17] Opposite to other galectins, further chemical modifications of this simple disaccharide epitope do not improve binding affinity for Gal-1.

Herein, we provide further experimental and theoretical evidences of allosterism operating in Gal-1. Motivated by the unexpected positive entropy contribution measured for LacNAc binding, opposite to that observed for Gal-3/LacNAc recognition,[18] the changes in protein flexibility upon ligand binding have been scrutinized. Our results show that upon LacNAc binding, but not upon binding to other lower affinity LacNAc-containing glycans, such as the blood group antigen (**4**), local protein flexibility increases in the μs-ms time scale. This is a rather different time frame dynamics to that previously reported in the ps timescale.[15] This transition to slow dynamic motions upon LacNAc binding influences the energy balance for the recognition process, and thus the affinity, through a favourable contribution to the binding entropy term. Remarkably, the combined experimental (relaxation dispersion NMR) and theoretical analysis (μs-MD) performed herein allowed identifying specific residues with a concerted dynamic behaviour that cluster at the dimerization interface, revealing a communication pathway between the two Gal-1 domains.

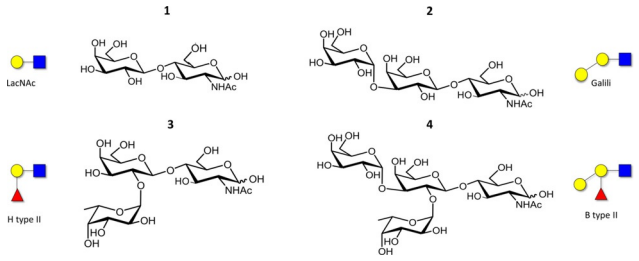

Figure 1. Structure and symbol representation of the oligosaccharides whose interaction with Gal-1 is studied herein.

Results and Discussion

The recognition of LacNAc (1), and LacNAc-containing glycans: blood group B antigen (4), and trisaccharide epitopes 2 and 3.

As mentioned above, individual galectins show a large variation in terms of affinity towards naturally occurring ligands as extensively and consistently highlighted in several studies.[19] Indeed, galectins exhibit rather different recognition patterns for sialylated glycans, polyLacNAc structures, and blood group antigens among others, and this fine specificity has been related with differential ligand interactions at regions adjacent to the canonical β-Gal binding site.[20] However, the detailed understanding of galectin-binding specificities is still modest given the lack of structural details for galectin complexes involving glycan structures larger than trisaccharides, as well as details regarding dynamics and flexibility of both interacting partners. We have recently addressed the impact of glycan flexibility on the binding of the A and B blood group tetrasaccharide antigens to Gal-3.[18] Interestingly, it has also been reported that Gal-1 and Gal-3 display opposing affinities towards these antigens.[21]

Isothermal titration calorimetry experiments

To obtain accurate information on the binding affinity and thermodynamics, isothermal titration calorimetry (ITC) experiments were performed for the four glycans (**1**–**4**). Fitting of the ITC binding isotherms to a single-site model yielded the dissociation binding constants (K_D) shown in Table 1. All of them are in the high-medium micromolar range, with the values for LacNAc (99 μM) in agreement with previously reported data.[17b] Data fitting to a sequential binding model (Table 2), as suggested by those previous studies,[15] was as good as or even better than the one-site model in terms of fitting quality (represented by χ^2). In this case, the first binding event displays better energetics than the second one, indicating a negative cooperativity between the two binding sites. This difference is

Table 1. Thermodynamic parameters for the binding of ligands **1**–**4** to Gal-1, as determined by ITC experiments. Data fitted to single-site binding model.

Ligand	ΔG [kcal mol^{-1}]	ΔH [kcal mol^{-1}]	$-T\Delta S$ [kcal mol^{-1}]	K_D [μM]
LacNAc (**1**)	−5.5	−5.3	−0.186	99
Galili (**2**)	−5.5	−5.2	−0.385	95
H type-II (**3**)	−4.3	−4.3	−0.469	319
B type-II (**4**)	−4.7	−4.3	−0.388	379

Table 2. Affinity constants obtained from ITC data fitting to one-site and sequential binding models. The quality of the fitting is provided by χ^2.

	One-site model		Sequential model		
Ligand	K_D	χ^2	K_{D1}	K_{D2}	χ^2
LacNAc (**1**)	99	1365	17	264	317
Galili (**2**)	95	2309	34	536	1151
H type-II (**3**)	319	169	196	1100	173
B type-II (**4**)	379	69	116	422	74

maximum for the LacNAc (**1**) and galili (**2**) ligands, for which $K_D2 = K_D \times 15$, while for the fucose (Fuc) containing ligands, this difference is smaller.

For either binding model, the affinity for the galili trisaccharide (**2**), which incorporates an additional Galα residue with respect to LacNAc **1**, was very similar to that obtained for **1**. Indeed, the binding enthalpy remained unaltered, strongly suggesting the absence of significant stabilizing intermolecular contacts provided by the Galα residue (see below in the NMR analysis). In contrast, the H type-II analogue (**3**), and especially the tetrasaccharide B type-II antigen (**4**), displayed somehow lower binding affinities compared to LacNAc **1.** In fact, the enthalpy contribution decreased for the fucosylated ligands, **3** and **4**. Also, variations in the binding enthalpy and entropy terms are not correlated, deviating from the commonly observed enthalpy-entropy compensation paradigm. Intriguingly, independently on the binding model used, the thermodynamic analysis (Table 1 and Supporting Information) revealed a positive entropy contribution to the binding for all the four ligands.[22] Although this entropy gain is always moderate (below 0.5 kcal mol^{-1}), it strongly contrasts with the loss of entropy (ca. 5 kcal mol^{-1}) observed for LacNAc binding to Gal-3.[18] This highlights the different molecular recognition mechanisms operating in both lectins.

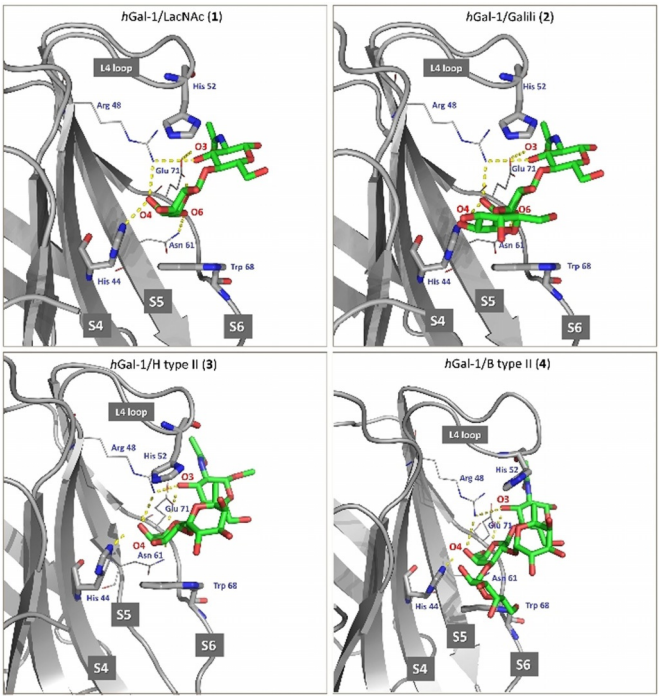

Figure 2. Molecular models for the complexes of Gal-1 with glycans **1**, **2**, **3**, and **4**, according to MD simulations (AMBER).

Generating the initial 3D models of the complexes

Initial 3D models of the ligand/Gal-1 complexes were built using the X-ray crystallographic structure reported for Gal-1:lactose,[23] by pair-fitting the binding residue(s) of each studied ligand to lactose followed by MD simulations as described in the experimental section. The complex formed with LacNAc (**1**) (Figure 2) is basically identical to that described in the X-ray crystallographic structure with lactose. Briefly, the Gal residue stacks on top of the indole moiety of Trp68, establishing key CH–π interactions,[24] with additional hydrogen bonding interactions involving residues His44, Arg48, Asn61 and Glu71 of the lectin and atoms Gal O4, Gal O5 and GlcNAc O3 of the ligand. The loop L4, which connects strands S4 and S5, is folded towards the ligand and narrows the binding site cavity. This loop has been shown to exhibit high conformational flexibility in *apo* structures, populating open and closed conformations.[15,25] According to the X-Ray crystallography data, His52, located at this loop, participates in hydrogen bonding with the Gal 2-OH of lactose. In our models for **1** and **2** (Figure 2), however, this is a transient interaction, occurring only ca. 25% along the 100 ns MD trajectories.

For **2** and **4**, the models show that the Galα residue is fairly close to the protein surface, although it does not provide additional van der Waals and/or hydrogen bonding interactions. Alternatively, the Fuc moiety, present in **3** and **4**, is close to the L4 loop, and His52 establishes transient (ca. 25%) hydrogen bonding interactions with Fuc O5 and/or Fuc 4-OH.

NMR experiments

STD-NMR*:* As an initial experimental validation of the proposed 3D complexes and to back up the experimental ITC results, information on the molecular basis of the interaction between Gal-1 and glycans **1**–**4** was obtained through NMR experiments,[26] starting with ^{1}H-STD-NMR (STD = saturation transfer difference),[27] which allows obtaining information on the ligand binding epitope. For all the ligands, significant STD signals were detected for the protons of the central β-Gal unit, in particular for H4, H5 and H6, which is the typical pattern for the interaction of β-Gal-containing saccharides with galectins.[18,28] Additionally, ligands **2** and **4**, which contain the Galα unit, showed evident STD signals for some protons of this residue (H1, H2 and H3), while the ligands containing Fuc (**3** and **4**) showed additional and significant STD effects for Fuc H1 (Figure 3 and Supporting Information). These data provide experimental evidence on the binding epitope of glycans **1**–**4**, which involves primarily the β-Gal ring, and with the Galα (in **2** and **4**) and Fuc moieties (in **3** and **4**) also in close proximity to the lectin surface, and can be satisfactorily explained involving

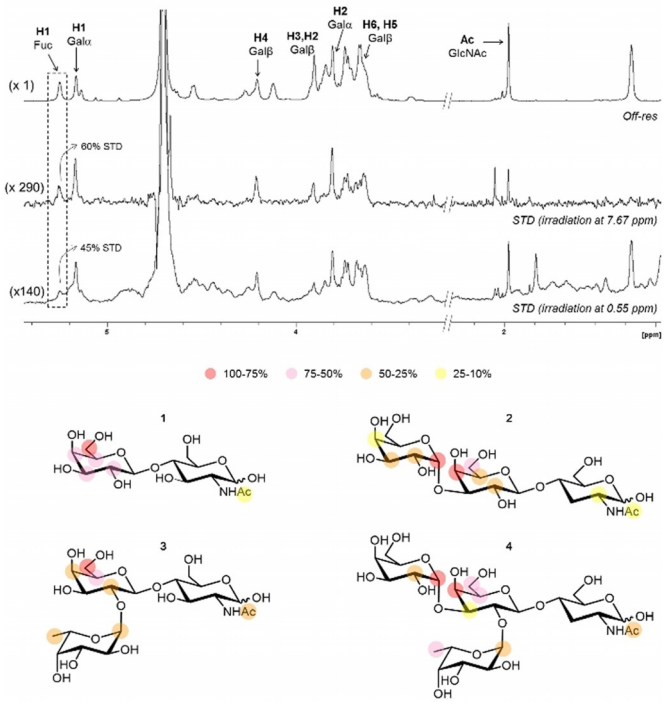

Figure 3. [1]H-STD-NMR results. Above: NMR spectra for the interaction of Gal-1 with tetrasaccharide **4**. On top, reference spectrum with annotations of the [1]H signals showing STD effect. Middle: STD spectrum with protein irradiation at the aromatic region. Below: STD spectrum with protein irradiation at the aliphatic region. Relative STD amplification factor is indicated for H1Fuc, which is larger for the aromatic irradiation STD. Below: Epitope mapping derived from [1]H- STD-NMR (irradiation at the aliphatic region) for the interaction of Gal-1 with ligands **1**, **2**, **3**, and **4**.

the recognition modes predicted by MD described above (Figure 2).

The presence of the loop L4 close to the binding site is a unique feature of Gal-1[23] and permits explaining the STD NMR effects observed for the Fuc residue. In fact, irradiation at the aromatic region of the protein (δ 7.7 ppm), increased the relative STD intensities for Fuc H1 with respect to the aliphatic irradiation, corroborating its proximity to His52. This latter result is in sharp contrast with that reported for the interaction of Gal-3 with **4**,[18] which demonstrated that the Fuc residue is exposed to the solvent, and does not interact with that lectin.

Chemical shift perturbation analysis: ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) NMR spectroscopy experiments were employed to analyse the chemical shift perturbation (CSP) of the amide signals of the lectin upon ligand addition and to obtain additional structural information on the sugar-protein molecular recognition processes from the lectin perspective.[25,29] The addition of 0.5, 1, 3, 5 and 10 equivalents of LacNAc (**1**) provided a progressive perturbation of the signals of specific amino acids. Most of them are included in the region Asn46-Val76, in β-strands S4, S5, S6 and L4 loop. This observation is again in agreement with the proposed binding mode described above (Figure 2) and reported by X-ray crystallography (PDB IDs 4Y1U, 4Q26 and 1W6P). Intriguingly, perturbations on several amino acids far beyond the binding site, especially on those located in F3-F4 sheets and close to the dimer interface (S1 and loop connecting S1-F2) were also detected (Figure 4A and Supporting Information). These results strongly suggest that the interaction with LacNAc induces changes on the whole structure of the protein. Similar observations have been described for the interaction with lactose[15,16] and lacto-*N*-neotetraose.[30]

The chemical shift perturbation profile for galili **2** was very similar to that of LacNAc, indicating comparable binding modes (Figure 4A and Supporting Information). This fact is in agreement with the MD simulations that show that the Galα residue establishes only short-lived interactions with the protein (see Supporting Information for details). Again, these observations contrast with our previous results for Gal-3,[18] where **2** displayed additional stabilizing contacts with several amino acids located at β-strand S3 and thus impacting on the measured CSP for these residues.

The CSP for the fucosylated glycans **3** and **4** (Figure 4B and C) were again similar to that of LacNAc. However, the perturbations corresponding to amino acids located at the L4 loop, in particular Ala51-Ala55, were markedly different (Figure 5). This fact indicates a different interaction of this loop region with the non fucosylated and fucosylated glycans, as shown by the STD-NMR analysis and predicted by the MD simulations. (Figure 4B,C and E,F and Supporting Information).

Additional structural information on the role of the Fuc residue in the binding process was inferred from the behaviour of the histidine side chain signals (His44 and His52) upon binding. His44 is conserved among galectins, located at strand S4 and consistently involved as hydrogen bonding acceptor from Galβ 4-OH, while His52 is located at the L4 loop, unique for Gal-1.

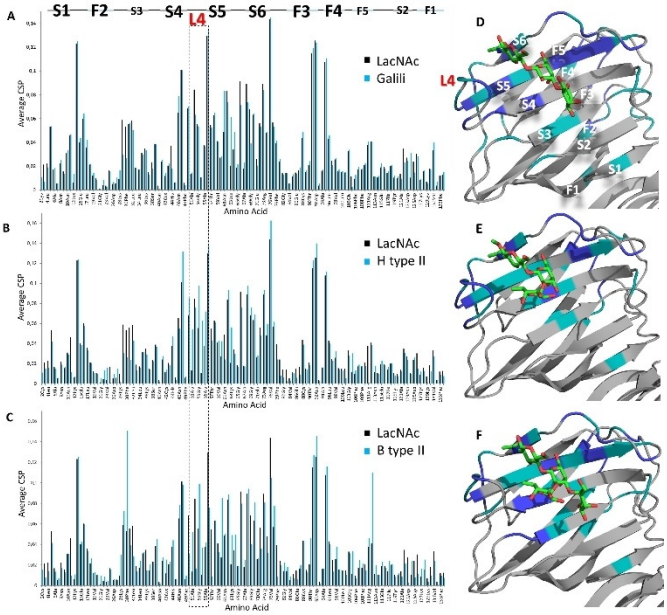

Figure 4. Comparison between the average CSP of the backbone amide signals of Gal-1 produced with **1** (in black, 10 equivalents) and in blue with A) ligand **2** (10 equiv), B) ligand **3** (15 equiv), and C) ligand **4** (15 equiv). The corresponding 3D models obtained through docking and MD simulations are shown on the right panels, D–F. The most perturbed amino acids are highlighted in dark blue (CSP over 2σ) and light blue (CSP 1–2σ).

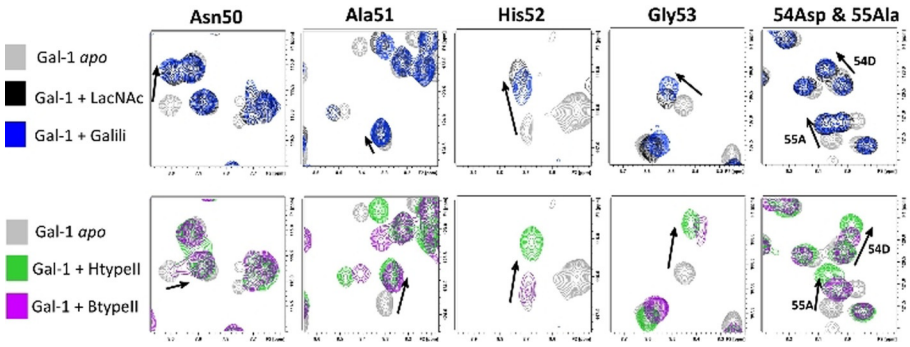

Figure 5. Expansions of ^{1}H-^{15}N HSQC spectra indicating the perturbations of amino acids at the L4 loop: apo Gal-1 (grey), upon the addition of 10 equivalents of **1** (black) and 10 equivalents of **2** (blue) (above), and comparison (below) with the same expanded regions recorded upon addition of 15 equivalents of **3** (green) and **4** (purple). The trend observed for the CSP measured for the L4 loop region (amino acids 51–55) is different upon addition of **1** and **2** versus **3** and **4**.

Thus, long-range ($^{2}J_{NH}$) ^{1}H-^{15}N HSQC experiments were acquired for Gal-1 *apo* and in the presence of **1** (without Fuc) and **3** (with Fuc) (Figure 6). In the *apo* form, only the signals corresponding to His52 were observed. Its pattern (Figure 5 A, see Supporting Information for details) revealed the existence of an equilibrium among the Nε2-H and the Nδ1-H tautomers and the protonated form.[31] Interestingly, addition of LacNAc **1** did not produce substantial changes on the shape of the His52 signals, suggesting no major changes on the equilibrium state (Figure 6 B). In contrast, the signals for His44 were now detected and the pattern pointed out to the presence of a very major Nε2-H tautomer, as expected for its role as hydrogen bond acceptor. Upon addition of **3**, the situation for His44 did not change with respect to the addition of **1**. In contrast, the signals for His52 became broader, even displaying multiple peaks, evidencing the presence of multiple states in slow-medium exchange regime in the chemical shift timescale (Figure 6 C). Thus, upon binding to ligand **3**, the chemical equilibrium for His52 is kept, although its dynamics is clearly altered, probably reflecting that instead of providing further contacts with the ligand, the loop L4 precludes a proper accommodation of the Fuc moiety.

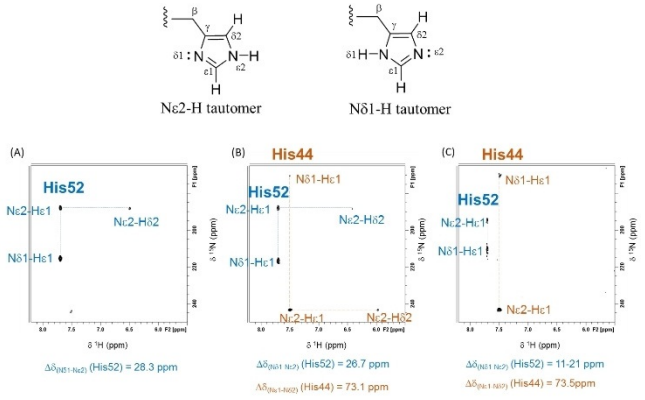

Figure 6. Above: tautomeric forms for the His side chains. Below: expansions of long range ^{1}H-^{15}N HSQC spectra of *apo* Gal-1 (A), upon addition of 10 equivalents of 1 (B) and 10 equivalents of 3 (C).

In summary of this section, ligands **1**–**4** share a similar binding mode to Gal-1, as deduced from STD and HSQC NMR experiments. Although the Galα and Fuc epitopes are located close to the protein surface in the so-called subsite B (strand S3) and close to the loop L4, respectively, MD simulations and NMR results support that the contacts of these moieties with the lectin are merely transient, with no clear stabilizing interactions taking place. These evidences are also in agreement with the ITC results described above, which show no enthalpy gain when the Galα and Fuc moieties are present. Moreover, the presence of the Fuc unit even decreases the enthalpy contribution. However, there is no clear explanation for the observed moderate entropy enhancement observed by ITC. Therefore, additional experiments and simulations focused on protein flexibility and dynamics were carried out.

Protein dynamics upon ligand binding: CLEANEX experiments

The long-range CSP observed for Gal-1 HSQC titration experiments with ligands **1**–**4**, together with the observed favourable binding entropy, are indicative of structural and dynamic changes in the whole structure of the protein upon ligand binding. Fast motions (in the ps timescale) of Gal-1[15] have been previously investigated by NMR through standard R_1 and R_2 experiments and highlighted the conformational entropy of the protein as a favourable contribution to the free energy of binding. However, the effects mentioned above regarding long-range chemical shift perturbations strongly suggest the presence of conformational fluctuations in a much slower timescale.[16,32] In order to detect local structural fluctuations and their potential relationship with sugar recognition, phase-modulated CLEAN chemical exchange spectroscopy NMR experiments (CLEANEX-PM)[33] were performed for the *apo* and bound forms of Gal-1. CLEANEX experiments allow detecting NH protons with fast exchange rates with water (exchange lifetimes in the 5–500 ms range) and are employed to estimate changes on the hydrogen bond stability or solvent accessibility of the backbone amides. In particular, the changes in exchange rates upon addition of medium/high- and low-affinity ligands such as LacNAc **1** and B type-II **4** were analysed. The CLEANEX

spectrum of Gal-1 showed 13 amide NH cross-peaks out of the 135 total ones (10%). They belong to amino acids located at the dimer interface and the loops connecting S2-F5, S3-S4, S4-S5, S6-F3 and F3-F4, which correspond to solvent exposed regions of the protein (Figure 7). Interestingly, they comprise residues directly involved in the binding as well as amino acids located far away from the binding site, rendering them as suitable probes to monitor changes on the protein structure.

The obtained average exchange rates for Gal-1 were $k_{ex} = 23\ s^{-1}$ for the *apo* form, and $10\ s^{-1}$ and $20\ s^{-1}$ for the LacNAc (**1**) and B type-II (**4**) bound forms, respectively. In fact, high protein:ligand ratios were employed in order to assure complete saturation of the protein. Thus, binding to the higher affinity ligand produced a global reduction on the exchange rates, while binding to the weaker affinity ligand produced minor changes. Remarkably, the four residues at the L4 loop were differently affected in the presence of both ligands. These results clearly demonstrate a different dynamic behaviour of the S4-S5 connecting loop in the presence of the fucosylated and non-fucosylated ligands, as also described above in the HSQC NMR analysis of His52. In fact, both results likely indicate that, in the presence of fucosylated sugars, His52, and in turn, the L4 loop, populates different conformations, which are less protected in average from water exchange than for non-fucosylated ligands.

As in the HSQC-based CSP experiments, significant changes in exchange rates were also detected for residues that are far away from the binding site. Particularly, the exchange rates of residues Ala1-Cys3 at the dimer interface, Ser38 at the S3-S4 loop, Ala94 at the F3-F4 loop, and Asn113 and Glu115 at the S2-F5 loop were reduced upon addition of **1**. The effect due to the presence of **4** was less pronounced, and did not follow a single trend. These results confirm that the whole structure of the protein is perturbed upon ligand binding and demonstrated that these effects are larger in the presence of the stronger binder. Similarly, previously reported NMR-HDX experiments indicated that lactose binding modulates HDX protection factors also for residues far beyond the binding site.[16]

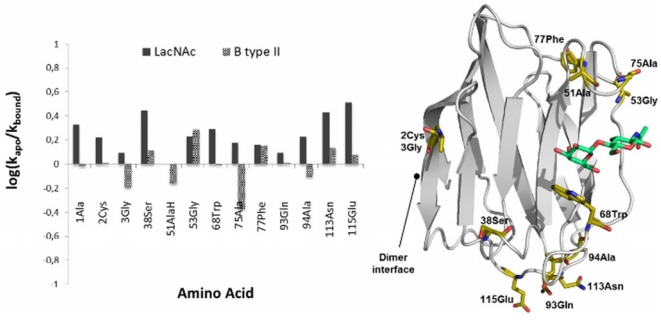

Figure 7. Left: NH-water exchange rates, k_{ex} (s^{-1}), obtained from CLEANEX-PM experiments for Gal-1 in the apo and bound forms upon addition of LacNAc (**1**) and B type-II (**2**). Right: X-ray structure of Gal-1:LacNAc (PDB ID: 1W6P). Amino acids with detected exchangeable NH amide protons are highlighted in yellow.

Relaxation dispersion NMR experiments

To fully discern the conformational fluctuations of Gal-1 in the *apo* state and in the presence of the ligands, relaxation dispersion (RD) NMR experiments were acquired for the backbone amides. The analysis of these experiments allowed identifying a large number of Gal-1 residues (up to 34) showing µs-ms dynamics upon LacNAc binding, with line-broadenings ranging from 102 to 768 Hz. Since the individual fitting of the RD profiles showed that, for a number of residues, there was a high degree of consistency in the obtained parameters (homogeneous k_{ex} and p_B values), a collective fitting procedure was employed. In the end, a group of 13 residues (Leu4, Ser7, Leu9, Arg18, Asp54, Ala55, Val76, Asp92, Ala121, Ala122, Asp123, Phe126, and Phe133) showed concerted dynamics at $380\ s^{-1}$ (k_{ex}) with an excited state showing a population (p_B) of about 1.5% (Figure 8C). Remarkably, this group of residues naturally clusters in the dimerization region of the protein, distal from the LacNAc binding site. The same RD experiment was performed for the Gal-1:**4** complex, and for Gal-3 in its *apo* and LacNAc bound forms, as control. For all these cases, only a limited number of residues (between 6 and 13) showed dispersion. Moreover, the RD-dispersions failed to statistically cluster into collective motions, indicating that they can be attributed to residual thermal motion.

Hence, the RD experiments support the notion that there is a conformational entropy gain of the protein upon ligand binding, consistent with the previous report.[15] Yet, the previous study focused in fast librations in the ps-ns timescale, more prone to capture thermal motion and less associated to functional dynamics. Herein, the observed µs dynamics associated to LacNAc binding provides the adequate experimental framework to support the idea of an allosteric transmission induced upon LacNAc binding.

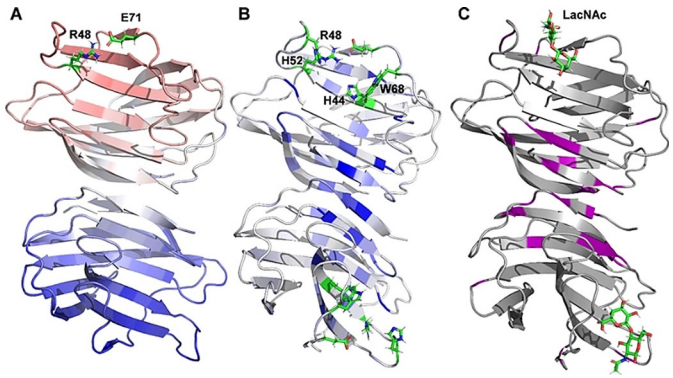

Figure 8. Long-range concerted dynamics in Gal-1. A) Color-coded distribution of whole-protein allosteric pathways determined from selected residues (shown as green sticks) in one binding site of *apo* Gal-1 through µs-MD simulations; red and blue colours indicate shorter/efficient and longer/inefficient pathways, respectively. B) Residues most frequently involved in the optimal and suboptimal pathways (color-coded in a blue gradient) calculated from selected residues (shown as green sticks) in both binding sites of *apo* Gal-1. C) Residues (in purple) showing concerted dynamics at $380\ s^{-1}$ as determined by transversal relaxation dispersion (RD) NMR experiments.

Allosteric communication analysis through MD simulations

In order to further support the NMR findings and analyse in detail the existence of allosteric effects, microsecond molecular dynamics simulations (μs-MD) were carried out, paying attention to possible pathways for dynamic correlation between amino acids at the binding site and any other amino acids in the protein, as described in the experimental section. As in the previous section, Gal-3 was also included as control, since it lacks changes in internal dynamics in the μs timescale upon LacNAc binding.

The analysis showed that, for Gal-1, the motion of the residues at the binding site propagates throughout the whole protomer even reaching the homodimeric interface (Figure 8A). Remarkably, the amino acids appearing at the highest frequency in the calculated pathways are concentrated in the internal β-strands, constituting the spine of the homodimer (Figure 8B). Fittingly, they match those determined experimentally to show concerted dynamics in the micro-to-millisecond time scale (Figure 8C). In contrast, for Gal-3, the correlated motions dissipated nearby the binding site (Figure S18–S20 in Supporting information). Accordingly, remarkable differences in the flexibility of the whole protein were also calculated for the *apo* and bound states of Gal-1 and Gal-3 with different ligands (Figure S21).

Conclusions

The interaction of human galectin-1 with *N*-acetyllactosamine (**1**), the blood B type-II antigen tetrasaccharide (**4**) and its two constituting trisaccharides (**2** and **3**) is favoured by entropy, in strong contrast with the observations for galectin-3 and most of lectin-sugar interaction events. In fact, the smaller disaccharide displays the best affinity for the lectin. The addition of the Fuc and Gal moieties (from **1** to **2** & from **1** to **3** and **4**) provides similar or weaker binding affinities, strongly suggesting that the Fuc and Gal do not establish stabilizing contacts with the lectin. Indeed, ligand-based NMR experiments indicate that these residues only provide, if any, minor contacts with galectin-1. Receptor-based HSQC chemical shift perturbation experiments, on the other hand revealed important effects for amino acids far from the binding site, which have been further assessed by water-exchange CLEANEX-PM experiments. Interestingly, the magnitude of those effects correlated with ligand affinity, very significant for the best affinity ligand, LacNAc (**1**). Moreover, relaxation dispersion NMR experiments have shown that there is important motion, in the microseconds-milliseconds time scale, for more than 30 amino acid residues upon LacNAc binding, many of them located distant from the binding site. More than ten of these residues cluster at the dimer interface. This behaviour is neither observed in the presence of the lowest affinity ligand (**4**) nor for LacNAc binding to galectin-3. Molecular dynamics simulations also predict the existence of dynamic correlation between the binding site and distant amino acids, reaching the lectin dimer interface upon LacNAc binding. In fact, once the first glycan molecule is bound, the second one is bound with smaller affinity, as deduced by the ITC measurements. The results presented herein show that sugar recognition by galectins is an extremely complex process that depends on many factors. Motions in the proteins may take place at different timescales, ligands display different flexibility and presentation of the epitopes both partners are important features to consider that affect the experimental observations. Indeed, despite their similarity, the prototype galectin-1 and the chimera-type galectin-3 show rather distinct features in their molecular recognition events. For instance, Gal-1 shows a noticeable preference to bind to terminal LacNAc structures in complex *N*-glycans, whereas Gal-3 preferentially recognizes internal LacNAc moieties.[17a] Their different conformational flexibility and protein architecture (dimer versus monomer) could underlie the observed features. In fact, binding enthalpies, binding entropies, and motion features are drastically different, highlighting the difficulty for achieving the full control of protein-sugar interactions. These findings shed light on structural and thermodynamic binding features of the analysed systems that, overall, can be used as clues for the rational design of compounds capable of selectively binding Gal-1.

Experimental Section

Materials

Ligands. Compound **1** (LacNAc), **2** (Galili), **3** (H type-II) and **4** (B type-II) were purchased from Elycityl (references GLY008, GLY074-2, GLY031-2 and GLY038-2).

Expression of unlabelled Gal-1. The gene encoding the carbohydrate recognition domain (CRD, 135 amino acids) of human galectin-1 was inserted into the pET21a expression vector. BL21 (D3) E. coli competent cells were transformed with the expression vector by heat shock method (42 °C for 90 s, 5 min in ice). After one night of incubation on agar plates in the presence of ampicillin at 37 °C, a single colony harbouring the expression construct was inoculated into 200 mL Luria Broth (LB) medium containing 100 μg mL^{-1} ampicillin and it was cultured overnight at 37 °C with shaking. A precise amount of the culture was then added to 2 L of fresh LB medium containing ampicillin so as to achieve a final OD600 of 0.1. Cells were grown at 37 °C until OD600 reached 0.6–1.2 and subsequently induced with 1 mM Isopropyl β-D-1-thio-galactopyranoside (IPTG). Growth continued for 3 h at 37 °C. The induced culture was harvested by centrifugation at 5500 rpm for 20 min. The pellet was then purified as explained in Purification of Gal-1 section.

Expression of ^{15}N labelled Gal-1. The gene encoding the carbohydrate recognition domain (CRD) of human galectin-1 was inserted into the pET21a expression vector. BL21 (D3) E. coli competent cells were transformed with the expression vector by heat shock method (42 °C for 90 s, 5 min in ice). After one night of incubation on agar plates in presence of ampicillin at 37 °C, a single colony harbouring the expression construct was inoculated into 5 mL of LB medium containing 100 μg mL^{-1} ampicillin for 6 h at 37 °C with shaking. The culture was centrifuged at 4400 rpm for 5 min. and the pellet was re-suspended in 1 mL of M9 medium containing ampicillin, transferred in a flask with 200 mL of the same medium and then incubated overnight at 37 °C with shaking. A precise amount of the overnight culture was then added to 2 L of fresh M9 labelled (^{15}N-NH_4Cl as nitrogen source) medium containing ampicillin so as to achieve a final OD600 of 0.1. Cells were grown at

37 °C until OD600 reached 0.6–1.2, then induced with 1 mM IPTG and again grown 3 h at 37 °C. The induced culture was harvested by centrifugation at 5500 rpm for 20 min. The pellet was purified as explained below.

Purification of Gal-1. The pellet obtained from the expression in BL21 E. coli cells was suspended in lysis buffer containing 22 mM Tris-HCl pH 7.5, 5 mM EDTA, 1 mM PMSF and 1 mM DTT (10 mL of lysis buffer were used per g of pellet) and left in ice for 30 min with shaking. The cell suspension was lysed by sonication in ice (60% amplitude, 12×20 s, with 59 s intervals between each burst). The crude extract was clarified by ultracentrifugation at 35000 rpm for 1 h at 4 °C. The soluble fraction was loaded onto 5 mL α-Lactose-Agarose resin (Sigma-Aldrich) previously equilibrated with equilibration buffer (50 mM TRIS pH 7.2, 150 mM NaCl). The column loaded was washed with 100 mL of equilibration buffer and then the lectin was eluted with 7 mL of elution buffer (150 mM α-Lactose pH 7.4 in PBS 1X). Gal-1 purity was checked by 4–12% SDS-PAGE and by LC-MS. To eliminate the lactose from the protein sample, a series of dialysis and washes with centrifuge filters (Sartorius Vivaspin 6 5000 MWCO) using fresh buffer (50 mM sodium phosphate, 150 mM NaCl, 2 mM DTT pH 7.4) were performed. The absence of lactose was checked by NMR. The addition of the reducing agent in the buffer of Gal-1 is justified by the presence of cysteine residues exposed to the solvent that could cause the formation of non-specific dimers or aggregates through intermolecular disulphide bonds.

NMR experiments. General information. The total volume for the NMR samples was 500 μL, using a precision NMR tube with 5 mm outer diameter (New Era Enterprises, Vineland, USA). The pH of the buffer was measured with pH meter Crison Basic 20 (Crison Instruments SA, Barcelona, Spain) and adjusted with the required amount of NaOH and HCl or NaOD and DCl.

Saturation Transfer Difference (STD) NMR. All the STD experiments[27] were acquired using Bruker AVANCE 2 600 MHz spectrometer equipped with standard triple-channel probe. The samples were prepared in deuterated phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 2 mM of dithiothreitol-d_{10} (DTT-d_{10}). The standard ratio ligand/Gal-1 used was 1:50, with the concentration of protein (unlabelled) set at 50 μM. Experiments at higher equivalents of ligand (ratio 1:100) were performed to amplify the STD effect and to confirm the preliminary results. In the case of **3**, the ratio used was 1:138. The temperature during the acquisition was 298 K for all the set of experiments. The 1D STD sequence from Bruker library with spoil and T2 filter (stddiff.3) was employed for STD experiments. STD spectra were acquired with 1028 scans, 2 s of saturation time using a train of 50 ms Gaussian-shaped pulses and 2 s of relaxation delay. The spin-lock filter used to remove the NMR signals of the macromolecule was set at 20 ms. The on- and off-resonance spectra were registered in an interleaved mode with the same number of scans. The on-resonance frequency was set for the aliphatic region between 0.55 and 0.85 ppm and for the aromatic region between 7.67 and 7.73 ppm, while the off-resonance frequency was set at 100 ppm. The STD NMR spectra were obtained by subtracting the on-resonance spectrum from the off-resonance spectrum. The STD Amplification Factor (STD-AF) and the percentage of STD (STD%) were calculated on the basis of the STD spectra. Reference experiments were acquired on samples containing only the protein as well as only the ligands under the same experimental conditions to verify the authenticity of the binding. No signals were detected in the blank STD NMR spectra of the ligands alone, made exception for the acetyl and methyl group of GlcNAc, GalNAc and Fuc moieties, respectively, as highlighted in Figures S3B and S4B, which displayed weak STD signals likely due to direct irradiation effects. The analysis of the spectra was carried out using the proton signal with the strongest STD effect as reference (100% of STD effect). On this basis the relative STD intensities for the others protons of the molecules were calculated.

Chemical Shift Perturbation (CSP) Analysis. The ^{1}H-^{15}N HSQC experiments were acquired using Bruker AVANCE 2 800 MHz spectrometer equipped with cryoprobe. The samples were prepared using ^{15}N-labelled Gal-1 at 100 μM in 90% phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 2 mM of dithiothreitol (DTT) and 10% D_2O. The ligands were titrated in the protein sample and at each intermediate point (0.5-1-3–5–10–15 equivalents of the corresponding sugar), an HSQC experiment was acquired. For this purpose, a standard HSQC sequence was set with 200 (T1)×1024 (T2) complex data points in ^{15}N and ^{1}H dimensions, respectively. The temperature during the acquisition was 298 K for all the experiments. The CSP data of the titrations were analysed using CcpNmr Analysis software. The average ^{1}H and ^{15}N CSP were calculated for NH groups of the protein backbone using the formula: $\Delta\delta\ (\mathrm{ppm}) = [(\Delta\delta_H{}^2 + (0.14\cdot\Delta\delta_N)^2)/2]^{1/2}$) and the results were plotted in graphics with the respective standard deviation.

CLEAN Chemical Exchange (CLEANEX-PM). The samples were prepared using ^{15}N uniformly labelled Gal-1 at 1 mM in 90% phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 2 mM DTT and 10% D_2O. The CLEANEX spectra[33] were acquired for Gal-1 alone, as well as for Gal-1 in presence of ligands **1** (12 equivalents) and **4** (10 equivalents). CLEANEX-PM experiments were performed using Bruker AVANCE 2 800 MHz spectrometer equipped with a cryoprobe. These experiments provided information about the exchange rates of NH groups of the protein with the bulk water. The setup has been optimized with TD of 2048 (F3) *x* 128 (F1) ^{15}N and ^{1}H dimensions, respectively and 4 points in the F2 dimension, corresponding to 3 different mixing times (25, 50 and 75 ms) and the reference spectrum. The temperature during the acquisition was set at 298 K for all the experiments. The ratio between the peak intensities in the CLEANEX spectra (V_i) and reference HSQC spectra (V_0) was considered as a measurement of the exchange rate. The analysis was carried out using CcpNmr Analysis software and $\frac{Vi}{V0}$ were fitted as function of the mixing time. Exchange rates, k_{ex}, were obtained from the fitting to Equation (1).

$$\frac{Vi}{Vo} = \frac{k}{k + R_1}\left[1 - e^{-(R_1+k)t}\right] \quad (1)$$

Where k is k_{ex} and R_1 is the effective NH relaxation rate during CLEANEX mixing time.

^{1}H-^{15}N long-range HMQC experiments. Analysis of histidine side chains.[31] In order to deeply analyse the role of two histidines of Gal-1 (His44 and His52) in the ligand binding process long-range HMQC experiments were performed. The spectra were acquired with TD of 1024 (F2)×160 (F1) for ^{1}H and ^{15}N dimensions, respectively and 96 scans were acquired. The temperature during the acquisition was set at 298 K. The samples were prepared using ^{15}N uniformly labelled Gal-1 at a concentration between 160 and 260 μM in 90% phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 2 mM DTT and 10% D_2O. ^{1}H-^{15}N Long-range HMQC were acquired for Gal-1 alone, as well as for Gal-1 in presence of ligands **1**, and **3** with a lectin:ligand ratio of 1:10.

^{15}N CPMG Relaxation Dispersion. Transversal relaxation dispersion experiments[34] were acquired on a Bruker AVANCE 2 800 MHz spectrometer equipped with a cryoprobe at 298 K using a relaxation compensated pulse CPMG sequence (80 ms total CPMG time) and

variable effective fields: 25, 50 (x2), 75, 100, 125, 150, 200, 250, 350, 450, 600, 800 (x2) and 1000 Hz. The experiments were set with 32 scans and a TD of 2048 (F3) x 180 (F2) for ^{1}H and ^{15}N dimensions,[35] respectively and 16 points in the F1 dimension. Different datasets were collected for the following samples: Gal-1:Apo, Gal-1:**1**, Gal-1:4, Gal-3:Apo and Gal-3:LacNAc. The samples were prepared using 15N uniformly labelled Gal-1 or Gal-3 at a concentration between 450 and 650 μM in 90% phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 10% D2O and the addition of 2 mM DTT as reducing agent only in the case of Gal-1. The 15N CPMG Relaxation Dispersion experiments were acquired for Gal-1 and Gal-3 alone and in the presence of ligands **1** and **4** with a ratio lectin:ligand of 1:20. Dispersion data were fit to the Carver and Richards equation using in-house Matlab© scripts, either to one field alone or simultaneously using data from two fields. Finally a collective fitting was done using different clustering residues. Duplicate data were used to obtain an estimation of the error and F-test statistics to validate the suitability of the different models adjusted.

Molecular modeling and MD simulations. General information about 100 ns MD simulations of Gal1/sugar complexes. The starting geometries for the initial modelling procedures were built based on the X-ray structures of Gal-1 complexed with *N*-acetyllactosamine (PDB ID: 1W6P). The structure of the sugar was superimposed in the binding site, employing the most populated conformation found for the free state (according to a standard NOE/molecular modelling approach). The complex structure was then submitted to 100 ns molecular dynamics simulation. Then, the MD simulations were performed using Amber16 program with the ff14SB force field parameters for protein and GLYCAM06j-1 for the oligosaccharides.[36] Thereafter, the starting 3D geometries were placed into a 12 Å octahedral box of explicit TIP3P waters, and counterions were added to maintain electroneutrality. Two consecutive minimizations were performed: 1) involving only the water molecules and ions, and 2) involving the whole system. The system was then heated and equilibrated in two steps: 1) 20 ps of MD heating the whole system from 0 to 300 K, followed by 2) equilibration of the entire system during 100 ps at 300 K. The equilibrated structures were the starting points for MD simulations (100 ns) at constant temperature (300 K) and pressure (1 atm) and constant volume. A detailed analysis of each MD trajectory (r.m.s.d., dihedral angles and hydrogen-bond evaluation) was accomplished using the cpptraj module included in Amber-Tools 16 package.

Microsecond Molecular Dynamics (μs-MD) simulations. These simulations were carried out with AMBER 18 package[37] implemented with ff14SB[38] and GLYCAM 06j-1[36] force fields for the proteins and carbohydrate ligands, respectively. Binding histidine residues (H44 and H52 in Gal-1 and H158 in *Gal*-3) were modeled in their Nδ1-H tautomeric state (reside name HID in Amber). Protein complexes were immersed in a water box with a 10 Å buffer of TIP3P[39] water molecules and neutralized by adding explicit Na^{+} or Cl^{-} counterions. A two-stage geometry optimization approach was performed. The first stage minimizes only the positions of solvent molecules and ions, and the second stage is an unrestrained minimization of all the atoms in the simulation cell. The systems were then heated by incrementing the temperature from 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. Harmonic restraints of 10 kcal mol^{-1} were applied to the solute, and the Andersen temperature coupling Scheme[40] was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Water molecules were treated with the SHAKE algorithm[41] such that the angle between the hydrogen atoms is kept fixed through the simulations. Long-range electrostatic effects were modelled using the particle mesh Ewald method.[42] An 8 Å cut-off was applied to Lennard–Jones interactions. Each system was equilibrated for 2 ns with a 2 fs time step at a constant volume and temperature of 300 K. Five independent production trajectories were then run for additional 1.0 μs under the same simulation conditions, leading to accumulated simulation times of 5.0 μs for each system.

Allosteric communication analysis through MD simulations. The Weighted Implementation of Suboptimal Paths (WISP)[43] was used for the analysis of dynamical networks. First, a correlation matrix (C_{ij}) is generated from 1.000 snapshots extracted every 1.0 ns from a converged μs-MD trajectory, by calculating the correlation motion among node-node pairs with Equation (2). In our model, nodes are defined by the whole-residue centre of mass, and two nodes are considered to be in contact if the mean distance between them along the MD simulation is 6 Å or less. The length of the edges connecting these nodes quantifies the degree of dynamic communication between pairs of connected nodes as defined in Equation (3). This pathway length is inversely proportional to the correlation motion between nodes, meaning that shorter w_{ij} values indicate tightly correlated or anticorrelated nodes, whereas larger values indicate less correlated nodes.

$$C_{ij} = \frac{\langle \Delta\vec{r}_i(t) \cdot \Delta\vec{r}_j(t) \rangle}{\sqrt{\langle \Delta\vec{r}_i(t)^2 \rangle \langle \Delta\vec{r}_j(t)^2 \rangle}} \quad (2)$$

$$w_{ij} = -\log(|C_{ij}|) \quad (3)$$

Then, Dijkstra's algorithm is used to generate all force-node paths, finding the shortest (i.e., optimal) path. To identify not only the optimal but also suboptimal pathways, WISP employs a bidirectional search. Suboptimal pathways are defined as those closest in length to the optimal one, but not including it. The available code rapidly calculates both optimal and suboptimal communication pathways between two user-specified residues of a protein. For each galectin, 100 pathways were calculated between selected binding site resides (H44, R48, H52, W68, E71 for *Gal*-1; R144, H158, R162, W181, E184 for *Gal*-3) and all the other residues of the protein. These paths were recalculated for the *apo* and *bound* forms with to two different ligands (LacNAc **1** and B type-II **4**).

Isothermal titration calorimetry (ITC). Isothermal Titration Calorimetry experiments were performed using MicroCal PEAQ-ITC calorimeter. *Gal*-1 in the presence of the ligand samples (**1**–**4**) were prepared in phosphate-buffered saline (50 mM sodium phosphate, 150 mM NaCl, pH 7.4) with 1 mM TCEP as reducing agent. The concentration of the protein solution was set between 100–200 μM and that of the sugar stock between 5–9 mM. During the automated experiment, small amounts of the sugar solution (2–5 μL) were titrated into a cell containing the protein solution and the heat dispersed was detected. The analysis of the curves was accomplished using the MicroCal Origin 7 software. The association constants and the thermodynamic parameters obtained from the fit of the titration profile to a single-site binding and to a sequential binding model are reported in Tables 1 and 2. Examples of titration profiles for each complex with data fitting to single-site binding model are reported in Figure S22.

Acknowledgements

This research was supported by the European Research Council (ERC-2017-AdG, project 788143-RECGLYCANMR to J.J.-B.), Agencia Estatal Investigacion of Spain (AEI; grant RTI2018-094751-B-C21 to J.J.-B., RTI2018-099592-B-C22 to G.J.O, RTI2018-101269-B-I00 to O. M., and Ramón *y* Cajal Contract to A. A.) and the Severo Ochoa Excellence Accreditation (SEV-2016-0644 to J.J.-B.). We also thank Instituto de Salud Carlos III of Spain, ISCIII (grant PRB3 IPT17/0019 to A. G.) and the Mizutani Foundation for Glycoscience (grant 200077 to G.J.O.).

Conflict of interest

The authors declare no conflict of interest.

Keywords: allostery · blood group antigens · galectin · glycan molecular recognition · NMR spectroscopy

[1] a) N. Giovannone, L. K. Smith, B. Treanor, C. J. Dimitroff, *Front. Immunol.* **2019**, *9*, 2839; b) V. C. M. Allo, M. A. Toscano, N. Pinto, G. A. Rabinovich, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE97–SE107; c) S. Thiemann, L. G. Baum, *Annu. Rev. Immunol.* **2016**, *34*, 243–264; d) S. Sato, C. St-Pierre, P. Bhaumik, J. Nieminen, *Immunol. Rev.* **2009**, *230*, 172–187.

[2] a) A. J. Cagnoni, J. M. Pérez Sáez, G. A. Rabinovich, K. V. Mariño, *Front. Oncol.* **2016**, *6*, 109; b) C. J. Dimitroff, *Cancer Res.* **2015**, *75*, 3195–202.

[3] a) J. Nio-Kobayashi, *Anat. Sci. Int.* **2017**, *92*, 25–36; b) L. Johannes, R. Jacob, H. Leffler, *J. Cell Sci.* **2018**, *131*, jcs.208884.

[4] a) L. Gauthier, B. Rossi, F. Roux, E. Termine, C. Schiff, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13014–13019; b) F. Mourcin, C. Breton, J. Tellier, P. Narang, L. Chasson, A. Jorquera, M. Coles, C. Shiff, S. J. C. Mancini, *Blood* **2011**, *117*, 6552–6561.

[5] F. Cedeno-Laurent, C. J. Dimitroff, *Clin Immunol.* **2012**, *142*, 107–16.

[6] a) V. Sundblad, L. G. Morosi, J. R. Geffner, G. A. Rabinovich, *J. Immunol.* **2017**, *199*, 3721–3730; b) V. C. Martínez Allo, V. Hauk, N. Sarbia, N. A. Pinto, D. O. Croci, T. Dalotto-Moreno, R. M. Morales, S. G. Gatto, M. N. Manselle Cocco, J. C. Stupirski, A. Deladoey, E. Maronna, P. Marcaida, V. Durigan, A. Secco, M. Mamani, A. Dos Santos, A. Catalán Pellet, C. Pérez Leiros, A. Rabinovich, M. A. Toscano, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 6630–6639.

[7] A. L. Lujan, D. O. Croci, J. A. Gambarte, Tudela, A. D. Losinno, A. J. Cagnoni, K. V. Mariño, M. T. Damiani, G. A. Rabinovich, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E6000–E6009.

[8] J. M. Cousin, M. J. Cloninger, *Int. J. Mol. Sci.* **2016**, *17*, 1566.

[9] D. K. Nambiar, T. Aguilera, H. Cao, S. Kwok, C. Kong, J. Bloomstein, Z. Wang, V. S. Rangan, D. Jiang, R. von Eyben, R. Liang, S. Agarwal, A. D. Colevas, A. Korman, C. T. Allen, R. Uppaluri, A. C. Koong, A. Giaccia, Q. T. Le, *J. Clin. Invest.* **2019**, *129*, 5553–5567.

[10] S. Kamitori, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE41-SE50.

[11] a) C. M. A. Guardia, D. F. Gauto, S. Di Lella, G. A. Rabinovich, M. A. Martí, D. A. Estrin, *J. Chem. Inf. Model.* **2011**, *51*, 1918–1930; b) A. Diniz, H. Coelho, J. S. Dias, S. J. van Vliet, J. Jiménez-Barbero, F. Corzana, E. J. Cabrita, F. Marcelo, *Chem. Eur. J.* **2019**, *25*, 13945–13955.

[12] a) M. Cho, R. D. Cummings, *J. Biol. Chem.* **1995**, *270*, 5198–5206; b) V. Giudicelli, D. Lutomski, M. Levi-Strauss, D. Bladier, R. Joubert-Caron, R. M. Caron, *Glycobiology* **1997**, *7*, 323.

[13] a) M. Dias-Baruffi, H. Zhu, M. Cho, S. Karmakar, R. P. McEver, R. D. Cummings, *J. Biol. Chem.* **2003**, *278*, 41282–93; b) S. R. Stowell, M. Cho, C. L. Feasley, C. M. Arthur, X. Song, J. K. Colucci, S. Karmakar, P. Mehta, M. Dias-Baruffi, R. P. McEver, R. D. Cummings, *J. Biol. Chem.* **2009**, *284*, 4989–4999.

[14] a) L. A. Earl, S. Bi, L. G. Baum, *Glycobiology* **2011**, *21*, 6–12; b) T. Miura, M. Takahashi, H. Horie, H. Kurushima, D. Tsuchimoto, K. Sakumi, Y. Nakabeppu, *Cell Death Differ.* **2004**, *11*, 1076–1083.

[15] I. V. Nesmelova, E. Ermakova, V. A. Daragan, M. Pang, M. Menéndez, L. Lagartera, D. Solís, L. G. Baum, K. H. Mayo, *J. Mol. Biol.* **2010**, *397*, 1209–30.

[16] C. H. Chien, M. R. Ho, C. H. Lin, S. D. Hsu, *Molecules* **2017**, *22*, 1357.

[17] a) N. A. Kamili, C. M. Arthur, C. Gerner-Smidt, E. Tafesse, A. Blenda, M. Dias-Baruffi, S. R. Stowell, *Proteomics* **2016**, *16*, 3111–312; b) K. Shimura, Y. Arata, N. Uchiyama, J. Hirabayashi, K. Kasai, *J. Chromatogr. B* **2002**, *768*, 199–210.

[18] A. Gimeno, S. Delgado, P. Valverde, S. Bertuzzi, M. A. Berbís, J. Echavarren, A. Lacetera, S. Martín-Santamaría, A. Surolia, F. J. Cañada, J. Jiménez-Barbero, A. Ardá, *Angew. Chem. Int. Ed.* **2019**, *58*, 7268–7272; *Angew. Chem.* **2019**, *131*, 7346–7350.

[19] a) J. Hirabayashi, T. Hashidate, Y. Arata, N. Nishi, T. Nakamura, M. Hirashima, T. Urashima, T. Oka, M. Futai, W. E. Muller, F. Yagi, K. Kasai, *Biochim. Biophys. Acta.* **2002**, *1572*, 232–254; b) M. I. Nielsen, J. Stegmayr, O. C. Grant, Z. Yang, U. J. Nilsson, I. Boos, M. C. Carlsson, R. J. Woods, C. Unverzagt, H. Leffler, H. H. Wandall, *J. Biol. Chem.* **2018**, *293*, 20249–20262.

[20] P. Modenutti, J. I. Blanco Capurro, S. Di Lella, M. A. Martí, *Front. Chem.* **2019**, *7*, 823.

[21] S. R. Stowell, C. M. Arthur, P. Mehta, K. A. Slanina, O. Blixt, H. Leffler, D. F. Smith, R. D. Cummings, *J. Biol. Chem.* **2008**, *283*, 10109–10123.

[22] a) T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387–429; b) E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719–725.

[23] M. F. López-Lucendo, D. Solís, S. André, J. Hirabayashi, K. I. Kasai, H. Kaltner, A. Romero, *J. Mol. Biol.* **2004**, *343*, 957–970.

[24] J. L. Asensio, A. Arda, F. J. Cañada, J. Jimenez-Barbero, *Acc. Chem. Res.* **2013**, *46*, 946–954.

[25] PDB ID: 3W59. H. Suburi, T. Tanaka, N. Kunishima, *Crystal structure of Galectin-1 in the lactose-unbound state* **2014**, https://doi.org/10.2210/pdb3W59/pdb.

[26] a) P. Valverde, J. I. Quintana, J. I. Santos, A. Ardá, J. Jiménez-Barbero, *ACS Omega* **2019**, *4*, 13618–13630; b) A. Gimeno, P. Valverde, A. Arda, J. Jimenez-Barbero, *Curr. Opin. Struct. Biol.* **2020**, *62*, 22–30.

[27] a) M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784–1788; *Angew. Chem.* **1999**, *111*, 1902–1906; b) B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864–890; *Angew. Chem.* **2003**, *115*, 890–918; c) A. Viegas, J. Manso, F. L. Nobrega, E. J. Cabrita, *J. Chem. Educ.* **2011**, *88*, 990–994.

[28] S. Bertuzzi, J. I. Quintana, A. Arda, A. Gimeno, J. Jimenez-Barbero, *Front. Chem.* **2020**, *8*, 593.

[29] A. Arda, J. Jimenez-Barbero, *Chem. Commun.* **2018**, *54*, 4761–4769.

[30] J. Bonzi, O. Bornet, S. Betzi, B. T. Kasper, L. K. Mahal, S. J. Mancini, C. Schiff, C. Sebban-Kreuzer, F. Guerlesquin, L. Elantak, *Nat. Commun.* **2015**, *6*, 6194–6206.

[31] a) A. A. Van Dijk, R. M. Scheek, K. Dijkstra, G. K. Wolters, G. T. Robillard, *Biochemistry* **1992**, *31*, 9063–9072; b) J. G. Pelton, D. A. Torchia, N. D. Meadow, S. Roseman, *Prot. Sci.* **1993**, *2*, 543–558.

[32] F. A. Chao, R. A. Byrd, *Emerging Top. Life Sci.* **2018**, *2*, 93–105.

[33] a) T. L. Hwang, S. Mori, A. J. Shaka, *J. Am. Chem. Soc.* **1997**, *119*, 6203–6204; b) T. L. Hwang, P. C. M. van Zijl, S. Mori, *J. Biomol. NMR* **1998**, *11*, 221–226.

[34] P. Neudecker, P. Lündstrom, L. E. Kay, *Biophys. J.* **2009**, *96*, 2045–2054.

[35] J. P. Loria, M. Rance, A. G. Palmer, *J. Biomol. NMR* **1999**, *15*, 151–155.

[36] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622–655.

[37] D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, AMBER 2018, University of California, San Francisco, **2018**.

[38] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[39] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.

[40] a) H. C. Andersen, *J. Chem. Phys.* **1980**, *72*, 2384 – 2393; b) T. A. Andrea, W. C. Swope, H. C. Andersen, *J. Chem. Phys.* **1983**, *79*, 4576 – 4584.
[41] S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 952 – 962.
[42] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.
[43] A. T. Van Wart, J. Durrant, L. Votapka, R. E. Amaro, *J. Chem. Theory Comput.* **2014**, *10*, 511 – 517.

Manuscript received: July 7, 2020
Accepted manuscript online: August 11, 2020
Version of record online: October 29, 2020