Supporting Information

Small Lectin Ligands as a Basis for Applications in Glycoscience and Glycomedicine

Paul V. Murphy,*[a,b,c] Ashis Dhara,[a] Liam S. Fitzgerald,[a,b,c] Eoin Hever,[a] Saidulu Konda,[a] Kishan Mandal[a]

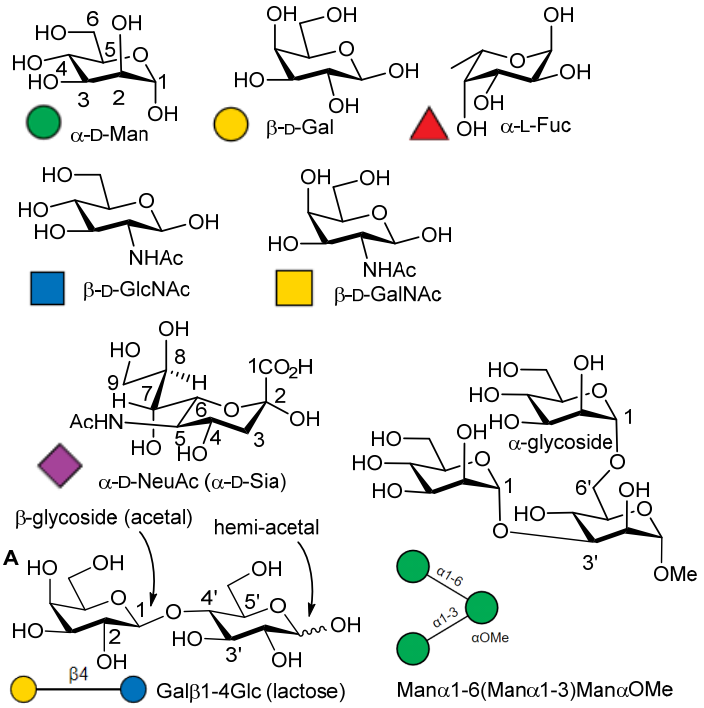

Fig. S1 Haworth chair perspectives of selected monosaccharides and oligosaccharides and corresponding symbol nomenclature for glycans (SNFG).

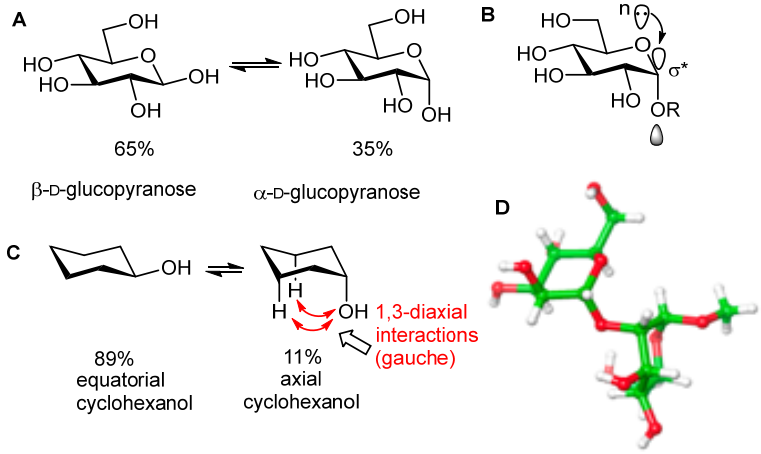

Fig. S2 A. Equilibrium concentrations of glucopyranoses. B. Hyperconjugation explanation for increased stabilisation for axial anomer (endo-anomeric effect). C. Equilibrium concentrations of cyclohexanol conformers. D. Crystal structure for Manα1-2Manα1-OMe (Cambridge Crystallographic Database Centre identifier = FABYOW11, www.ccdc.cam.ac.uk) shows ϕ= -g (-55, -60) at both glycosidic linkages consistent with influence of the exo-anomeric effect and minimisation of steric repulsive interactions. The angle for the 1-2 linkage in the X-ray structure is +17 which approaches ψ= syn.

Supplementary notes on interglycosidic torsional angles. For a fuller account on glycan structural preferences please consult sources referred to in the main text.

Φ for α-glycopyranosides: Fig. S3 depicts staggered conformers that result from rotation about the C1-O1 bond of α-Man*p*; the exo-anomeric effect stabilises exo-anti-Φ and exo-syn-Φ conformers, but exo-syn-Φ is preferred as it has only one steric repulsive gauche interaction, compared to two for exo-anti, showing the combined effect of stereoelectronic and steric repulsive interaction. The glycoside torsion Φ, defined herein by atoms H1-C1-O1-R', approximates to -60° for the exo-syn-Φ. 1 These values are estimates typical for conformational analysis of hydrocarbons. Deviation is expected when there are oxygen substituents due to differences in C-O bond lengths and COC bond angles. The exo-syn conformer could be defined as within the range -60±30°. ϕ is ~+60° if defined for O5-C1-O1-R' as is the definition in many publications.

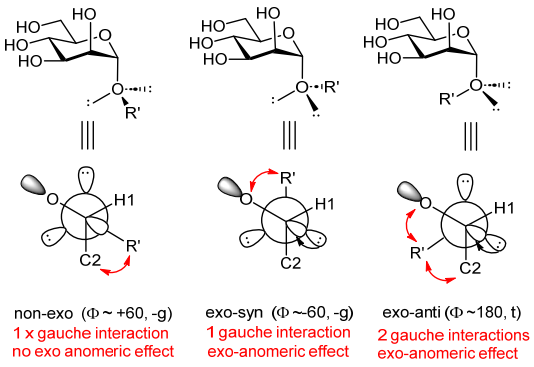

Fig. S3 The exo-anomeric effect is the stabilisation of exo-syn and exo-anti over non-exo, by hyperconjugation. Steric repulsive interactions imply that exo-syn is preferred.

Φ for β-D-glycopyranosides: Then exo-syn-Φ conformer is again preferred (see Fig. S4C), with Φ, defined by atoms H1-C1-O1-R', approximating in this case to +60°.

Φ for NeuAc-Gal linkages. The presence of the carboxylic acid substituent at the glycosidic carbon, reduces the energy difference between NeuAc's exo-syn-Φ and exo-anti-Φ conformers.[1] Thus, Φ is potentially more flexible in NeuAc(α2-3)Gal and NeuAc(α2-6)Gal, than Φ for β-galactopyranosides and α-mannopyranosides. The crystal structure analysis by Wormwald et al[13] showed exo-syn-Φ frequently found in bound conformers of NeuAc(α2-3)Gal.

Ψ when the glycoside is derived from a secondary alcohol. While there is no exo-anomeric effect for Ψ, (defined by CY-OY- CX'-HX', where Y = anomeric C/O), there is a preference for syn-Ψ (Fig. 4D), observed in disaccharides such as lactose (Galβ1-4Glc), NeuAc(α2-3)Gal, Manα1-2Man, where the CY-OY and CX'-HX' bonds have a near eclipsed conformation (syn-Ψ). Organic Chemistry students learn that eclipsed conformations of ethane are disfavoured compared to staggered conformers. We understand the near eclipsing syn (non-staggered) preference for Ψ is influenced by (i) repulsive syn-1,3-diaxial (syn pentane/Hassel Ottar) interactions that occur in staggered disaccharide Ψ conformers and (ii) a lower barrier to rotation about a C-O bond linkage (~1 kcal/mol for the C-O bond in methanol, compared to ~3 kcal/mol for the ethane C-C bond).[2] If there is a higher barrier to rotation or if repulsive syn-1,-3-diaxial interactions are less significant then Ψ (anti) may have significant population in such disaccharides.

ψ when the glycoside is derived from a primary alcohol. The staggered anti-Ψ conformer, which places large substituents anti-periplanar, has increased preference compared to syn-Ψ in NeuAc(α2-6)Gal and Manα1-6Man[3] and related linkages, as repulsive syn-1,3-diaxial type interactions are absent (Fig. 4E).

[1] C. O. Soares, A. S. Grosso, J. Ereno-Orbea, H. Coelho and F. Marcelo, Front. Mol. Biosci., 2021, 8, 727847.

[2] Theoretical and experimental barriers to rotation in methanol and ethane are reported here. A. Chung-Phillips and K. A. Jebber, J. Chem. Phys., 1995, 102, 7080-7087.

[3] E. W. Sayers, J. H. Prestegard, Biophys. J. 2000, 79, 33133329

The ω torsion. Free rotation about the C5-C6 single bond (ω-torsion) of pyranosides is influential for 1,6-linked disaccharides and NeuAcα-2,6Gal derivatives. The ω-torsion adopts the gauche-gauche (*gg*), trans-gauche (*tg*) or gauche-trans (*gt*) conformers (Fig. 5). There are high preferences for *gg* and *gt* conformers, and a low population of *tg*, for glucopyranose.[4,5] In contrast, for galactopyranose in water, that there is increased preference (>70%) for the *gt* conformer with a low *gg* population.[6] Molecular dynamics (MD) calculations using explicit water accurately predicts the ω torsion population, showing that water disrupts intramolecular hydrogen bonding allowing the *gg*, *gt* and *tg* populations, to be determined, not by intramolecular H-bonding, but by internal electronic and steric repulsions.[7] For Man and GlcNAc, the conformer preferences are like those of glucopyranoses.[8]

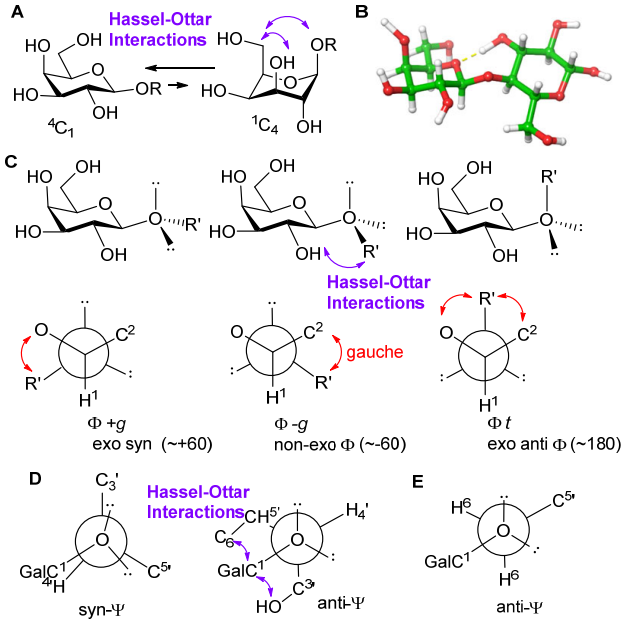

Fig. S4 A. Hassel Ottar (syn 1,3-diaxial) interactions,[9] destabilise $^1C_4$ relative to $^4C_1$ B. X-ray crystal structure of lactose has exo-syn-Φ; syn-Ψ. C. Exo-anomeric effect and steric repulsive interactions favour exo-syn Φ for β-Galp. D. Low energy conformers from O1-C4′ bond rotation in disaccharides like lactose, favours syn-Ψ in 1-4 linkage; 1,3-syn-diaxial interactions disfavours staggered anti-Ψ. E. 1-6 linkage favours anti-Ψ.

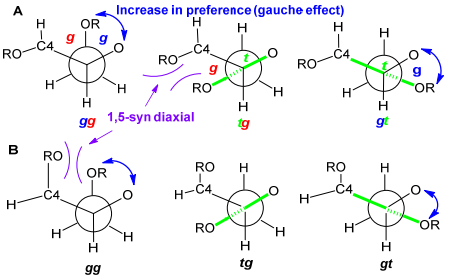

Fig. S5 Newman projections showing *gg*, *tg* and *gt* conformers viewing along C-5 to C-6 bonds of: A. Glc/Man/GlcNAc (*gg* and *gt* preferred); B. Gal/GalNAc (*tg* and *gt* preferred).

[4] (a) H Ohrui, Y Nishida, M Watanabe, H Hori, H Meguro *Tetrahedron Lett* 26, 3251–3254 (1985). (b) Rockwell, G. D.; Grindley, T. B. J. Am. Chem. Soc. 1998, 120, 10953-10963.
[5] I. Tvaroška, F. R. Taravel, J. P. Utille, J. P. Carver Carbohydr. Res. 2002, 337, 353–367.
[6] C Thibaudeau, R. Stenutz, B. Hertz, T. Klepach, S. Zhao, Q. Wu, I. Carmichael, A. S. Serianni, J. Am. Chem. Soc. 2004, 126, 15668–15685.
[7] K. N. Kirschner, R. J. Woods, Proc. Natl. Acad. Sci. (USA), 2001, 98, 10541–10545
[8] K. Bock and J. Ø. Duus, *J. Carbohydr. Chem.*, 1994, **13**, 513-543.
[9] O. Hassel, B. Ottar, Acta Chem. Scand. 1947, 1, 929-942.

Table S1: Φ and Ψ (°) for interglycosidic torsions between galactopyranose and glucopyranose residues in galectin ligands from co-crystal structures

Ligand	Galectin	Φ	Ψ	PDB code
Lactose	1	61	10	1W6O
Lactose	3	51	12	1KJL
Lactose	4C	53	10	4YM3
3'-Sulfo-lactose	4C	61	10	4YM2
3'-Sulfo-lactose	4N	55	19	5DUW
Lactose	7	61	4	4GAL
Lactose	8C	61	-7	3VKM
Lactose	8N	50	18	3AP4
3'-Sialyl Lactose	8N	54	14	3AP7
3'-Sulfo-lactose	8N	53	11	3AP6
3'-Sialyl Lactose	9C	54	19	3NV4
Lactose	9N	60	20	3EAK
Lactose	10	54	-20	6A1T

Data is taken from co-crystal structures with coordinates downloaded from the protein databank (www.rcsb.org). Φ is defined by H1'-C1'-O'-C4 and Ψ by C'1-O1'-C4-H4. While there are no H atoms in the crystal structures, the H atoms were added to the coordinates downloaded from www.rcsb.org using Maestro (freely available for academic use from schrodinger.com) and this was used to measure the torsion angles.

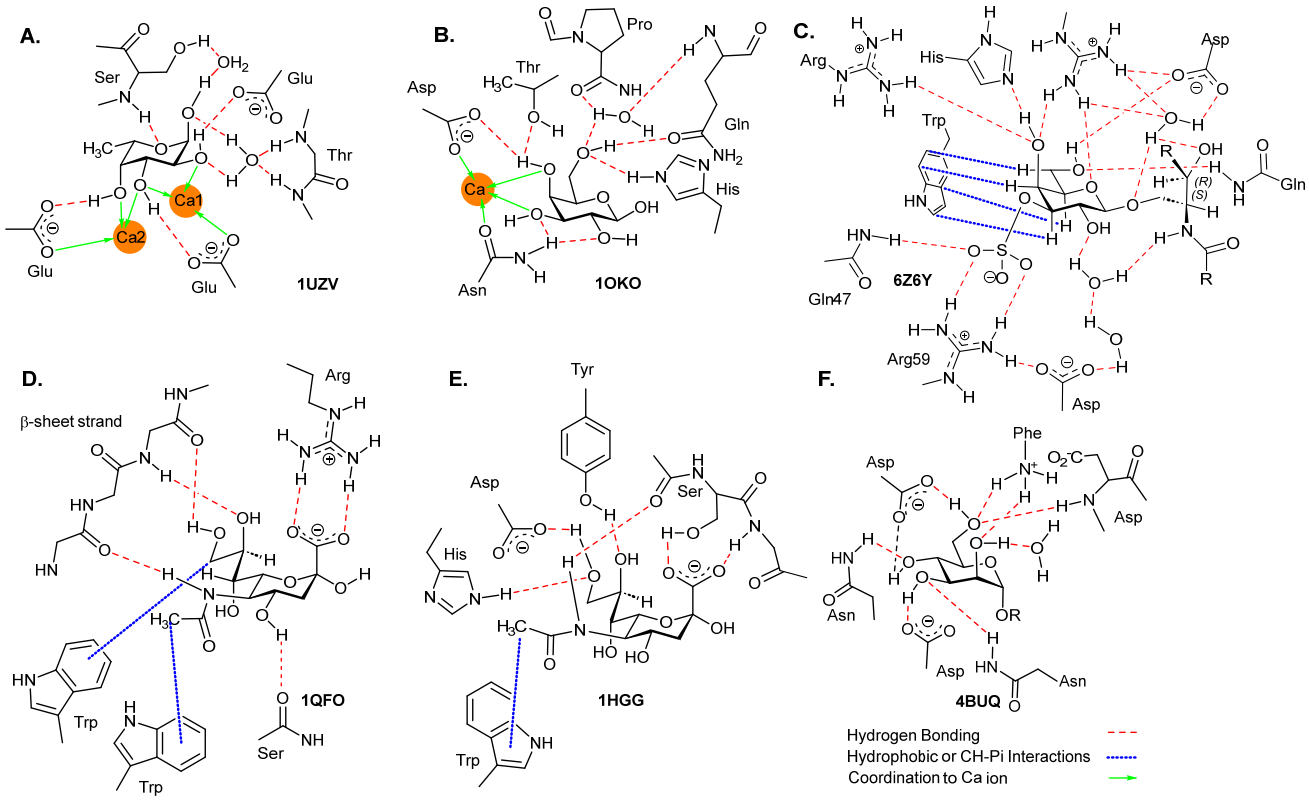

Fig. S6. A selection of potential lectin-ligand binding interactions depicted with Chemdraw. Shown are A. Fuc-LecB; B. Gal-LecA; C. β-Galactoside (sulfatide)-galectin-8N; D. Neu5Ac-mouse siglec-1); E. Neu5Ac-hemagglutinin; F. Man-FimHLD. The ligand-lectin interactions have affinities from mM to mM range. PDB identifier codes are provided.

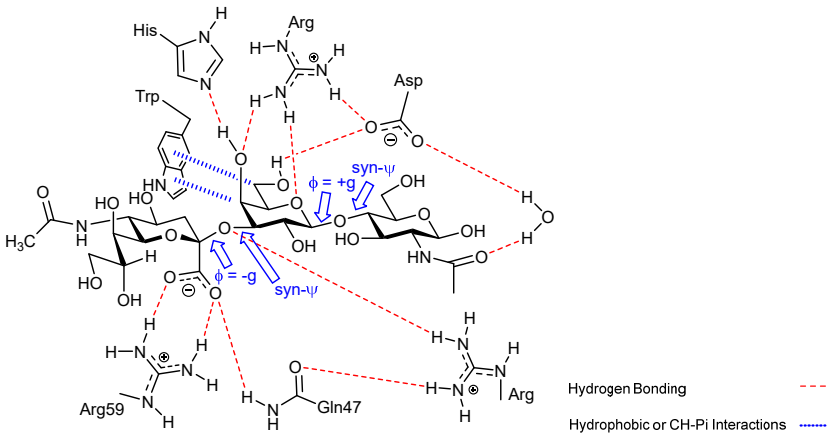

Fig. S7. 3'SL binding mode with galectin-8N. Charged H-bonding interactions with arginine from the carboxylate give higher affinity interactions with galectin-8. The wider H-bonding network arises and CH-p interactions common for galectin recognition of Gal. For sulfatide, the CH-Pi interactions involves the α-Gal face defined by H atoms ~3A distance from indole atoms are shown. The ligand bound structure has glycosidic torsion angles that correspond to a low energy ligand conformation.

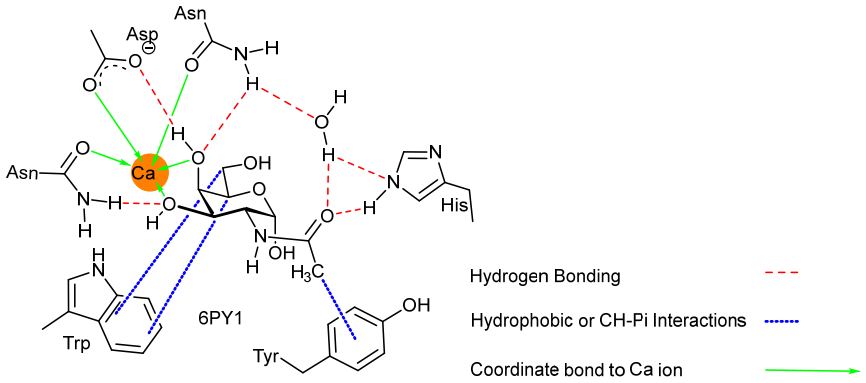

Fig. S8. GalNAc binding mode to human macrophage galactose C-type lectin (6PY1).

Table S2: Ligand Binding Affinities for Galectins (K_d μM) References specific to data cited are given below Table.

Compound	Gal-1	Gal-2	Gal-3	Gal-4C	Gal-4N	Gal-7	Gal-8C	Gal-8N	Gal-9C	Gal-9N
Methyl β-D-Gal	10000[1]	13000[10]	4100[1]	10000[1]	6600[1]	4800[9]	>30000[1]	6300[1] 4370[4]	8600[1]	3300[1]
Methyl α-D-Gal	>10000[2]	>20000[2]	2700[2]	>20000[2]	>20000[2]	11000[2]	>20000[2]	6300[2]	6200 ±220[2]	2800[2]
Methyl β-D-Glc	10000[1]	10000[1]	NB	11000[1]	ND	ND	NB	NB	NB	NB
Man	>10000[14]	ND	>10000[14]	ND	ND	>10000[14]	ND	>10000[14]	ND	>10000[14]
Lactose (Galβ1-4Glc)	104[3]	ND	231[3]	1900[7]	1300[11]	ND	331.1[13]	91[3] 47.4 ±1.1[13]	ND	ND
Galβ1-4GlcNAc	ND	ND	1.8[15]	21800[7]	ND	ND	ND	420[15](SPR) 9.7[15](FA)	ND	ND
Galβ1-4Glcβ1-OMe	187[8] 190[12]	ND	160[8] 220[12]	1200[12]	540[12]	110[8] 91[12]	ND	109[4] 62[8] 52[12]	ND	23[8]
Galβ1-4GlcNAcβ1-OMe	65[8]	ND	59[8]	ND	ND	550[8]	ND	1000[8]	ND	490[8]
TDG (Thiodigalactoside)	24±11[5,8]	340±19[6]	49±11[5,8]	1450±260[5]	410±21[5]	160±18[6,8]	ND	61±17[5]	42±1.1[6]	38±8[6,8]

ND = not determined or not reported

References for Table S2:

1) M. Mahanti, K. B. Pal, A. P. Sundin, H. Leffler and U. J. Nilsson, *ACS Med. Chem. Lett.,* 2020, **11**, 34-39. (DOI: 10.1021/acsmedchemlett.9b00396)

2) F. R. Zetterberg, K. Peterson, R. E. Johnsson, T. Brimert, M. H. Kansson, D. T. Logan, H. Leffler and U. J. Nilsson, *ChemMedChem.,* 2018, **13**, 133-137. (doi: 10.1002/cmdc.201700744)

3) S. V. Klaveren, J. Dernovsek, Z. Jakopin, M. Anderluh, H. Leffler, U. J. Nilsson and T. Tomasic, *RSC Adv*., 2022, **12**, 18973-18984. (doi: 10.1039/d2ra03163a)

4) S. Carlsson, C. T. Oberg, M. C. Carlsson, A. Sundin, U. J. Niosson, D. Smith, R. D. Cummings, J. Almkvist, A. Karlsson and H. Leffler, *Glycobiology*, 2007, **17**, 663-676. (doi:10.1093/glycob/cwm026)

5) K. Bum-Erdene, P. M. Collins, M. W. Hugo, S. S. Tarighat, F. Fei, C. Kishor, H. Leffler, U. J. Nilsson, J. Groffen, I. D. Grice, N. Heisterkamp and H. Blanchard, *J. Med. Chem.,* 2022, **65**, 5975-5989. (https://doi.org/10.1021/acs.jmedchem.1c01296)

6) T. Delaine, P. Collins, A. MacKinnon, G. Sharma, J. Stegmayr, V. K. Rajput, S. Mandal, I. Cumpstey, A. Larumbe, B. A. Salameh, B. K. Knutsson, H. Hattum, M. V. Scherpenzeel, R. Pieters, T. Sethi, H. Schambye,

S. Oredsson, H. Leffler, H. Blanchard and U. J. Nilsson, *ChemBioChem.*, 2016, **17**, 1759-1770. (doi: 10.1002/cbic.201600285)

7) K. B. Erdene, H. Leffler, U. J. Nilsson and H. Blanchard, *FEBS Journal*., 2015, **282**, 3348-3367. (doi:10.1111/febs.13348)

8) I. Cumpstey, E. Salomonsson, A. Sundin, H. Leffler and U. J. Nilsson, *Chem. Eur. J*., 2008, **14**, 4233-4245. (doi: 10.1002/chem.200701932)

9) A. Dahlqvist, S. Mandal, K. Peterson, M. Hakansson, D. T. Logan, F. R. Zetterberg, H. Leffler and U. J. Nilsson, *Molecules,* 2019, **24,** 4554. (doi:10.3390/molecules24244554)

10) K. B. Pal, M. Mahanti, H. Leffler and U. J. Nilsson, *Int. J. Mol. Sci.*, 2019, **20**, 3786. (doi:10.3390/ijms20153786)

11) K. B. Erdene, H. Leffler, U. J. Nilsson and H. Blanchard, *Sci. Rep*., 2016, **6**, 20289 (doi: 10.1038/srep20289)

12) J. Tejler, E. Tulberg, T. Frejd, H. Leffler and U. J. Nilsson, *Carbohydr. Res*., 2006, **341**, 1353–1362. (doi:10.1016/j.carres.2006.04.028)

13) M. H. Bohari, X. Yu, C. Kishor, B. Patel, R. M. Go, H. A. E. Seyedi, Y. Vinik, I. D. Grice, Y. Zick and H. Blanchard, *ChemMedChem*., 2018, **13**, 1664 –1672. (doi:10.1002/cmdc.201800224)

14) J. Tejler, F. Skogman, H. Leffler and U. J. Nilsson, *Carbohydr. Res*., 2007, **342**, 1869–1875. (doi:10.1016/j.carres.2007.03.012)

15) M. H. Bohari, X. Yu, Y. Zick and H. Blanchard, *Sci. Rep*., 2016, **6**, 39556. (doi: 10.1038/srep39556)

Energy surface plots for various disaccharides: Figures S6-S9:

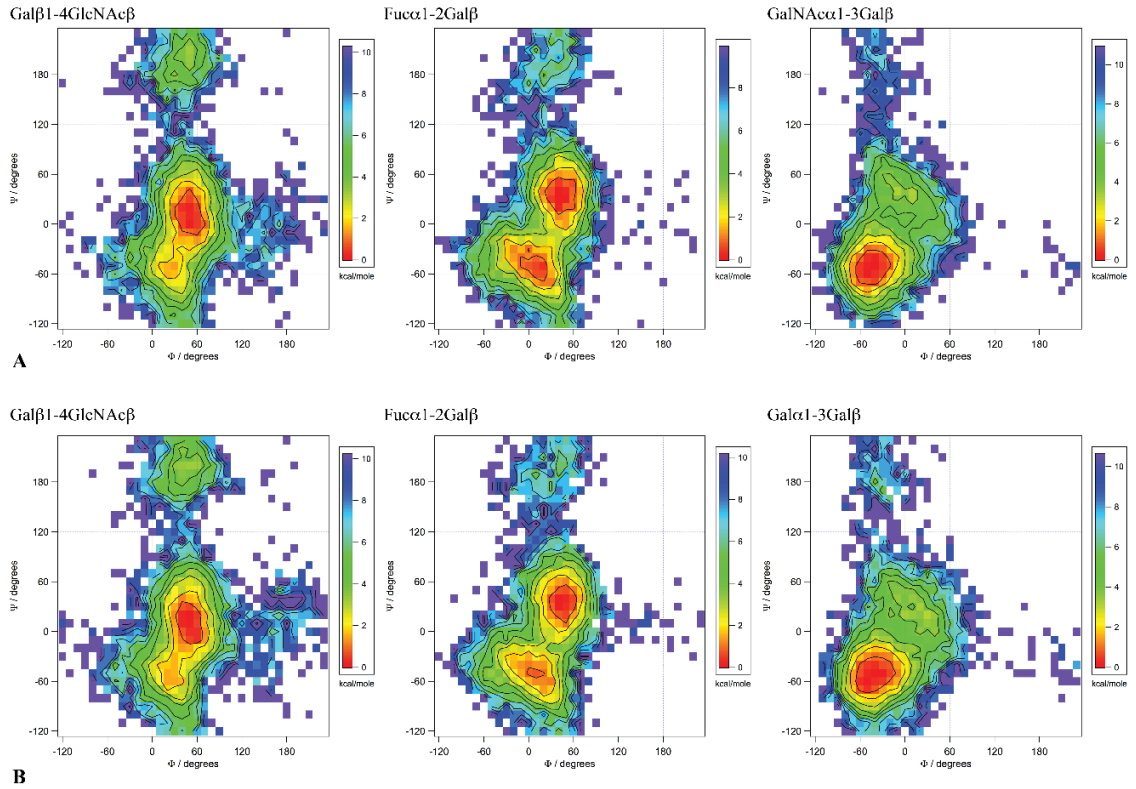

Fig. S9: This figure was reproduced from S. Kumar, M. Frank, R. Schwartz-Albiez, PLoS ONE, 2013, 8(3):e59761 (doi.org/10.1371/journal.pone.0059761). The conformational space of glycosidic linkages in blood group antigens was analyzed over 10 ns of MD simulations in the gas phase. Panel A shows the space for blood group antigen A (BGA), and Panel B for blood group antigen B (BGB). The ϕ and ψ values were determined using NMR definitions where ϕ is defined by H1-C1-O1-C_x and ψ as C1-O1-C_x-H_x .

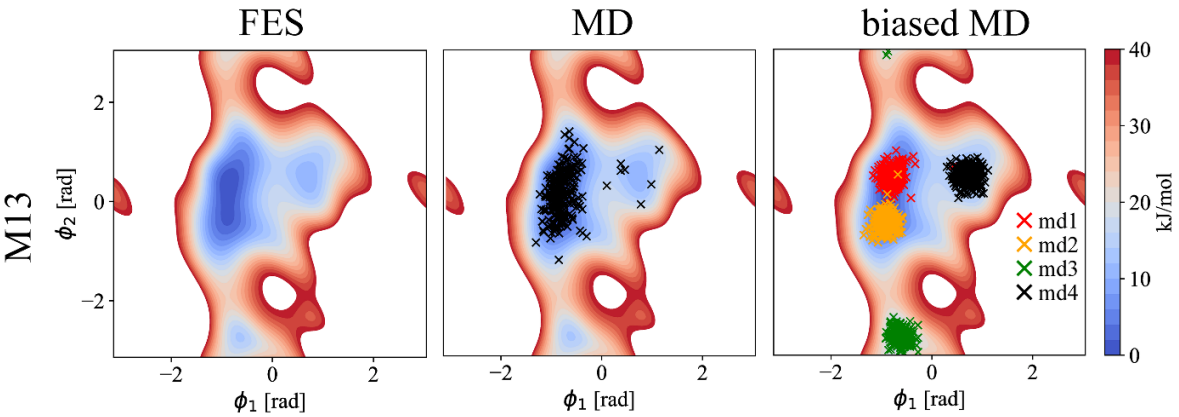

Fig. S10: This figure was reproduced from V. Palivec, C. Johannessen, J. Kaminsky, H. Martinez-Seara, PLoS Computational Biol, 2022, 18(1):e1009678 (doi.org/10.1371/journal.pcbi.1009678). Left: Calculated free energy surface of Manα(1-3)ManαOMe. Middle: Calculated FES, together with 250 sampled structures from unbiased 500 ns MD simulations. Right: Calculated FES, together with 250 sampled structures for various biased 200 ns MD simulations. White regions represent area with the free energy >40 kJ/mol. $\phi_1$is defined by atoms H1-C1-O1-C3'; ϕ_2 is defined by C1-O1-C3'-H3'. 1Rad = 180/π = 57.296°.

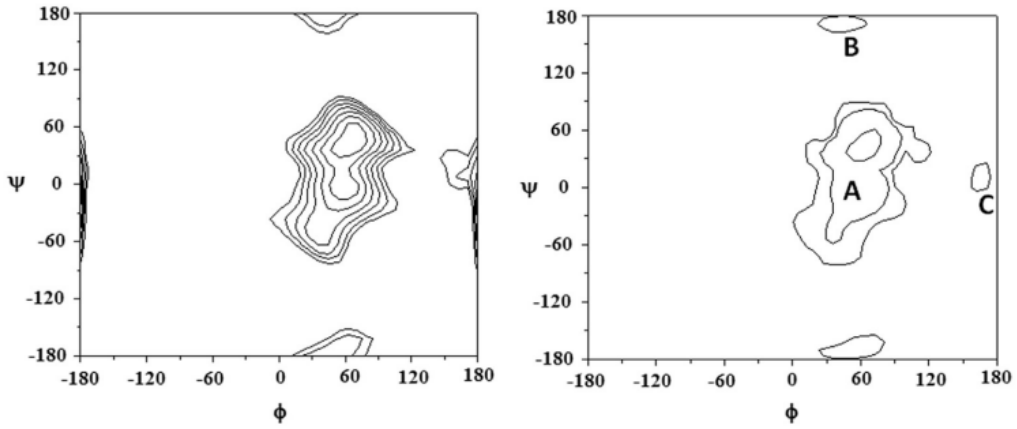

Fig. S11: This figure was reprinted from P. Vidal, J. Jiménez-Barbero, J. F. Espinosa, Carbohydrate Research 2016, 433, 36-40 (doi.org/10.1016/j.carres.2016.06.009) with permission from Elsevier. The left side shows an adiabatic map and the right side displays the population distribution of Galβ(1-3)GlcβOMe, calculated using the MM3* force field (ε = 80). Isoenergy contours are marked at 0.5 kcal/mol intervals. Probability contours indicate populations at 10%, 1%, and 0.1%. The three possible conformational families are A (syn), B (anti-ψ), and C (anti-ϕ). The ϕ and ψ torsions correspond to H1 -C1 -O1 -C3' and C1 -O1 -C3'-H3' dihedral angles respectively.

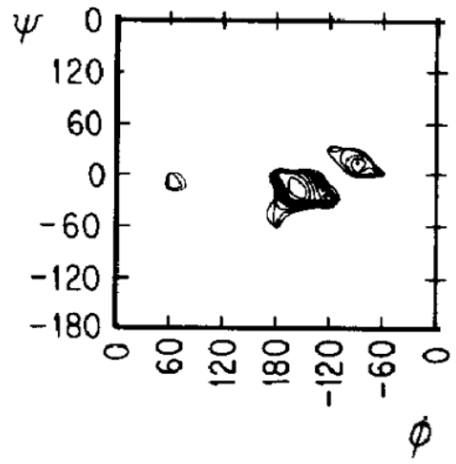

Fig. S12: This figure was reprinted from S. Sabesan, K. Bock, J. C. Paulson, Carbohydrate Research., 1991, 218, 27-54 (doi: 10.1016/0008-6215(91)84084-r) with permission from Elsevier. It shows the energy-contour map for sialoside linkage of D-NeuAc(2-3)β-D-Gal-OMe. Contour lines indicate energy differences of 1 kcal/mol, with the outermost line around each energy well representing a level 10 kcal/mol above the minimum-energy conformer. ϕ is defined by C1-C2-O3'-C3'; ψ is defined by C2-O3'-C3'-H3'.