Serianni Anthony S (Orcid ID: 0000-0001-6114-1446)

Methyl α-D-Galactopyranosyl-(1→3)-β-D-galactopyranoside and Methyl β-D-Galactopyranosyl-(1→3)-β-D-galactopyranoside: Glycosidic Linkage Conformation Determined From *MA'AT* Analysis

Reagan Meredith,[1] Yuping Zhu,[2] Mi-Kyung Yoon,[1] Timothy Tetrault,[1] Jieye Lin,[1] Wenhui Zhang,[1] Margaret McGurn,[1] Evan Cook,[1] Reed Popp,[1] Pradip Shit,[1] Ian Carmichael,[3] and Anthony S. Serianni[1]*

[1]Department of Chemistry and Biochemistry, and [3]Radiation Laboratory, University of Notre Dame, Notre Dame, IN 46556-5670 USA; [2]Discovery Chemistry, Merck Research Laboratories, Kenilworth, NJ 07033 USA

Keywords: redundant NMR spin-coupling constants; Gal-Gal disaccharides; *O*-glycosidic linkage conformation; ^{13}C-labeled disaccharides; *MA'AT* analysis

*Corresponding author: aseriann@nd.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/mrc.5424

Abstract

MA'AT analysis has been applied to two biologically-important *O*-glycosidic linkages in two disaccharides, α-D-Gal*p*-(1→3)-β-D-Gal*p*OMe (**3**) and β-D-Gal*p*-(1→3)-β-D-Gal*p*OMe (**4**). Using density functional theory (DFT) to obtain parameterized equations relating a group of trans-*O*-glycosidic NMR spin-couplings to either *phi* (ϕ') or *psi* (ψ'), and experimental $^{3}J_{COCH}$, $^{2}J_{COC}$ and $^{3}J_{COCC}$ spin-couplings measured in aqueous solution in ^{13}C-labeled isotopomers, probability distributions of ϕ' and ψ' in each linkage were determined and compared to those determined by aqueous 1-μs molecular dynamics (MD) simulation. Good agreement was found between the *MA'AT* and single-state MD conformational models of these linkages for the most part, with modest (<~15°) differences in the mean values of ϕ' and ψ', although the envelope of allowed angles (encoded in circular standard deviations or CSDs) is consistently larger for ϕ' determined from *MA'AT* analysis than from MD for both linkages. The *MA'AT* model of the αGal*p*-(1→3)-βGal*p* linkage agrees well with those determined previously using conventional NMR methods ($^{3}J_{COCH}$ values and/or ^{1}H-^{1}H NOEs), but some discrepancy was observed for the βGal*p*-(1→3)-βGal*p* linkage, which may arise from errors in the conventions used to describe the linkage torsion angles. Statistical analyses of X-ray crystal structures show ranges of ϕ' and ψ' for both linkages that include the mean angles determined from *MA'AT* analyses, although both angles adopt a wide range of values in the crystalline state, with ϕ' in βGal*p*-(1→3)-βGal*p* linkages showing greater-than-expected conformational variability.

Introduction

Galactopyranosyl residues (Gal*p*) involved in (1→3) *O*-glycosidic linkage to other sugars are important constituents of several biologically important molecules. Hyperacute rejection that occurs during discordant xenotransplantation of an organ from a pig to a primate is triggered by the binding of naturally occurring antibodies to antigens expressed on xenograft cells.[1,2] This binding destroys the vascular endothelium of the donor organ within minutes. Similar to the

αGal*p*-(1→3)-βGal*p*-(1→4)-βGlcNAc*p*-(1→3)-βGal*p*-(1→4)-βGlc*p*-protein
1

[βGluA*p*-(1→3)-βGalNAc[SO_4]*p*-(1→4)]$_n$-βGlcA*p*-(1→3)-βGal*p*-(1→3)-βGal*p*-(1→4)-βXyl*p*-O-Ser
2

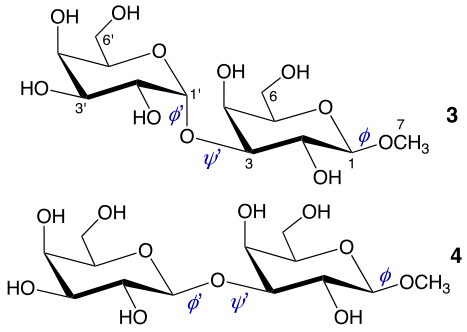
Scheme 1. An α-Gal epitope **1** recognized by human anti-Gal antibodies during xenotransplantation, and an oligosaccharide **2** in chondroitin sulfate. The linkages highlighted in blue were investigated in this work.

destruction of red blood cells in ABO-mismatched organ allotransplants, the major antigen against the human natural antibodies on porcine endothelium are oligosaccharides containing an α-D-Gal*p*-(1→3)-β-D-Gal*p* terminus (Scheme 1).[3–9] Human anti α-Gal antibodies (anti-Gal) bind trisaccharides containing α-D-Gal*p*-(1→3)-β-D-Gal*p* substructures. The α-Gal*p* epitopes are synthesized *in vivo* by α-(1→3) galactosyltransferase (α-(1→3)-GalT; EC 2.4.1.151) that is expressed in most mammals but not in humans and other Old World primates.[10–12] α-(1→3)-GalT transfers D-galactose from uridine diphosphate galactose (UDP-Gal) with retention of anomeric configuration to *N*-acetyllactosamine-containing oligosaccharides and glycolipids.[13] An unconventional food allergy, Alpha-gal syndrome (red meat allergy), is characterized by immune-mediated hypersensitivity (via IgE) to the αGal-(1→3)-Gal moiety, triggered by tick bites.[14]

β-(1→3)-Linked Gal*p* residues are found *in vivo* as constituents of human proteoglycan structure (*O*-glycans), specifically in the sequence X-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-β-D-Xyl*p*–protein near the linkage between the oligosaccharide chain and protein (Scheme 1).[15–17]

Conformational investigations of the αGal*p*-(1→3)-βGal*p* and βGal*p*-(1→3)-βGal*p* linkages have been reported using conventional NMR spectroscopic methods (e.g., ^{1}H-^{1}H NOEs

[18], ^{13}C chemical shifts[19]) and molecular modeling[20–22]. While inter-residue $^{3}J_{COCH}$ values for the αGal*p*-(1→3)-βGal*p* linkage have been determined at natural abundance to assist in conformational analysis,[23] other trans-*O*-glycoside spin-couplings (e.g., $^{2}J_{COC}$ and $^{3}J_{COCC}$)[24–27] were not measured since the compounds were not suitably ^{13}C-labeled. Without these additional carbon-based *J*-couplings, it is not possible to assign linkage conformation without relying on input from theoretical calculations. X-ray crystal structures of oligosaccharides containing both linkages have been reported,[23] often in complex with protein, but no crystal structures of disaccharides **3** and **4** are presently available, although that of the peracetate of the α-anomer of **4** has been reported.[28]



Scheme 2. Methyl α-D-galactopyranosyl-(1→3)-β-D-galactopyranoside (**3**) and methyl β-D-galactopyranosyl-(1→3)-β-D-galactopyranoside (**4**). Atom numbering and linkage torsion angles ϕ, ϕ' and ψ' are shown.

In this report, α-D-Gal*p*-(1→3)-β-D-Gal*p*OMe (**3**) and β-D-Gal*p*-(1→3)-β-D-Gal*p*OMe (**4**) (Scheme 2) were prepared chemo-enzymatically with site-specific ^{13}C-labeling and investigated by *MA'AT* analysis.[29,30] New equations for redundant trans-glycoside spin-coupling constants sensitive to the *phi* (ϕ') and *psi* (ψ') torsion angles across the internal *O*-glycosidic linkages of **3** and **4** were parameterized using density functional theory (DFT), and *MA'AT* models of ϕ' and ψ' in both linkages were compared to those obtained by aqueous molecular dynamics (MD) simulation. The results are discussed in the context of prior conformational assignments of these linkages.

Experimental

A. Preparation of Unlabeled and ^{13}C-Labeled Disaccharides. Methods used to prepare unlabeled and ^{13}C-labeled disaccharides **3** and **4** are described in the Supporting Information (Schemes S1 and S2).

B. NMR Spectroscopy. 1D ^{1}H and ^{13}C NMR spectra were obtained in $^{2}H_2O$ at ~298 K on a 600-MHz NMR spectrometer operating at 599.89 MHz for ^{1}H and 150.85 MHz for ^{13}C. 1D ^{1}H

spectra were obtained with spectral windows of 15 ppm and 26,920 data points, and FIDs were zero-filled to 65,536 points to give final digital resolutions of 0.14 Hz/point. 1D $^{13}C\{^{1}H\}$ spectra were obtained with spectral windows of 250 ppm and 75,472 data points, and FIDs were zero-filled to 131,072 points to give final digital resolutions of 0.28 Hz/point. ^{1}H and ^{13}C resonances for **3** were assigned using a combination of 1D ^{1}H and ^{13}C spectra, and 2D ^{1}H-^{1}H DQF-COSY[31,32], ^{1}H-^{1}H TOCSY,[33,34] and ^{1}H-^{13}C gHSQC[35,36] spectra recorded on a 600-MHz NMR spectrometer. ^{1}H and ^{13}C chemical shifts for **3** and **3**$^{1',3}$ (throughout this manuscript, the superscripted number on a compound number identifies an atom that was labeled with ^{13}C) are reported in Tables S1 and S2, respectively (see Supporting Information). Intra-residue J_{HH}, J_{CH} and J_{CC} values in **3** and **3**$^{1',3}$ are given in Tables S3–S5 (see Supporting Information). Representative 1D and 2D NMR spectra of **3** and **3**$^{1',3}$ are available in the Supporting Information (Figures S3–S13).

For disaccharides **4**, **4**$^{1'}$ and **4**$^{2'}$, 1D ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were obtained on a 600-MHz spectrometer equipped with a 3-mm Nalorac microprobe. ^{1}H and ^{13}C chemical shifts, and J_{HH}, J_{CH} and J_{CC} values were obtained from direct analyses of 1D spectra (Table S6, Supporting Information). In **4**, $^{3}J_{C3,H1'}$ was not measurable directly from 1D ^{1}H NMR spectra since **4** was not ^{13}C-labeled at C3. In this case, $^{3}J_{C3,H1'}$ was measured at natural abundance.[37]

Calculations

A. Geometry Optimizations. Density functional theory (DFT) calculations were conducted on model structures **3**^{c} and **4**^{c} (Scheme 2; *in silico* DFT structures are identified by a "c" superscript to distinguish them from compounds **3** and **4**) within *Gaussian*16[38] using the B3LYP functional[39,40] and the 6-31G* basis set.[41] In all geometric optimizations, the effects of solvent water were treated using the Self-Consistent Reaction Field (SCRF)[42] and the Integral Equation Formalism (polarizable continuum) model (IEFPCM).[43] For calculations on **3**^{c} and **4**^{c}, the *phi* (ϕ') and *psi* (ψ') *O*-glycosidic torsion angles, defined as C2'–C1'–O1'–C3 and C1'–O1'–C3–C2, respectively, were each rotated in 15^{o} increments through 360^{o}, giving a 24 x 24 matrix or 576

optimized structures (Schemes S3 and S4, Supporting Information). The remaining ten exocyclic torsion angles were set initially at 180° and allowed to optimize during the calculations (Schemes S3 and S4, Supporting Information). As indicated in Schemes S3 and S4, only one combination of exocyclic C–O and C–C torsion angles in **3**^{c} and **4**^{c} were treated in this work. This limitation, however, is not expected to affect the parameterization of trans-glycosidic *J*-coupling equations to any appreciable extent based on prior work. Furthermore, given the nature of *MA'AT* analysis, the DFT calculations do not need to provide accurate conformational energies (energies are not required as inputs) or recapitulate hydrogen bonding that may occur in solution. In all calculations, only structures containing unstrained $^{4}C_{1}$ ring conformations of the Gal*p* residues were used in the *J*-coupling calculations. Cremer-Pople parameters were calculated for each galactopyranosyl ring in all optimized structures to identify those structures containing unstrained chair forms (a θ value of 35° was used as the cutoff).

B. Calculations of NMR Spin-Coupling Constants. J_{CH} and J_{CC} *s*pin-coupling constants that depend on ϕ' or ψ' were calculated in geometry-optimized structures of **3**^{c} and **4**^{c} by DFT using the B3LYP functional[39,40] in *Gaussian*16. The Fermi contact,[44–46] diamagnetic and paramagnetic spin-orbit, and spin-dipole terms[44] were recovered using a tailored basis set, [5s2p1d|3s1p],[47,48] and raw (unscaled) calculated spin-couplings are reported and are accurate to within ± 0.2–0.3 Hz based on prior work.[47,48] The Self-Consistent Reaction Field (SCRF)[42] and the Integral Equation Formalism (polarizable continuum) model (IEFPCM)[43] were used to treat the effects of solvent water during the *J*-coupling calculations.

C. Spin-Coupling Equation Parameterization. Equations relating DFT-calculated J_{CH} and J_{CC} values to ϕ' or ψ' in **3**^{c} and **4**^{c} were parameterized using the software package R. Equations were parameterized using the full set of conformers (full equations), or using a subpopulation of conformers that was selected by applying a 10 kcal/mol energy cut-off to the full dataset to remove a limited number of highly strained conformers (trimmed equations).[29] A secondary constraint was also applied to the trimmed equations when necessary to remove DFT-optimized structures containing distorted aldohexopyranosyl rings; Cremer-Pople puckering parameters were

calculated from DFT-generated Cartesian coordinates and a θ value of 35° was used as the cut-off.[29]

For *MA'AT* analyses of ϕ' in **3** and **4**, an additional set of equations was parameterized in which the secondary effects of ψ' were truncated by including only those values of ψ' in the population of conformers determined from *MA'AT* analysis (trimmed and *psi*-constrained equations). All of the parameterized equations are given in the Supporting Information (eqns. [S1]–[S30]). The goodness-of-fit of each equation is reported as a root-mean-squared deviation (RMSD). Equation parameterization was further evaluated using the Akaike Information Criterion (AIC),[50] resulting in truncated forms when justified.

D. Conformational Models of ϕ' and ψ' in **3**^{c} and **4**^{c} *From MA'AT Analysis*. NMR *J*-couplings and DFT-parameterized equations were combined with Fredholm theory and circular statistics to obtain conformational models of ϕ' *and* ψ' *in* **3**^{c} and **4**^{c} using the in-house statistical software, *MA'AT*.[30] The current version of *MA'AT* can be accessed at: https://rmeredit.shinyapps.io/maat24/. A User's Manual is available on the software's webpage.

E. Aqueous Molecular Dynamics Simulations of **3**^{c} *and* **4**^{c}. Initial structures of **3**^{c} and **4**^{c} were built using the Carbohydrate Builder module available at the GLYCAM website http://www.glycam.org.[51] The GLYCAM06[52] (version j) force field was employed in all simulations. Structures **3**^{c} and **4**^{c} were solvated with TIP3P[53] water using a 12Å buffer in a cubic box, using the LEaP module in the AMBER14 software package.[54] Energy minimizations for solvated **3**^{c} and **4**^{c} were performed separately under constant volume (500 steps steepest descent, followed by 24500 steps of conjugate-gradient minimization). Each system was subsequently heated to 300 K over a period of 50 ps, followed by equilibration at 300 K for a further 0.5 ns using the nPT condition, with the Berendsen thermostat[55] for temperature control. All covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm,[56] allowing a simulation time step of 2 fs throughout the simulation. After equilibration, production simulations were carried out with the GPU implementation[57] of the PMEMD.MPI module, and trajectory frames were collected every 1 ps for a total simulation time of 1 μs. One to four non-

bonded interactions were not scaled[58] and a non-bonded cut-off of 8 Å was applied to van der Waals interactions, with long-range electrostatics treated with the particle mesh Ewald approximation. Output from each MD simulation was imported into *Prism*[59] for visualization.

Results and Discussion

*A. Preparation of Unlabeled and ^{13}C-Labeled **3** and **4***. Enzyme-catalyzed transglycosylation[60–64] was used to prepare unlabeled and ^{13}C-labeled disaccharides **3** and **4**. Glycosidases are commonly used to hydrolyze *O*- and *N*-glycosidic linkages, but they can also be used to catalyze the formation of *O*-glycosidic linkages under certain reaction conditions. These conditions involve the use of an activated donor that is hydrolyzed by the enzyme, producing a transient covalently-modified enzyme bearing the saccharide donor. Transfer of the enzyme-bound donor occurs subsequently to a suitable acceptor. The reaction occurs with retention of configuration, and linkage regioselectivity can often be controlled to achieve highly specific synthesis of the desired linkages.[60–63]

Enzyme-catalyzed transglycosylation to prepare **3** (Scheme S1, Supporting Information) gave α-(1→6)-linked **J** and α-(1→3)-linked **3**, which were separated by chromatography on a Sephadex G-10 column followed by chromatography on a Dowex-1 x 8 (OH^-) column.[65] Preparation of the glycosyl acceptor, methyl β-D-galactopyranoside (**I**), was first attempted by Fischer glycosidation in methanol,[66–68] but gave the α-pyranoside as the heavily favored product. An alternate pathway through glycosyl bromide **G** was then adopted that gave **I** in ~60% yield from D-galactose **E**. Considering the reaction yield to prepare donor **D**, the reaction regiochemistry that often produces more than one linkage, and the need for excess acceptor (which may preclude isotopic labeling of this component), enzyme-catalyzed transglycosylation will not be a generally useful method to prepare ^{13}C-labeled disaccharides, although it has the advantage of simplicity.

Transglycosylation was also used to prepare **4** (Scheme S2, Supporting Information). In this case, the anomeric configurations of acceptor **L** and donor **K** are the same (βGal). While **K**

is a more reactive donor than **L**, **L** also serves as a donor due to its high concentration relative to **K**. In the present application, the same two disaccharides **4** and **M** form, *but each is produced as a mixture of unlabeled and ^{13}C-labeled compounds*. This complication impedes the measurement of *J*-couplings in the products due to overlapping signals in ^{1}H and ^{13}C NMR spectra arising from both labeled and unlabeled products. *From a practical standpoint, the use of transglycosylation to prepare isotopically labeled disaccharides containing identical residues in which only one residue is to be labeled should be confined to reactions where the anomeric configurations of the donor and acceptor differ.*

*B. NMR Signal Assignments and J_{HH}, J_{CH} and J_{CC} Values in **3** and **4***. ^{1}H NMR signal assignments (Figure S3 and S4, Tables S1 and S6, Supporting Information) were based on NMR analyses of unlabeled and ^{13}C isotopomers of **3** and **4**. For example, the H1′ signal (5.145 ppm, $^{3}J_{H1',H2'}$ = 3.9 Hz) in **3** was assigned from observation of the large $^{1}J_{C1',H1'}$ (170.0 Hz) in **3**$^{1',3}$ and the smaller $^{3}J_{H1',H2'}$ expected in αGal anomers .[68,69] The H1 signal (4.375 ppm, $^{3}J_{H1,H2}$ = 8.0 Hz) was assigned based on a $^{3}J_{H1,H2}$ value consistent with βGal anomers.[68,69] Signal assignments of the remaining hydrogens in both residues of **3** were made based on analyses of J_{HH} values extracted from the 1D ^{1}H NMR spectrum (Table S3, Figures S3 and S4, Supporting Information) and analyses of 2D ^{1}H-^{1}H DFQ-COSY and ^{1}H-^{1}H TOCSY spectra (Figures S7–S9, Supporting Information).

Table 1. Trans-Glycosidic Spin-Coupling Constants[a] in Disaccharides **3** and **4.**

J-Coupling	αGal-(1→3)-βGalOCH$_3$ (**3**)	βGal-(1→3)-βGalOCH$_3$ (**4**)
phi (ϕ')		
$^{3}J_{C3,H1'}$	2.8 (2.9–3.1)[b]	3.9
$^{2}J_{C1',C3}$	–1.8	–2.0
$^{3}J_{C2',C3}$	2.6	3.1
psi (ψ')		
$^{3}J_{C1',H3}$	3.3 (3.3–3.4)	4.1
$^{3}J_{C1',C2}$	2.3	2.6
$^{3}J_{C1',C4}$	~0	0.5

[a]In Hz ± 0.1 Hz, in $^{2}H_2O$ at ~25 °C. [b]Values in parentheses those reported previously.[76] See Scheme 2 for the definitions of torsion angles ϕ' and ψ'.

^{13}C Signal assignments in **3** (Figures S5 and S6, Table S2, Supporting Information) were made from the 2D ^{1}H-^{13}C gHSQC spectrum (Figure S10, Supporting Information) and from the magnitudes of intra-residue J_{CC} values observed in the 1D ^{13}C{^{1}H} spectrum of **3**$^{1',3}$ (Figure S6, Table S5, Supporting Information).

Methods similar to those used to assign ^{1}H and ^{13}C signals and measure J_{HH}, J_{CH} and J_{CC} values in NMR spectra of **3** and **3**$^{1',3}$ were applied to **4**, **4**$^{1'}$ and **4**$^{2'}$. These results are summarized in Table S6 (Supporting Information).

*C. Trans-Glycosidic Spin-Couplings in **3** and **4***. The trans-glycosidic $^{3}J_{C3,H1'}$ in **3** was obtained from analysis of the H1′ signal in the ^{1}H NMR spectrum of **3**$^{1',3}$, assisted by selective decoupling of H2′ (Figure S11, Supporting Information). $^{3}J_{C1',H3}$ was obtained from analysis of the H3 signal, which was split by the large $^{1}J_{C3,H3}$, producing two doublets which were further split by spin-coupling to H2, H4 and C1′ to give a pair of multiplets. $^{3}J_{H2,H3}$ and $^{3}J_{H3,H4}$ were determined from analyses of the H2 and H4 signals, respectively, allowing an initial estimate of $^{3}J_{C1',H3}$ (~3 Hz) from one of these multiplets. This value was refined through selective decoupling of H4, giving a more precise $^{3}J_{C1',H3}$ value of 3.3 Hz (Figure S12, Supporting Information). $^{2}J_{C1',C3}$ was obtained from the C1' and C3 doublets in the ^{13}C{^{1}H} NMR spectrum of **3**$^{1',3}$ (Figure S6, Supporting Information). $^{3}J_{C1',C2}$, $^{3}J_{C2',C3}$ and $^{3}J_{C1',C4}$ were measured from the splitting of the C2, C3 and C4 signals, respectively. Trans-glycosidic J_{CH} and J_{CC} values in **3** are summarized in Table 1.

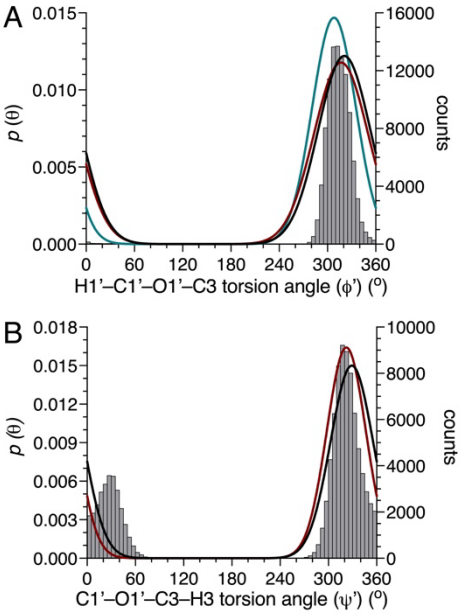


Figure 1. *MA'AT* and MD analyses of ϕ' (A) and ψ' (B) in αGal-(1→3)-βGalOCH$_3$ (**3**). Black, full dataset. Red, trimmed dataset. Green, trimmed and constrained dataset. The hatched grey areas in (A) and (B) are models determined from a 1-μs aqueous MD simulation. See text for details.

Methods similar to those used to measure trans-glycosidic *J*-couplings in **3** and **3**$^{1',3}$ were applied to **4**, giving values summarized in Table 1. Since $^{3}J_{C3,H1'}$ could not be measured directly from the ^{1}H NMR spectra of either **4**$^{1'}$ or **4**$^{2'}$, it was measured at natural abundance in unlabeled **4**.[37]

*D. MA'AT and MD Analyses of ϕ' and ψ' in **3** and **4**.* Trans-glycosidic *J*-couplings (Table 1) and parameterized eqns [S1]–[S30] (see Supporting Information) were used as inputs in *MA'AT* analyses of ϕ' and ψ' in **3** and **4**. The results are shown in Figures 1 and 2, and the statistics from these analyses are summarized in Table 2.

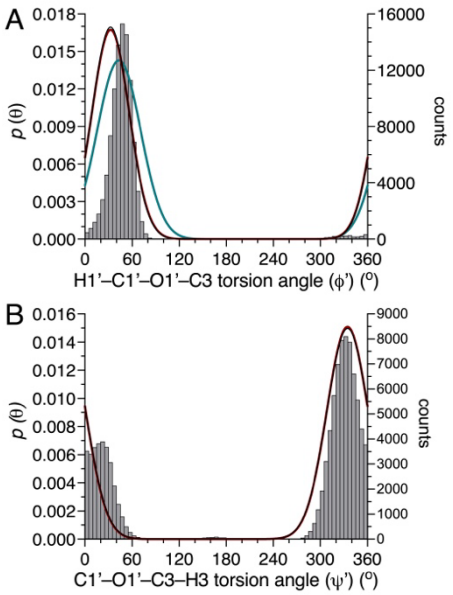


Figure 2. *MA'AT* and MD analyses of ϕ' (A) and ψ' (B) in βGal-(1→3)-βGalOCH$_3$ (**4**). Black, full dataset. Red, trimmed dataset. Green, trimmed and constrained dataset. The hatched grey areas in (A) and (B) are models determined from a 1-μs aqueous MD simulation. See text for details.

Depending on the parameterized equations used in the analysis, the mean value of ϕ' in **3** ranged from –39.2 – –51.9°, with CSDs ranging from 27.1–33.7° and RMSDs ranging from 0.06–0.46 Hz. The smallest RMSD was observed when the ψ'-constrained equations were used. This "best fit" to the experimental data gave a mean ϕ' of –51.9° and a CSD of 27.1°, with the former in good agreement with that determined from MD simulation (–45.4°) when the data were fit to a single-state model and –46.0° when a two-state model was used (Table 3). CSDs obtained from MD (13–15°) were significantly smaller than CSDs determined from *MA'AT* analysis, an observation consistent with prior *MA'AT* studies of ϕ in *O*-glycosidic linkages.[29] The cause of this discrepancy is not known, but may be due to flaws in the *MA'AT* treatment, in the GLYCAM06 force field, or both. The RMSD

Table 2. Statistics From *MA'AT* Analyses of ϕ' and ψ' in αGal-(1→3)-βGalOCH$_3$ (**3**) and βGal-(1→3)-βGalOCH$_3$ (**4**).

αGal-(1→3)-βGalOCH$_3$ (3)						
Fit constraints	*phi* (ϕ')			*psi* (ψ')		
	mean ± SE[a] (°)	CSD[b] ± SE (°)	RMSD[c] (Hz)	mean ± SE (°)	CSD ± SE (°)	RMSD (Hz)
full	–39.2 (± 14.1)	32.5 (± 16.6)	0.46	–31.0 (± 9.8)	26.4 (± 14.1)	0.40
trimmed	–43.2 (± 14.5)	33.7 (± 17.2)	0.33	–37.8 (± 8.4)	24.2 (± 13.7)	0.60
psi-constrained	–51.9 (± 9.6)	27.1 (± 20.1)	0.06			

βGal-(1→3)-βGalOCH$_3$ (4)						
Fit constraints	*phi* (ϕ')			*psi* (ψ')		
	mean ± SE (°)	CSD ± SE (°)	RMSD (Hz)	mean ± SE (°)	CSD ± SE (°)	RMSD (Hz)
full	32.4 (± 10.4)	23.4 (± 13.4)	0.16	–25.3 (± 9.5)	26.4 (± 12.0)	0.05
trimmed	32.5 (± 10.7)	23.7 (± 13.4)	0.15	–25.4 (± 9.6)	26.2 (± 11.3)	0.07
psi-constrained	43.1 (± 9.2)	27.8 (± 15.3)	0.10			

[a]SE = standard error. [b]CSD = circular standard deviation. [c]RMSD = root mean squared deviation. See text for descriptions of the full, trimmed and *psi*-constrained fitting constraints.

calculated from the MD data (0.45 Hz) was considerably greater than that determined from the best *MA'AT* fit (0.06 Hz), indicating that the *MA'AT* model may better represent the torsional behavior of ϕ' in solution.

MA'AT analysis of ψ' in **3** gave mean values of either –31.0º or –37.8º depending on the equations used in the calculations, and CSDs of 24–26º. RMSDs ranged from 0.40–0.60 Hz, somewhat larger than observed for ϕ' but still indicative of a good fit to the experimental *J*-couplings. MD data fit to a single-state model gave a mean value of –20.5º, a CSD of 33.6º and a RMSD of 0.79 Hz. When the MD data were fit to a two-state model, which is more consistent with the MD histogram (Figure 1B), mean values of –37.4º (72%) and 27.6º (28%) were obtained. Given the limited number of redundant *J*-couplings used in *MA'AT* analysis of ψ', it is not possible to test using *MA'AT* analysis whether the two-state model indicated by MD provides a better fit to the experimental *J*-couplings than the single-state model. However, RMSDs calculated from the one- and two-state fits of the MD data are the same (0.79 Hz) and slightly larger than that obtained from a single-state *MA'AT* analysis (0.40–0.60 Hz), indicating that the single-state *MA'AT* model is a valid representation of the solution behavior of ψ'.

Table 3. Statistics From Molecular Dynamics Simulations of αGal-(1→3)-βGalOCH$_3$ (**3**^{c}) and βGal-(1→3)-βGalOCH$_3$ (**4**^{c}) For One- and Two-State Models.

One-State Model						
structure	*phi* (ϕ)			*psi* (ψ)		
	mean (º)	CSD[a] (º)	RMSD[b] (Hz)	mean (º)	CSD (º)	RMSD (Hz)
αGal-(1→3)-βGalOCH$_3$ (**3**^{c})	–45.4	14.5	0.45	–20.5	33.6	0.79
βGal-(1→3)-βGalOCH$_3$ (**4**^{c})	44.3	15.1	0.29	–14.1	29.8	0.37

Two-State Model							
structure	mean 1 (º)	CSD 1 (º)	peak 1 %	mean 2 (º)	CSD 2 (º)	peak 2 %	RMSD (Hz)
	phi (ϕ)						
αGal-(1→3)-βGalOCH$_3$ (**3**^{c})	–46.0	12.9	98.3	0.9	31.7	1.7	0.45
βGal-(1→3)-βGalOCH$_3$ (**4**^{c})	45.5	15.6	97.9	–34.8	15.2	2.1	0.29
	psi (ψ)						
αGal-(1→3)-βGalOCH$_3$ (**3**^{c})	–37.4	15.9	72.0	27.6	15.0	28.0	0.79
βGal-(1→3)-βGalOCH$_3$ (**4**^{c})	–29.5	15.9	69.2	22.1	14.1	30.8	0.37

[a]CSD = circular standard deviation. [b]RMSD = root mean squared deviation.

MA'AT analysis of ϕ' in **4** gave a good fit to the experimental *J*-couplings when a single-state model was assumed, giving RMSDs of 0.10–0.16 Hz depending on the equations used in the analysis. Mean values of ϕ' ranged from 32.4–43.1º, with the largest value obtained from the

psi-constrained equations, while CSDs ranged from 23.4–27.8º. MD simulation gave a mean value of 44.3º for ϕ', in good agreement with *MA'AT* analysis when the *psi*-constrained equations were applied. As found for **3**, the CSD from MD (15.1º) was considerably smaller than those obtained from *MA'AT* analysis. When the MD data were fit with a two-state model, the mean value of ϕ' in the dominant conformer (45.5º; 98%) and its associated CSD (15.6º) were essentially the same as corresponding values obtained from a single-state fit of the MD data.

MA'AT analyses of ψ' in **4** gave mean values of –25º and CSDs of 26º, and neither value is much affected by the equations used in the analysis. RMSDs were small (0.05–0.07 Hz), indicating good fits of the calculated model to the experimental *J*-couplings. Single-state fits of the MD data for **4** gave a mean value of ψ' of –14.1º , a CSD of 29.8º, and a RMSD (0.37 Hz). When a two-state model was used to fit the MD data, a mean value of ψ' of –29.5º was obtained for the dominant conformation (69%), with the minor conformer (31%) having a mean of 22.1º. CSDs of 14–16º for both conformers are smaller than that found from *MA'AT* analysis. The RMSD for the two-state fit of the MD data (0.37 Hz) was identical to that found for the single-state fit. As found for **3**, the population distribution of ψ' in **4** obtained from *MA'AT* analysis overlaps the two populations observed by MD, with the *MA'AT* mean lying between the two means determined by MD. The envelope of *MA'AT*-determined ψ' values includes nearly all angles indicated by MD, but only one broad envelope is needed to fit the experimental data well, yielding a smaller RMSD but larger CSD than MD. Based on the

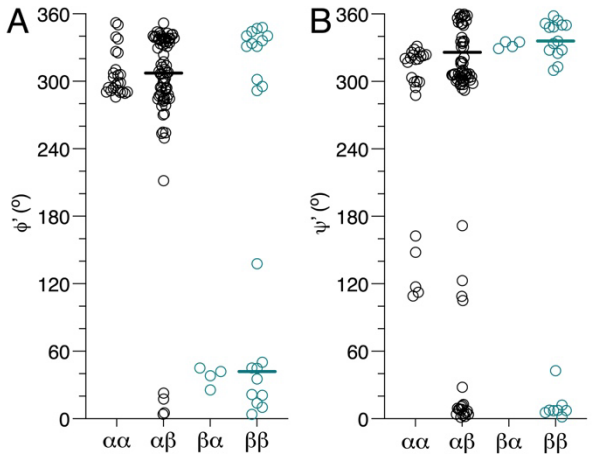


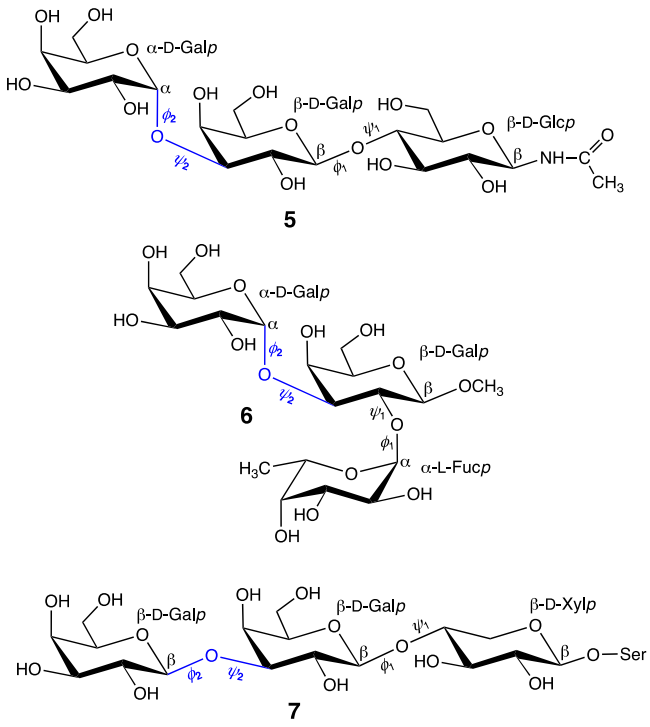
Figure 3. Distributions of ϕ' (A) and ψ' (B) in X-ray crystal structures containing αGal-(1→3)-αGal (αα), αGal-(1→3)-βGal (αβ), βGal-(1→3)-αGal (βα), and βGal-(1→3)-βGal (ββ) *O*-glycosidic linkages. Black, α-linkages. Green, β-linkages. Total structures in each group were 24, 73, 4 and 22, respectively. Horizontal bars indicate the mean values of ϕ' and ψ' determined from *MA'AT* analysis (Table 2). ϕ' = H1'–C1'–O1'–C3. ψ' = C1'–O1'–C3–H3.

available experimental *J*-couplings, a single-state model of ψ' may pertain to **4** in solution, as found for **3**.

E. Linkage Conformations in X-ray Crystal Structures. X-ray crystal structures containing αGal-(1→3)-αGal, αGal-(1→3)-βGal, βGal-(1→3)-αGal and βGal-(1→3)-βGal *O*-glycosidic linkages were extracted from the PDB to determine the preferred values of ϕ' and ψ' in the crystalline state. Only structures with resolutions of 1.2–2.9 Å were considered, and structures bound to protein or other macromolecular receptors were excluded from the analysis. These constraints produced 24, 73, 4 and 22 structures containing αGal-(1→3)-αGal, αGal-(1→3)-βGal, βGal-(1→3)-αGal and βGal-(1→3)-βGal linkages, respectively (Figure 3). The number of βGal-(1→3)-αGal structures was too low to allow a meaningful analysis. For the αGal-(1→3)-αGal and αGal-(1→3)-βGal linkages, ϕ' ranges from 320 ± 40º, which includes the mean value determined from *MA'AT* analysis (–51.9º or 308.1º). The ψ' values show significant variability, but most are 330 ± 40º, which includes the mean value determined from *MA'AT* analysis (~–34º or 326º). While there are fewer structures containing βGal-(1→3)-βGal linkages, the crystallographic data indicate that ϕ' adopts a wide range of values (350 ± 60º), indicating that it can deform relatively easily in the crystalline lattice despite stereoelectronic constraints.[70–75] *MA'AT* analysis gave a mean value of 43.1º, which lies at the edge of this range. These observations indicate that solvation factors could play a more significant role in dictating ϕ' in **4** than in **3**. On the other hand, ψ' in structures containing βGal-(1→3)-βGal linkages ranges from 350 ± 40º (similar to the range observed for αGal-(1→3)-βGal linkages) and includes the mean value determined from *MA'AT* analysis (–25.4º or 334.6º).

F. Prior Conformational Analyses of αGal-(1→3)-βGal and βGal-(1→3)-βGal Linkages in Solution. Wang and coworkers[76] examined the αGal-(1→3)-βGal linkage in trisaccharide **5** (Scheme 5) using two trans-*O*-glycosidic *J*-couplings, ^{1}H-^{1}H NOEs and HSEA/MM calculations. $^{3}J_{H1',C3}$ and $^{3}J_{C1',H3}$ values of 3.1 Hz and 3.4 Hz, respectively, were measured, which are similar to those reported in this study (Table 1). Based mainly on computational data, preferred values of –50º and –34º were reported for ϕ' and ψ', respectively. These angles are in good agreement

with those found from *MA'AT* analysis (–51.9º and –38º, respectively; Table 2). Bundle and coworkers[23] examined the same αGal-(1→3)-βGal linkage in trisaccharide **6** and reported $^3J_{H1',C3}$ and $^3J_{C1',H3}$ values of 1.4 Hz and 1.5 Hz, respectively, which differ significantly from corresponding values in **3** and **5**. They concluded that this linkage prefers ϕ and ψ values of –63º and –62º, respectively. Computational studies of the αGal-(1→3)-βGalOCH$_3$ disaccharide gave preferred ϕ and ψ values of –49º and –32º, respectively, in good agreement with those found in this report (Table 2). The authors concluded that the addition of the αFuc-(1→2) linkage to the βGal residue affects both ϕ and ψ (context effect). Future *MA'AT* analyses of **6** may provide a more quantitative determination of this conformational change. Finally, Imberty and coworkers[77] studied the αGal-(1→3)-βGal linkage using computational methods and reported two dominant conformations characterized by ϕ/ψ values of –40º/–40º and –20º/20º, and a minor conformer with ϕ/ψ of –20º/–190º. Both dominant forms fall within the envelope of conformations determined by *MA'AT* analysis, with the –40º/–40º form having values similar to the mean values of –51.9º/–37.8º determined by *MA'AT* (Table 2). *MA'AT* analysis was not able to detect the minor –20º/–190º form, presumably due to its low abundance (the relative abundances of the three forms were not reported).



Scheme 5. Structures of Trisacharides **5–7.** The αGal-(1→3)-βGal linkages are highlighted in blue in **5** and **6**, and the βGal-(1→3)-βGal linkage is highlighted in blue in **7.**

Krishna *et al.*[78] studied trisaccharide **7** that contains a βGal-(1→3)-βGal linkage identical to that found in **4**. ^{1}H-^{1}H NOE measurements and MD/EM calculations were interpreted to indicate preferred ϕ angles (defined as O5–C1–O1–C3) ranging from –164.1— –172.0º, with a possible

minor conformer with ϕ = ~35º. The ψ angle, defined as C1–O1–C3–C2, was reported to favor values ranging from –168.9 — –173.2º (data taken from Table IV in ref. 78). These angles were converted to those used in this study by adding +120º to each, giving ϕ values of –44.1 – –52º for the major form (155º for the minor form) and –48.9 — –53.2º for ψ. This ϕ assignment differs significantly from that obtained from *MA'AT* analysis (mean value of +43.1º from *psi*-constrained equations), but the ψ assignment compares somewhat favorably with that determined from *MA'AT* analysis (~–25.4º using trimmed equations) (Table 2).

Prior NMR studies of the βGal-(1→4)-βXyl linkage in **7** by Krishna and coworkers[79,80] gave preferred ϕ and ψ values of –62.8º (O5–C1–O1–C4) and 128.3º (C1–O1–C4–C3), respectively. These values are similar qualitatively to those determined recently in the disaccharide, βGal-(1→4)-βXylOCH$_3$, by *MA'AT* analysis (–92.5 º for ϕ, 136.8º for ψ).[29] An X-ray crystallographic study of the same disaccharide gave values of –85.18º and 96.01º for ϕ and ψ, respectively.[81]

Conclusions

Conformational models of saccharides have relied heavily on input from computational methods, most often from MD simulation. Conventional assignments of conformation are typically made by comparing a group of NMR parameters (e.g., *J*-couplings, ^{1}H-^{1}H NOEs, RDCs, spin-relaxation times) to the MD model to test for internal consistency. While this approach has proven useful, it places a heavy emphasis on computational models despite the fact that the underlying force field from which they are obtained cannot be fully validated by experiment. *MA'AT* analysis provides a solution to this problem by analyzing sets of redundant NMR *J*-couplings (*J*-values that depend on the same molecular torsion angle) to provide probability distributions of molecular torsion angles similar to those determined by MD simulation. A physical picture of the behavior of the torsion angle in solution is thereby obtainable from experimental data, which heretofore has

proven difficult to achieve. In addition to providing these pictures, feedback provided by *MA'AT* analysis provides an effective means of validating MD force fields experimentally.

This investigation has applied *MA'AT* analysis to two disaccharides **3** and **4** containing two biologically-important *O*-glycosidic linkages involving αGal and βGal donors. This work is part of an ongoing longitudinal effort to test the ability of *MA'AT* analysis to model different types of conformational elements in saccharides and other biologically important molecules.[82–85] In the present work, we have shown that the method provides reliable conformational models of both linkages, with the degree of similarity between *MA'AT* and MD particularly high for ψ both in terms of mean values and CSDs. A similar comparison for ϕ showed good agreement between mean values, but different CSDs, as found in prior work.[29,82] The cause of the latter discrepancy has been investigated recently where secondary effects on the ϕ-dependent *J*-couplings were minimized by using parameterized equations that were constrained based on information obtained from *MA'AT* analysis of ψ (ψ-constrained ϕ-dependent equations). This modified equation parameterization resulted in essentially no change in the *MA'AT*-determined CSDs for ϕ (although mean values changed slightly as observed in the present work) (Table 2), eliminating secondary effects as the cause of the discrepancy.[86] Recent MD simulations using the CHARMM force field on methyl β-lactoside gave a CSD for ϕ similar to that obtained from the GLYCAM06 force field. In these MD studies, 1-μs simulations were assumed to be sufficient to capture the equilibrium distribution of ϕ in solution. We are currently examining the use of non-conventional two-bond $^2J_{CCH}$ values[87] as additional constraints of ϕ in an effort to resolve the CSD discrepancy, results of which will be reported in a future publication.

The DFT calculations performed in this study for geometry optimization made used of the B3LYP functional[39,40] and 6-31G* basis set[41]. More extended basis sets and modern functionals have been used in our work as discussed in a recent review.[88] However, use of the latter in most cases gives calculated ^2J and ^3J values and their associated Karplus-like equations that do not differ significantly from those calculated by the methods used in this work. The resulting *MA'AT* models are also virtually identical. The fact that reasonable *MA'AT* models of glycosidic linkage

torsion angles were obtained (comparable to those obtained by MD simulation) provides evidence, albeit indirect, that the DFT calculations must be giving reliable calculated *J*-values and parameterized equations, otherwise convergence during *MA'AT* analysis with the resulting low RMSDs would not occur. More extended basis sets and other functionals, however, must be used when accurate C–H bond lengths are required, since these lengths exert a large effect on DFT-calculated $^{1}J_{CH}$ and $^{1}J_{CC}$ values.[89,90]

This study extends the application of *MA'AT* analysis to (1→3) *O*-glycosidic linkages involving Gal*p* residues as doors in either the α- or β-configuration. Prior applications have been reported for β-(1→4),[29] α-(1→2)[82] and α-(1→3)[82] linkages, and the results for β-(1→2) and α/β-(1→6) linkages are forthcoming. The findings reported herein provide further evidence of the general applicability of the method to all types of *O*-glycosidic linkages, but additional studies are needed to substantiate this claim.

Acknowledgements

This work was supported by the National Science Foundation (CHE 1707660 and CHE 2002625 to A. S.) and by Omicron Biochemicals, Inc., South Bend, IN. The Notre Dame Radiation Laboratory is supported by the Department of Energy Office of Science, Office of Basic Energy Sciences, under Award Number DE-FC02-04ER15533. This is document number NDRL 5401.

Supporting Information Available

Preparation of disaccharides **3**$^{1',3}$, **4**$^{1'}$ and **4**$^{2'}$; ^{1}H Chemical shift assignments for **3;** ^{13}C Chemical shift assignments for **3;** ^{1}H-^{1}H Spin-coupling constants in **3**; Intra-residue ^{13}C-^{1}H spin-coupling constants in **3**; Intra-residue ^{13}C-^{13}C spin-coupling constants in **3**; ^{1}H and ^{13}C Chemical shifts, and NMR spin-coupling constants, for **4;** ^{1}H NMR spectrum of **3**; ^{1}H NMR spectrum of **3**$^{1',3}$; ^{13}C{^{1}H} NMR spectrum of **3**; ^{13}C{^{1}H} NMR spectrum of **3**$^{1',3}$; 2D ^{1}H-^{1}H DQF-COSY spectrum of

3; Expansion of the 2D ^{1}H-^{1}H DQF-COSY spectrum of **3**; 2D ^{1}H-^{1}H TOCSY spectrum of **3**; 2D ^{1}H-^{13}C gHSQC spectrum of **3**; Measurement of $^{3}J_{C3,H1'}$ in the ^{1}H NMR spectrum of **3**$^{1',3}$; Measurement of $^{3}J_{C1',H3}$ in the ^{1}H NMR spectrum of **3**$^{1',3}$; Measurement of $^{3}J_{C3,H5}$ in the ^{1}H NMR spectrum of **3**$^{1',3}$; Torsion angle constraints in DFT calculations of **3**^{c}; Torsion angle constraints in DFT calculations of **4**^{c}; DFT-parameterized *J*-coupling equations: αGal-(1→3)-βGalOCH$_3$ (**3**^{c}); DFT-parameterized *J*-coupling equations: βGal-(1→3)-βGalOCH$_3$ (**4**^{c}); Back-calculated *J*-couplings from *MA'AT* models of ϕ' and ψ' in **3** and **4;** Back-calculated spin-coupling constants from MD models of ϕ' and ψ' in **3**^{c} and **4**^{c}; Cartesian coordinates for **3**^{c} and **4**^{c}; Complete references 38 and 54.

References

1. M. Sykes, D.H. Sachs, *Nature Reviews Nephology* **2022**, *18*, 745–761.

2. M. Sandrin, H. A. Vaughan, P. L. Dabkowski, I. F. C. McKenzi, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11391–11395.

3. D. K. C. Cooper, E. Koren, R. Oriol, *Immunol. Rev*. **1994**, *141*, 31–58.

4. U. Galili, *Immunol. Today* **1993**, *14*, 480–482.

5. M. S. Sandrin, I. F. McKenzie, *Immun. Rev*. **1994**, *141*, 169–190.

6. D. K. C. Cooper, *Glycobiology* **2016**, *26*, 571–581.

7. L. Kreft, A. Schepers, M. Hils, K. Swiontek, A. Flatley, R. Janowski, M. K. Mirzaei, M. Dittmar, N. Chakrapani, M. S. Desai, S. Eyerich, L. Deng, D. Niessing, K. Fischer, R. Feederle, S. Blank, C. B. Schmidt-Weber, C. Hilger, T. Biedermann, C. Ohnmacht, *Front. Immunol*. **2022**, *13*, 958952. doi.org/10.3389/fimmu.2022.958952

8. S. Chinnakotla, I. J. Fox, *Curr. Opin. Organ Transplant*. **2002**, *7*, 35–40.

9. T. Z. Arabi, B. N. Sabbah, A. Lerman, X.-Y. Zhu, L. O. Lerman, *Cell Transplantation* **2023**, *32*, 1–10.

10. B. E. Samuelsson, L. Rydberg, M. E. Breimer, A. Backer, M. Gustavsson, J. Holgersson, E. Karlsson, A.-C. Uyterwaal, T. Cairns, K. Welsh, *Immunol. Rev*. **1994**, *141*, 151–168.

11. U. Galili, M. R. Cark, S. B. Shohet, J. Buehler, B. A. Macher, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1369–1373.

12. U. Galili, S. B. Shohet, E. Kobrin, C. L. M. Stults, B. A. Macher, *J. Biol. Chem*. **1988**, *263*, 17755–17762.

13. J. Fang, J. Li, X. Chen, Y. Zhang, J. Wang, Z. Guo, W. Zhang, L. Yu, K. Brew, P. G. Wang, *J. Am. Chem. Soc*. **1998**, *120*, 6635–6638.

14. J. D. Macdougall, K. O. Thomas, O. I. Iweala, *ImmunoTargets and Therapy* **2022**, *11*, 37–54.

15. J. R. Couchman, C. A. Pataki, *J. Histochem. Cytochem*. **2012**, *60*, 885–897.

16. N. S. Gandhi, R. L. Mancera, *Chem. Biol. Drug Des*. **2008**, *72*, 455–482.

17. T. R. O'Leary, M. Critcher, T. N. Stephenson, X. Yang, A. A. Hassan, N. M. Bartfield, R. Hawkins, M. L. Huang, *Nat. Chem. Biol*. **2022**, *18*, 634–642.

18. R. U. Lemieux, K. Bock, L. T. J. Delbaere, S. Koto, V. S. Rao, *Can. J. Chem*. **1980**, *58*, 631–653.

19. K. Bock, A. Brignole, B. W. Sigurskjold, *J. Chem. Soc., Perkin Trans*. *2* **1986**, 1711–1713.

20. A. Imberty, E. Mikros, J. Koca, R. Mollicone, R. Oriol, S. Perez, *Glycoconjugate J*. **1995**, *12*, 331–349.

21. C. A. Stortz, A. S. Cerezo, *J. Carbohydr. Chem*. **1994**, *13*, 235–247.

22. F. Bizid, I. Tvaroska, *Chem. Pap*. **1995**, *49*, 202–214.

23. A. Otter, R. U. Lemieux, R. G. Ball, A. P. Venot, O. Hindsgaul, D. R. Bundle, *Eur. J. Biochem.* **1999**, *259*, 295–303.

24. B. Bose, S. Zhao, R. Stenutz, F. Cloran, P. B. Bondo, G. Bondo, B. Hertz, I. Carmichael A. S. Serianni, *J. Am. Chem. Soc.* **1998**, *120*, 11158–11173.

25. F. Cloran, I. Carmichael, A. S. Serianni, *J. Am. Chem. Soc*. **1999**, *121*, 9843–9851.

26. F. Cloran, I. Carmichael, A. S. Serianni, *J. Am. Chem. Soc*. **2000**, *122*, 396–397.

27. M. J. Hadad, W. Zhang, T. Turney, L. Sernau, X. Wang, R. J. Woods, A. Incandela, I. Surjancev, A. Wang, M.-K. Yoon, A. Coscia, C. Euell, R. Meredith, I. Carmichael, A. S. Serianni, in New Developments in NMR 10: NMR in Glycoscience and Glycotechnology, ed. T. Peters and K. Kato, Royal Society of Chemistry, 2017, pp. 20–100.

28. C. Foces-Foces, F. H. Cano, S. Garcia-Blanco, *Acta. Cryst*. **1980**, *B36*, 377–384.

29. W. Zhang, T. Turney, R. Meredith, Q. Pan, L. Sernau, X. Wang, X. Hu, R. J. Woods, I. Carmichael, A. S. Serianni, *J. Phys. Chem. B* **2017**, *121*, 3042–3058.

30. R. J. Meredith, L. Sernau, A. S. Serianni, *J. Chem. Inf. Model.* **2022***, 62,* 3135–3141.

31. A. A. Shaw, C. Salaun, J.-F. Dauphin, B. Ancian, J*. Magn. Reson. A* **1996**, *120*, 110–115.

32. B. Ancian, I. Bourgeois, J. F. Dauphin, A. A. Shaw, *J. Magn. Reson. A* **1997**, *125*, 348–354.

33. L. Braunschweiler, R. R. Ernst*, J. Magn. Reson*. **1983**, *53*, 521–528.

34. J. Cavanagh, M. Rance, *J. Magn. Reson*. **1990**, *88*, 72–85.

35. G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett*. **1980**, *69*, 185–189.

36. A. Palmer III, J. Cavanagh, P. E. Wright, M. Rance, *J. Magn. Reson*. **1991**, *93*, 151–170.

37. A. Meissner, O. W. Sørensen, *Magn. Reson. Chem*. **2001**, *39*, 49−52.

38. M. J. Frisch, *et al*., *Gaussian*16, Revision B.01, Gaussian Inc., Wallingford, CT, 2016, https://gaussian.com/gaussian16/.

39. A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.

40. A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 1372–1377.

41. W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys*. **1972**, *56*, 2257–2261.

42. E. Cances, B. Mennucci, J. Tomasi, *J. Chem. Phys*. **1997**, *107*, 3032–3041.

43. G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, 132, 114110. doi.org/10.1063/1.3359469

44. V. Sychrovsky, J. Grafenstein, D. Cremer, *J. Phys. Chem*. **2000**, *113*, 3530–3547.

45. T. Helgaker, M. Watson, N. C. Handy, *J. Chem. Phys*. **2000**, *113*, 9402–9409.

46. V. Barone, J. E. Peralta, R. H. Contreras, J. P Snyder, *J. Phys. Chem. A.* **2002**, *106*, 5607–5612.

47. T. Klepach, H. Zhao, X. Hu, W. Zhang, R. Stenutz, M. J. Hadad, I. Carmichael, A. S. Serianni, in Glycoinformatics: Methods in Molecular Biology, ed. T. Lütteke and M. Frank, Springer, New York, 2015, pp. 289–331.

48. R. Stenutz, I. Carmichael, G. Widmalm, A. S. Serianni, *J. Org. Chem*. **2002**, *67*, 949–958.

49. B. Bose-Basu, T. Klepach, G. Bondo, P. B. Bondo, W. Zhang, I. Carmichael, A. S. Serianni, *J. Org. Chem*. **2007**, *72*, 7511–7522.

50. H. Akaike, *IEEE Transactions on Automatic Control* **1974**, *AC-19*, 716–723.

51. Complex Carbohydrate Research Center (CRCC), University of Georgia. http://www.glycam.org

52. K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González- Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622–655.

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

54. D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. I. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, H. *et al*. AMBER 14, **2014**, University of California, San Francisco.

55. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.

56. W. F. van Gunsteren, H. J. C. Berendsen, *Mol. Phys.* **1977**, *34*, 1311–1327.

57. A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555.

58. K. N. Kirschner, R. J. Woods, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541–10545.

59. Prism 8 for Mac OS X, GraphPad Software, Version 8.4.2 (464), April 18, 2020.

60. K. G. I. Nilsson, *Trends Biotechnol.*, **1988**, *6*, 256–264.

61. K. G. I. Nilsson, *Carbohydr. Res.*, **1987**, *167*, 95–103.

62. K. G. I. Nilsson, *Tetrahedron Lett.* **1997**, *38 (1)*, 133–136.

63. B. Zeuner, C. Jers, J. D. Mikkelsen, A. S. Meyer, *J. Agri. Food Chem*. **2014**, *62*, 9615–9631.

64. Z. Guo, L. Wang, L. Su, S. Chen, W. Xia, I. André, C. Rovira, B. Wang, J. Wu, *J. Phys. Chem. Lett.* **2022**, *13*, 56526–5632.

65. P. W. Austin, F. E Hardy, J. C. Buchanan, J. Baddiley, *J. Chem. Soc*. **1963**, 5350–5353.

66. C. T. Bishop, F. P. Cooper, *Can. J. Chem*. **1963**, *41*, 2743–2758.

67. M. Haese, K. Winterhalter, J. Jung, M. S. Schmidt, *Top. Curr. Chem*. **2022**, *380*, 26. doi.org/10.1007/s41061-022-00383-9.

68. C. A. Podlasek, J. Wu, W. A. Stripe, P. B. Bondo, A. S. Serianni, *J. Am. Chem. Soc*. **1995**, *117*, 8635–8644.

69. 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.

70. R. Lemieux, *Pure Appl. Chem*. **1971**, *25*, 527−548.

71. J.-P. Praly, R. Lemieux, *Can. J. Chem*. **1987**, *65*, 213−223.

72. H. Thøgersen, R. Lemieux, K. Bock, B. Meyer, *Can. J. Chem*. **1982**, *60*, 44−57.

73. A. J. Kirby, The Anomeric Effect and Related Stereoelectronic Effects at Oxygen; Springer Verlag: Berlin, 1983.

74. I. V. Alabugin, Stereoelectronic Effects. A Bridge Between Structure and Reactivity; John Wiley & Sons: West Sussex, U.K., 2016.

75. E. Juaristi, G. Cuevas, The Anomeric Effect, CRC Press, Boca Raton, 1995, pp. 95–111.

76. J. Li, M. B. Ksebati, W. Zhang, Z. Guo, J. Wang, L. Yu, J. Fang, P. G. Wang, *Carbohydr. Res*. **1999**, *315*, 76–88.

77. B. Blanchard, A. Nurisso, E. Hollville, C. Tétaud, J. Wiels, M. Pokorná, M. Wimmerová, A. Varrot, A. Imberty, *J. Mol. Biol*. **2008**, *383*, 837–853.

78. B. Y. Choe, G. C. Ekborg, L. Rodén, S. C. Harvey, N. R. Krishna, *J. Am. Chem. Soc*. **1991**, *113*, 3743–3749.

79. N. R. Krishna, B. Y. Choe, M. Prabhakaran, G. C. Ekborg, L. Rodén, S. C. Harvey, *J. Biol. Chem.* **1990**, *265*, 18256–18262.

80. N. R. Krishna, B. Y. Choe, S. C. Harvey, In Computer Modeling of Carbohydrate Molecules; A. French, J. W. Brady, Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 1990; Chapter 14, pp. 227–239.

81. W. Zhang, A. G. Oliver, A. S. Serianni, *Acta. Cryst*. **2012**, *C68*, o7–o11.

82. W. Zhang, R. Meredith, Q. Pan, X. Wang, R. J. Woods, I. Carmichael, A. S. Serianni, *Biochemistry* **2019**, *58*, 546–560.

83. R. J. Meredith, T. Tetrault, M.-K. Yoon, W. Zhang, I. Carmichael, A. S. Serianni, *J. Org. Chem*. **2022**, *87*, 8368–8379.

84. T. Turney, Q. Pan, L. Sernau, I. Carmichael, W. Zhang, X. Wang, R. J. Woods, A. S. Serianni, *J. Phys. Chem. B* **2017**, *121*, 66–77.

85. R. J. Meredith, M. McGurn, C. Euell, P. Rutkowski, E. Cook, I. Carmichael, A. S. Serianni, *Biochemistry* **2022**, *61*, 239–251.

86. R. J. Meredith, R. J. Woods, I. Carmichael, A. S. Serianni, *Phys. Chem. Chem. Phys*. **2020**, *22*, 14454–14457.

87. R. J. Meredith, I. Carmichael, A. S. Serianni, *ACS Omega* **2022**, *7*, 23950–23966.

88. R. J. Meredith, I. Carmichael, R. J. Woods, A. S. Serianni, *Acc. Chem. Res*. **2023**, *56*, 2313–2328.

89. T. Tetrault, R. J. Meredith, W. Zhang, I. Carmichael, A. S. Serianni, *J. Phys. Chem. B* **2022**, *126*, 9506–9515.

90. T. Tetrault, R. J. Meredith, M.-K. Yoon, C. Carizares, A. G. Oliver, I. Carmichael, A. S. Serianni, *Phys. Chem. Chem. Phys*. **2023**, *25*, 16048–16059.

TOC Graphic