



PCCP

**Reconciling *MA'AT* and Molecular Dynamics Models of Linkage Conformation in Oligosaccharides**

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-COM-03-2020-001389.R1
Article Type:	Communication
Date Submitted by the Author:	22-May-2020
Complete List of Authors:	Serianni, Anthony; University of Notre Dame, Department of Chemistry and Biochemistry; Omicron Biochemicals Inc., Meredith, Reagan; University of Notre Dame, Department of Chemistry and Biochemistry Woods, Robert; University of Georgia, Complex Carbohydrate Research Center Carmichael, Ian; University of Notre Dame, Radiation Laboratory and Department of Chemistry & Biochemistry

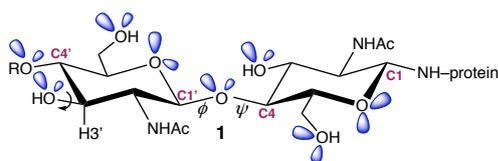
SCHOLARONE™  
Manuscripts

## COMMUNICATION

Reconciling *MA'AT* and Molecular Dynamics Models of Linkage Conformation in OligosaccharidesReagan J. Meredith,<sup>a</sup> Robert J. Woods,<sup>b</sup> Ian Carmichael,<sup>c</sup> and Anthony S. Serianni<sup>a\*</sup>Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

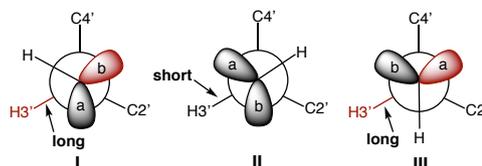
*MA'AT* conformational models of the  $\phi$  torsion angles of *O*-glycosidic linkages differ from those obtained from MD simulation. To determine the source of the discrepancy, *MA'AT* analyses were performed using DFT-derived equations obtained with and without *psi* constraints. The resulting  $\phi$  models were essentially the same, indicating a force-field problem. Circular standard deviations (CSDs) were found to provide reliable estimates of torsional averaging.



Scheme 1. The characteristic high abundance of oxygen lone-pair orbitals in saccharides, illustrated in the  $\beta$ GlcNAc-(1 $\rightarrow$ 4)- $\beta$ GlcNAc disaccharide fragment (1) found in human *N*-glycans.

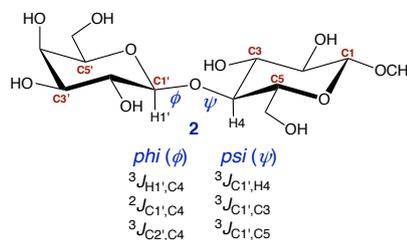
Stereoelectronic effects in saccharides play key roles in dictating their structural properties in solution.<sup>1–3</sup> These effects are pervasive due to the high abundance of oxygen lone-pair orbitals in these molecules (Scheme 1). The best characterized stereoelectronic effects are the *endo*- and *exo*-anomeric effects<sup>1–9</sup> that influence the geometry about R<sub>1</sub>–O–C–O–R<sub>2</sub> fragment such as found near the anomeric carbon of aldopyranosyl rings. The *endo*-anomeric effect influences preferred orientation of the C1–O1 bond in these rings (axial vs equatorial) and favors the former due to optimal  $n \rightarrow \sigma^*$  overlap. The *exo*-anomeric effect influences the rotational properties of the C1–O1 bond, and its strength is determined by C1–O1 bond orientation, with axial orientations weakening the effect due to competing  $n \rightarrow \sigma^*$  interactions (the equatorial orientation is devoid of this competition). In general, in *O*-glycosidic linkages such as the  $\beta$ -(1 $\rightarrow$ 4) linkage shown in Scheme 1, the  $\phi$  ( $\phi$ )

torsion angle favors a geometry that orients C2' roughly anti to C4, whereas the least favored geometry orients O5' anti to C4 in which the two lone-pair orbitals on O1' and O5' are eclipsed.



Scheme 2. Vicinal oxygen lone-pair effects on C–H bond lengths in saccharides, illustrated in Newman projections for  $r_{C3',H3'}$  in the three perfectly staggered rotamers of the C3'–O3' bond in 1 (see Scheme 1). Rotamers I and III orient a lone-pair orbital anti to the C3'–H3' bond, elongating this bond relative to that found in Rotamer II.

Lone-pair effects involving exocyclic hydroxyl oxygens also exist in saccharides, and depend on the rotational properties of the C–O bond. For example, when the oxygen lone-pair orbital on O3' of 1 is anti to the C3'–H3' bond, the latter is lengthened relative to that when both orbitals are gauche (Scheme 2).<sup>10,11</sup> The strengths of all of these lone-pair effects are solvent-dependent, partly due to the plasticity of hybridization at oxygen in different solvents, especially hydrogen bonding solvents. These solvent perturbations lead to different orbital overlaps depending on whether the implicated oxygens are  $sp^3$ - or  $sp^2$ -hybridized, with the residual orbital in the latter case being a  $p$ -orbital roughly perpendicular to the O5–C1–O1 plane.<sup>12,13</sup>



Scheme 3. Structure of methyl  $\beta$ -D-galactopyranosyl-(1 $\rightarrow$ 4)- $\beta$ -D-glucopyranoside (2) and the conventional redundant  $J$ -couplings sensitive to the  $\phi$  ( $\phi$ ) and  $\psi$  ( $\psi$ ) torsion angles that comprise its internal *O*-glycosidic linkage.

Given these structural and solution variables, quantifying the strengths of oxygen lone-pair effects on *O*-glycosidic linkage conformation and dynamics in solution has been difficult,

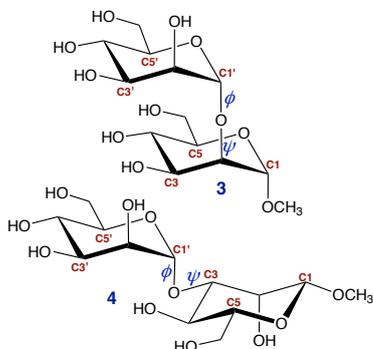
<sup>a</sup> Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN 46556-5670.

<sup>b</sup> Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602.

<sup>c</sup> Radiation Laboratory, University of Notre Dame, Notre Dame, IN 46556-5670.

Electronic Supplementary Information (ESI) available: Plots of calculated trans-*O*-glycoside  $J$ -couplings in 2–4 as a function of the  $\phi$  ( $\phi$ ) torsion angle;  $\phi$ -dependent  $J$ -coupling equations for 2–4 using the full, trimmed and constrained DFT datasets for parameterization; aqueous molecular dynamics simulations (1  $\mu$ s) of 2–4; DFT-derived hydroxymethyl equations 6. See DOI: 10.1039/x0xx00000x

resulting in uncertainties in the force-fields underlying MD simulations that are commonly used to predict their solution behaviors. However, a new experimental method, *MA'AT* analysis,<sup>14–17</sup> is capable of modelling molecular torsion angles in solution in a continuous manner, providing a new opportunity to directly compare the experimental models so obtained to those derived from MD simulation.



Scheme 4. Structures of methyl  $\alpha$ -D-mannopyranosyl-(1 $\rightarrow$ 2)- $\alpha$ -D-mannopyranoside (**3**) and methyl  $\alpha$ -D-mannopyranosyl-(1 $\rightarrow$ 3)- $\beta$ -D-mannopyranoside (**4**), and identification of the  $\phi$  ( $\phi$ ) and  $\psi$  ( $\psi$ ) torsion angles in their internal *O*-glycosidic linkages.

Recent *MA'AT* modelling of the  $\phi$  and  $\psi$  torsion angles in  $\beta$ -(1 $\rightarrow$ 2),  $\beta$ -(1 $\rightarrow$ 4),  $\alpha$ -(1 $\rightarrow$ 2) and  $\alpha$ -(1 $\rightarrow$ 3) *O*-glycosidic linkages (Scheme 3) has shown that, on average, *MA'AT* models of  $\psi$  are in good agreement with corresponding models obtained from aqueous MD simulation (GLYCAM06 force-field).<sup>18</sup> In contrast, and pertinent to this report, comparisons of *MA'AT* and MD models of  $\phi$  reveal consistent and significant differences.<sup>15–17</sup> The latter differences pertain both to the mean values and the circular standard deviations (CSDs), and appear to be independent of the anomeric configuration of the linkage. The *MA'AT*-derived CSDs are typically much larger than those obtained from MD simulation, implying greater librational motion about  $\phi$  than predicted by MD. Of importance to this study is the following question: Are the *MA'AT* models of

Table 1. Summary of Statistics for *MA'AT* and MD<sup>a</sup> Models of the  $\psi$  ( $\psi$ ) Torsion Angles in Disaccharides **2–4**.

$\psi$ Fit	mean (°)	CSD (°)	RMSD (Hz)
Disaccharide <b>2</b>			
full <sup>b,c</sup>	352.3 $\pm$ 7.0	17.9 $\pm$ 11.1	0.35
trimmed <sup>d</sup>	352.4 $\pm$ 6.8	18.3 $\pm$ 10.9	0.30
MD (major; ~94%)	357.8	18.5	0.48 <sup>e</sup> , 0.45 <sup>f</sup>
MD (minor; ~6%)	186.9	13.4	
Disaccharide <b>3</b>			
full <sup>c</sup>	20.9 $\pm$ 8.7	24.4 $\pm$ 10.1	0.21
trimmed <sup>d</sup>	20.1 $\pm$ 8.4	23.8 $\pm$ 10.1	0.20
MD (major; ~75%)	31.6	16.7	0.27 <sup>e</sup> , 0.30 <sup>f</sup>
MD (minor; ~25%)	342.1	15.6	
Disaccharide <b>4</b>			
full <sup>c</sup>	9.3 $\pm$ 10.5	25.0 $\pm$ 9.7	0.33
trimmed <sup>d</sup>	8.0 $\pm$ 11.5	28.8 $\pm$ 9.3	0.36
MD (major; ~66%)	26.1	15.4	0.37 <sup>e</sup> , 0.41 <sup>f</sup>
MD (minor; ~34%)	344.1	16.4	

<sup>a</sup>GLYCAM06 (ref. 18); aqueous 1- $\mu$ s simulation; see details in Supporting Information. <sup>b</sup>Fit gave two solutions; data shown are for the local minimum. <sup>c</sup>Full = fit using all DFT-determined conformers for equation parameterization; <sup>d</sup>trimmed = fit using all DFT-determined conformers except high-energy conformers (>10 kcal/mol), which were removed prior to equation parameterization. *MA'AT* analyses were conducted using the three  $\psi$ -dependent experimental *J*-couplings reported previously in **2–4** (refs. 15 and 17). <sup>e</sup>RMSD values obtained from back-calculated *J*-couplings using MD torsions and full or trimmed sets of equations.

$\phi$  accurate, that is, do they provide a more faithful picture of solution behavior than MD simulation?

Current *MA'AT* analyses of two-bond *O*-glycosidic linkages rely on six NMR *J*-couplings, three sensitive to  $\phi$  and three sensitive to  $\psi$  (Scheme 3).<sup>15–17,19,20</sup> The *J*-values sensitive to  $\psi$  are conventional Karplus-like vicinal (three-bond) values that show a minimal secondary dependence on  $\phi$  and can be parameterized reliably. *MA'AT* modelling of  $\psi$  typically gives very good single-state fits of the experimental data (i.e., RMS errors  $\leq$  0.3 Hz), the resulting models agree reasonably well with the MD models with respect to mean values and CSDs of  $\psi$  in the dominant conformers, and the method can discriminate between relatively small differences in mean values.<sup>15–17</sup> On the other hand, one of the three *J*-couplings sensitive to  $\phi$  is a geminal (two-bond) value with a more limited dynamic range than the remaining two vicinal *J*-values. More importantly, the geminal  $^2J_{C1',Cx}$  displays a notable secondary dependence on  $\psi$ , which results in greater error in the parameterized equation. Does the use of this geminal  $^2J_{COC}$  in *MA'AT* modelling give unreliable models of  $\phi$ , thus explaining the significant discrepancy between the *MA'AT* and MD models?

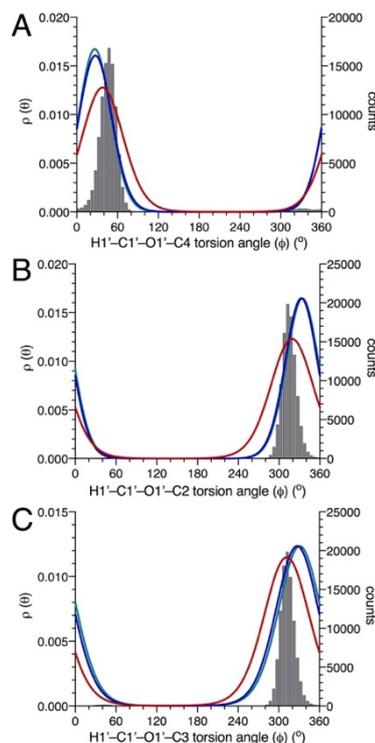


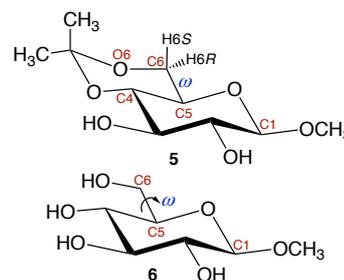
Figure 1. *MA'AT* and MD models of the  $\phi$  ( $\phi$ ) torsion angles in **2** (A), **3** (B) and **4** (C). In each plot, the hatched gray envelope is the MD model, and the green, blue and red lines are the *MA'AT* models obtained using full, trimmed, and  $\psi$ -constrained  $\phi$ -dependent parameterized equations, respectively.

To answer the above questions, equations for the three  $\phi$ -dependent *J*-couplings were parameterized in disaccharides **2–4** (Schemes 3 and 4) that contain three different types of internal *O*-glycosidic linkages,  $\alpha$ -(1 $\rightarrow$ 2),  $\alpha$ -(1 $\rightarrow$ 3) and  $\beta$ -(1 $\rightarrow$ 4) (Figures S1–S3, Supporting Information). Since the  $\psi$  torsion angles in **2–4** are well defined by *MA'AT* analyses (Table 1), we reasoned that a substantial reduction or elimination of

uncertainties in the  $\phi$ -dependent equations could be achieved by re-parameterizing them to include only those  $\psi$  conformations that MA'AT modelling indicates are found in solution; see Supporting Information for the  $\phi$ -dependent equations without (eqs. [1]–[18]; two sets using the full set of conformers and a trimmed set to eliminate high-energy conformers<sup>15–17</sup>) and with (eqs. [19]–[27]) the inclusion of the  $\psi$  constraints. These three sets of  $\phi$ -dependent equations were then used to obtain conformational models of  $\phi$  in **2–4** from MA'AT analysis.

An inspection of Figures S1–S3 reveals that the inclusion of  $\psi$  constraints in the parameterization of the  $\phi$ -dependent equations reduces data scatter in  $J$  vs  $\phi$  plots, leading to parameterized equations with significantly less error, especially for the  $^2J_{\text{COC}}$  and  $^3J_{\text{COC}}$  values. A summary of the statistics for the MA'AT and MD models of  $\phi$  in **2–4** using the three different sets of parameterized equations (full; trimmed,  $\psi$ -constrained; Figures S1–S3, Supporting Information) is shown in Table 2. The same results are illustrated in Figure 1 in which the MA'AT and MD models of  $\phi$  in **2–4** are superimposed for visual comparison. The results show that the quality of MA'AT analyses of  $\phi$  in **2–4** improves considerably when the  $\psi$ -constrained equations are used, as reflected in significantly reduced rms errors ( $\sim 0.3$  Hz to  $\sim 0.1$  Hz). Use of the  $\psi$ -constrained equations in MA'AT analyses gave mean values of  $\phi$  in **2–4** in better agreement with the MD-determined mean values (Figure 1). However, and importantly, the CSDs for  $\phi$  obtained from MA'AT analyses are either similar to, or greater than, corresponding CSDs obtained when either

the full or trimmed equations are used (Table 2). Thus, despite the better fits of the data when  $\psi$ -constrained equations are used to model  $\phi$ , the degree of rotational (librational) motion of



the C1'–O1' bonds in **2–4** determined by MA'AT analysis is considerably greater than that predicted by MD simulation. This discrepancy appears to pertain regardless of the anomeric configuration of the *O*-glycosidic linkage and linkage regiochemistry,<sup>16,17</sup> and has been recapitulated in other di- and oligosaccharides such as  $\beta\text{GlcNAc-(1}\rightarrow\text{2)-}\beta\text{ManOCH}_3$  not discussed here.

The question of whether the CSDs obtained from MA'AT analyses are reliable measures of librational motion about rotatable single bonds was addressed in a separate set of studies of methyl 4,6-*O*-isopropylidene- $\beta$ -D-glucopyranoside (**5**). The presence of the 4,6-*O*-isopropylidene group in **5** constrains rotation about the C5–C6 bond ( $\omega$ ) significantly, thus providing an ideal system to evaluate the reliability of MA'AT-derived CSDs as proxies of rotational dynamics. Three  $J$ -couplings whose magnitudes strongly correlate with  $\omega$  were measured in **5** ( $^2J_{\text{H6R,H6S}}$ ,  $^3J_{\text{H5,H6R}}$  and  $^3J_{\text{H5,H6S}}$ ), and an *in silico* model of **6** was used to parameterize three equations (see eqs. [28–30], Supporting Information) by DFT. The experimental values of  $^2J_{\text{H6R,H6S}}$  ( $-10.8$  Hz),  $^3J_{\text{H5,H6R}}$  (5.4 Hz) and  $^3J_{\text{H5,H6S}}$  (10.5 Hz) were then used in conjunction with eqs. [28–30] to model  $\omega$  in **5**, giving the results shown in Figure 2. The MA'AT model of  $\omega$  in **5** is consistent with the expected  $\omega$  torsion angle

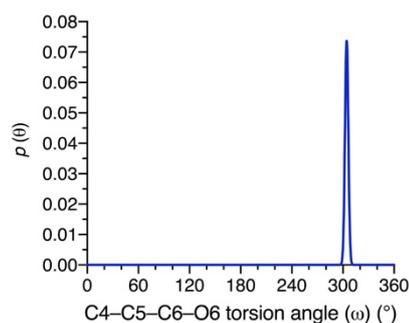


Figure 2. MA'AT model of the  $\omega$  torsion angle in **5** (defined as C4–C5–C6–O6) using  $^2J_{\text{H6R,H6S}}$ ,  $^3J_{\text{H5,H6R}}$  and  $^3J_{\text{H5,H6S}}$  and equations [28–30]. The following results were obtained: mean angle =  $304.3 \pm 4.3^\circ$ ; CSD =  $1.9 \pm 14.2^\circ$ , RMSD = 0.31 Hz.

of  $\sim 300^\circ$ , demonstrating the accuracy of the fit, and yields a very small CSD ( $1.9^\circ$ ), which is consistent with very limited rotational averaging about  $\omega$ .  $J$ -Couplings thus provide not only information on conformational equilibria but also on conformational dynamics in flexible systems, provided that  $J$ -coupling redundancy exists in the conformational element under examination. Redundant  $J$ -values also promise to

Table 2. Summary of Statistics for MA'AT and MD<sup>a</sup> Models of the  $\Phi$  ( $\phi$ ) Torsion Angles in Disaccharides **2–4**.

$\phi$ Fit	mean ( $^\circ$ )	CSD ( $^\circ$ )	RMSD (Hz)
Disaccharide <b>2</b>			
full <sup>b</sup>	$26.5 \pm 11.3$	$24.2 \pm 12.4$	0.31
trimmed <sup>c</sup>	$27.3 \pm 11.7$	$25.3 \pm 12.3$	0.32
$\psi$ -constrained	$37.9 \pm 11.5$	$32.4 \pm 13.3$	0.10
MD	44.1	18.1	$0.83^d$ , $0.79^e$ , $0.49^f$
Disaccharide <b>3</b>			
full <sup>b</sup>	$334.3 \pm 9.7$	$24.6 \pm 11.9$	0.32
trimmed <sup>c</sup>	$332.7 \pm 10.2$	$24.6 \pm 12.9$	0.33
$\psi$ -constrained	$318.7 \pm 11.1$	$33.9 \pm 14.4$	0.12
MD	315.4	12.3	$1.03^d$ , $0.93^e$ , $0.60^f$
Disaccharide <b>4</b>			
full <sup>b</sup>	$331.0 \pm 13.1$	$33.6 \pm 12.2$	0.35
trimmed <sup>c</sup>	$327.2 \pm 14.0$	$33.7 \pm 12.7$	0.28
$\psi$ -constrained	$311.7 \pm 12.5$	$36.7 \pm 14.8$	0.03
MD	313.1	10.7	$0.95^d$ , $0.85^e$ , $0.64^f$

<sup>a</sup>GLYCAM (ref. 18); aqueous 1- $\mu$ s simulation; see details in Supporting Information. <sup>b</sup>Full = fit using all DFT-determined conformers for equation parameterization; <sup>c</sup>trimmed = fit using all DFT-determined conformers except high-energy conformers ( $>10$  kcal/mol), which were removed prior to equation parameterization;  $\psi$ -constrained = fit using DFT-determined conformers consistent with the MA'AT models of  $\psi$ . MA'AT analyses were conducted using the three  $\phi$ -dependent experimental  $J$ -couplings reported previously in **2–4** (refs. 15 and 17). <sup>d</sup>RMSD values obtained from back-calculated  $J$ -couplings using MD torsions and full, <sup>e</sup>trimmed or <sup>f</sup> $\psi$ -constrained sets of equations.

minimize the virtual conformation problem common to interpretations of NMR data in fluxional systems.<sup>21</sup> The line-width of the MA'AT fits, captured in the CSDs, can be viewed as analogous to line-widths ( $T_2^*$ ) in an NMR spectrum, whose values are determined by the intrinsic properties of the molecule (e.g., molecular motion; chemical exchange; etc.) ( $T_2$ ) and by inherent limitations of the experimental measurement (e.g., field inhomogeneity contributions to the experimental line-width). In the present case, uncertainties in the parameterized equations used to fit the redundant  $J$ -couplings contribute to the CSD, with greater uncertainties leading to larger CSDs. As these uncertainties approach zero, the CSD approaches its true value and will accurately reflect the dynamics of the torsion angle. Other factors may also apply, however, such as the number and nature of the functions that describe the equations for the redundant  $J$ -couplings used in the analysis, and the degree to which these functions are unique (i.e., not overlapping). These factors will be treated in more detail in future work.

## Conclusions

This investigation provides new experimental evidence supporting the contention that the current GLYCAM06 force-field parameterization of the *exo*-anomeric effect underestimates librational motion of C1–O1 bonds in *O*-glycosidic linkages involving aldopyranosyl rings, and that force-field revision is needed to bring the MD models into better alignment with experimental models. The MA'AT approach described here is expected to provide a valuable means of validating MD force fields in a wide range of molecules, not only saccharides, if the molecules contain sufficient redundant  $J$ -values that can be parameterized reliably, and if the experimental  $J$ -values can be measured accurately.

## Conflicts of interest

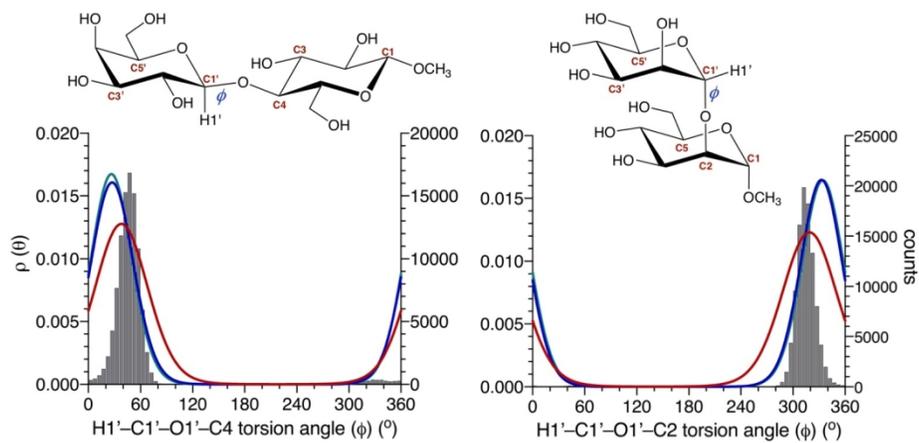
There are no conflicts to declare.

## Notes and references

This work was supported by the National Science Foundation (CHE 1707660 to A. S.) and by Omicron Biochemicals, Inc., South Bend, IN. R. J. W. thanks the National Institutes of Health (U01 CA207824 and P41 GM103390) for financial support. 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 5271.

- 1 A. J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*. Springer-Verlag, Berlin, 1983, p. 3–20.
- 2 E. Juaristi, and G. Cuevas. *The Anomeric Effect*. CRC Press, Boca Raton, 1995, pp. 1–13, 95–109.
- 3 I. V. Alabugin. *Stereoelectronic Effects – A Bridge Between Structure and Reactivity*. Wiley, Chichester, UK, 2016, pp. 129–133.
- 4 J. T. Edward, *Chem. & Ind. (London)* 1955, 1102.
- 5 R. U. Lemieux, R. K. Kullnig, H. J. Bernstein and W. G. Schneider, *J. Am. Chem. Soc.* 1958, **80**, 6098.

- 6 R. U. Lemieux, *Pure Appl. Chem.* 1971, **25**, 527.
- 7 J.-P. Praly and R. U. Lemieux, *Can. J. Chem.* 1987, **65**, 213.
- 8 H. Thogersen, R. U. Lemieux, K. Bock and B. Meyer, *Can. J. Chem.* 1982, **60**, 44.
- 9 P. P. Graczyk and M. Mikolajczyk. *Anomeric Effect: Origins and Consequences*, Topics in Stereochemistry, Vol. 21, Eliel, E. L. & Wilen, S. H., eds., Wiley, New York, 1994, 159–349.
- 10 R. Freymann and J. Gueron, *C. R Acad. Sci.* 1937, **205**, 859.
- 11 A. S. Serianni, J. Wu and I. Carmichael, *J. Am. Chem. Soc.* 1995, **117**, 8645.
- 12 M. Sinnott, *Carbohydrate Chemistry and Biochemistry - Structure and Mechanism*, RSC Publishing, Cambridge, UK, 2013, pp. 53–58.
- 13 A. J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*. Springer-Verlag, Berlin, 1983, pp. 39–48.
- 14 T. Turney, Q. Pan, L. Sernau, I. Carmichael, W. Zhang, X. Wang, R. J. Woods and A. S. Serianni, *J. Phys. Chem. B* 2017, **121**, 66.
- 15 W. Zhang, T. Turney, R. Meredith, Q. Pan, L. Sernau, X. Wang, X. Hu, R. J. Woods, I. Carmichael and A. S. Serianni, *J. Phys. Chem. B* 2017, **121**, 3042.
- 16 W. Zhang, R. Meredith, M.-K. Yoon, X. Wang, R. J. Woods, I. Carmichael and Serianni, *J. Org. Chem.* 2019, **84**, 1706.
- 17 W. Zhang, R. Meredith, Q. Pan, X. Wang, R. J. Woods, I. Carmichael and A. S. Serianni, *Biochemistry* 2019, **58**, 546.
- 18 K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. Gonzalez-Outeirino, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.* 2008, **29**, 622.
- 19 B. Bose, S. Zhao, R. Stenutz, F. Cloran, P. B. Bondo, G. Bondo, B. Hertz, I. Carmichael and A. S. Serianni, *J. Am. Chem. Soc.* 1998, **120**, 11158.
- 20 F. Cloran, I. Carmichael and A. S. Serianni, *J. Am. Chem. Soc.* 1999, **121**, 9843.
- 21 O. Jardetzky, *Biochim. Biophys. Acta* 1980, **621**, 227.



595x334mm (72 x 72 DPI)