

# Distinguishing Epimers Through Raman Optical Activity

Shaun T. Mutter,<sup>†</sup> François Zielinski,<sup>†</sup> Christian Johannessen,<sup>‡</sup> Paul L. A. Popelier,<sup>\*,†</sup> and Ewan W. Blanch<sup>\*,§</sup>

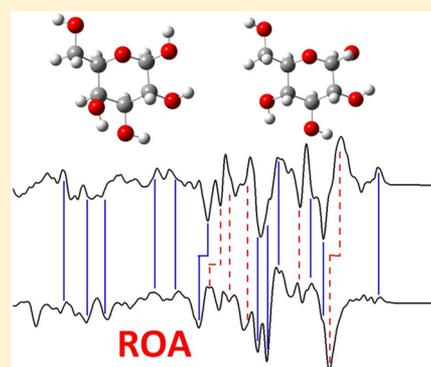
<sup>†</sup>Manchester Institute of Biotechnology and School of Chemistry, University of Manchester, 131 Princess Street, Manchester, M1 7DN, Great Britain

<sup>‡</sup>Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

<sup>§</sup>School of Applied Sciences, RMIT University, GPO Box 2476, Melbourne, Victoria 3001, Australia

## Supporting Information

**ABSTRACT:** The Raman optical activity spectra of the epimers  $\beta$ -D-glucose and  $\beta$ -D-galactose, two monosaccharides of biological importance, have been calculated using molecular dynamics combined with a quantum mechanics/molecular mechanics approach. Good agreement between theoretical and experimental spectra is observed for both monosaccharides. Full band assignments have been carried out, which has not previously been possible for carbohydrate epimers. For the regions where the spectral features are opposite in sign, the differences in the vibrational modes have been noted and ascribed to the band sign changes.



## INTRODUCTION

Carbohydrates are an important class of biomolecules with a wide ranging set of applications in biology, nature, and industry. Despite their immense importance, they are still relatively poorly understood, especially regarding structure–function activity relationships. This is in stark contrast to other biomolecules such as nucleotides and proteins, which have well-defined characteristic secondary and tertiary structures. This is in part due to the difficulty of studying carbohydrate structure. While the “traditional” structural biology techniques of nuclear magnetic resonance (NMR) and X-ray crystallography have been used to great success for analysis of proteins and related systems, the same cannot be said for carbohydrates. Many polysaccharides and complex carbohydrates do not readily form diffractable crystals, nor do they possess conformational motions that occur on time scales easily detectable by NMR. However, a technique that does offer great potential for the determination of carbohydrate structure is Raman Optical Activity (ROA).<sup>1–5</sup>

ROA is a powerful chiroptical technique that has been used to great effect when analyzing chiral molecules.<sup>6–8</sup> This method measures small differences in the Raman scattering of right- and left-circularly polarized light arising from chiral systems. As well as for structure determination, ROA is an informative tool for applications such as the assignment of absolute configuration and analysis of conformational dynamics. Both of these capabilities are important for the analysis of epimers.

Within chiral systems, ROA offers unrivalled sensitivity to molecular structure, sample conditions, and the molecular environment and as such greatly benefits from the support of

computational modeling.<sup>9</sup> In fact, there is a perfect complementarity between experiment and computation. With computed simulations, ROA is able to confirm structural parameters and conformational populations or reveal otherwise unobtainable data such as the specific vibrational nature of observed bands of any intensity.<sup>10–12</sup> The reverse is also true, as experiment can help computation. Indeed, experimental data are a gold standard for guiding force field design and effective solvation modeling.

The majority of combined computational and experimental ROA work has focused on the study of peptides and proteins<sup>4,13–19</sup> with comparatively fewer studies on carbohydrates, although this disparity has been changing through recent methodological developments.<sup>20–22</sup> Some of the early reports by Macleod et al.,<sup>20</sup> Lubber et al.,<sup>21</sup> and Kaminsky et al.<sup>22</sup> confirmed the importance of correctly modeling the solvent and of including conformational dynamics. Without the inclusion of these factors, agreement with experiment stays poor as carbohydrates are subject to strong solvent interactions and show high structural flexibility, which needs to be sampled by using multiple conformations.

Broadly, solvent interactions can be modeled in two ways, that is, using either implicit or explicit solvation methods. Implicit models, such as the polarizable continuum model (PCM), place the solute in a cavity surrounded by a structureless continuum with a dielectric charge. PCM has

**Received:** January 12, 2016

**Revised:** February 28, 2016

**Published:** March 1, 2016

been widely used in studies on peptides to great effect<sup>23–25</sup> but has severe shortcomings for studying carbohydrates, where very poor agreement can be seen with experiment.<sup>26,27</sup> As carbohydrates exhibit strong solvent interactions, an approach using explicit water molecules is required for accurate computation of ROA spectra. Cheeseman et al.<sup>28</sup> illustrated this in their study of methyl- $\beta$ -D-glucose using a combined molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) approach. MD simulations were carried out to account for conformational flexibility as well as for obtaining explicitly solvated snapshots that were optimized using QM/MM with the solvent molecules being treated with a force field method. Computational ROA spectra were generated from averaging the calculated spectra for each snapshot, ensuring that the conformational flexibility of the monosaccharide was retained. This combined approach resulted in excellent agreement between calculated and experimental ROA spectra of carbohydrates for the first time.

Mutter et al.<sup>26</sup> expanded on this protocol while studying two other monosaccharides, D-glucuronic acid and N-acetyl-D-glucosamine. Conformational analysis was included to carefully select only statistically relevant snapshots, thus enabling a reduction in the number of QM/MM calculations and increasing computational efficiency. For D-glucuronic acid, they found that an average over as few as six snapshots resulted in excellent agreement with experiment. This is significantly less than what has been required for good comparison with experiment in similar studies, using random sampling at intervals on the MD trajectory.<sup>29,30</sup> Not only did this study further the computational methodology, it also showed that it was possible to use this approach for larger monosaccharides and, for example, containing different functional groups.

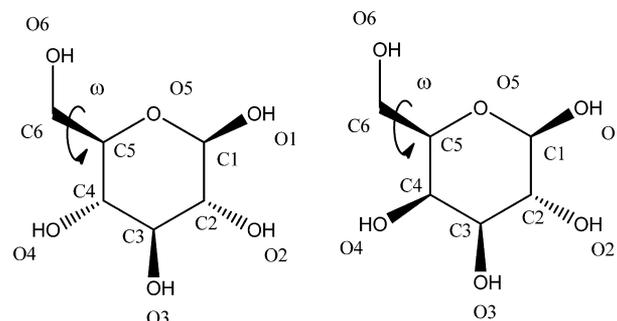
Recently Zielinski et al.<sup>27</sup> carried out a benchmark study using this protocol in order to model the minimum number of explicit solvent molecules required for accurate computation. Using  $\beta$ -D-xylose as a test system, it was found that a spherical solvation shell of radius 10 Å (approximately 120 water molecules) from the center mass of the monosaccharide was the minimum required to obtain spectral consistency. This study also exhibited the power of this approach by obtaining near perfect agreement with experiment.

These previous studies utilizing the MD QM/MM approach focused on methodological development and benchmarking to allow accurate and efficient computation of ROA spectra for monosaccharides. Since this has been successfully achieved, it is now possible to explore other avenues in more detail such as improving upon existing band assignments and understanding how they change with different stereochemistries. Therefore, we present here the first study to examine the sensitivity of ROA spectra to local chirality in monosaccharides and the resultant changes in their vibrational modes using explicitly solvated calculations.

A method for assigning vibrational modes to the calculated ROA spectra generated from multiple explicitly solvated QM/MM optimized snapshots has already been devised by Blanch and co-workers.<sup>31</sup> Definitive assignments followed for D-glucuronic acid, N-acetyl-D-glucosamine, and  $\beta$ -D-xylose with consistencies observed across all three systems.<sup>26,27</sup> In this study, we explore how the ROA spectra change between the epimers of monosaccharides, and the underlying variations in the vibrational modes responsible for the observed spectral responses are identified. While it is well established that ROA spectra of enantiomers are mirror images, many biological

systems feature multiple chiral centers, as such, changing the stereochemistry of a single chiral center can still result in marked differences between spectra.<sup>32</sup>

The monosaccharides chosen for this study are  $\beta$ -D-glucose and  $\beta$ -D-galactose, two fundamentally important sugar molecules (cf. Figure 1 for the atom numbering convention).  $\beta$ -D-



**Figure 1.** Structures of  $\beta$ -D-glucose (left) and  $\beta$ -D-galactose (right) with the numbering scheme used herein. The  $\omega$  torsion angle (O5–C5–C6–O6) is explicitly marked.

galactose is an epimer of  $\beta$ -D-glucose as they only differ in the stereochemistry of one chiral center which is the C4 carbon. Both monosaccharides exhibit three staggered rotamers around the  $\omega$  torsion angle, O5–C5–C6–O6, defined as gauche–gauche (gg:  $\omega \approx -60^\circ$ ), gauche–trans (gt:  $\omega \approx +60^\circ$ ), and trans–gauche (tg:  $\omega \approx 180^\circ$ ). As direct comparison between the two monosaccharides is required, the ratio of 60:40 gg/gt has been chosen for consistency instead of using the MD conformer populations as used in earlier work. However, we have included the results obtained using the MD conformer populations in the Supporting Information and very good agreement with experiment can still be observed. This ratio has been shown to be accurate for glucopyranosides in water, by analysis of X-ray structures<sup>33</sup> and NMR studies with long time scale MD simulations.<sup>34</sup> Only the  $\beta$  anomer of both monosaccharides has been included in this study. It is important to note that exclusion of the  $\alpha$  anomer seems to have little to no detrimental effect on the agreement between theoretical and experimental spectra. This assertion can be deduced from the fact that the level of agreement between experiment and computation is still excellent.

Establishing a deeper understanding of the vibrational nature of ROA bands is of paramount importance. The majority of the current studies of computational ROA of carbohydrates focus on monosaccharides but the current work can be used as the framework required for future studies on oligosaccharides and polysaccharides. A more detailed understanding of ROA band origins and how bands change between different sugar stereochemistries will assist in the observation and determination of spectral marker bands that can be used as signifiers of higher order carbohydrate structure in the future.

## METHODS

**Overall Protocol.** Our computational method has been outlined in detail in previous work and shown to be highly effective for the accurate computation of monosaccharide ROA spectra.<sup>26,27,31</sup> The general protocol can be summarized as follows. Explicitly solvated MD simulations, followed by accurate conformational analysis, give an optimal number of systems for further QM/MM optimization. Raman and ROA

spectra are then calculated for these optimized systems, which can then be compared to experimental spectra. This process can be modified and repeated until a satisfactory level of agreement is achieved.

All *ab initio* calculations were carried out using GAUSSIAN09<sup>35</sup> and all MD simulations were performed using DL POLY v4.<sup>36</sup>

**Molecular Dynamics Simulations.** DL\_FIELD v3.1<sup>37</sup> was used to create the explicitly solvated systems by immersing the monosaccharide within a cubic cell of preoptimized TIP3P water. GLYCAM06<sup>38</sup> force field parameters were included and the water molecules were constrained to a rigid body model to increase computational efficiency. The periodic cubic cell size of 40 Å, the minimum available within DL\_FIELD, was cropped to 30 Å, which was deemed suitable for monosaccharides.<sup>26,27</sup>

Electrostatic interactions were handled with the smooth particle-mesh Ewald method with a cut off of 7 Å ( $>2\sigma_{\max}$ ) for the real-space part and van der Waals interactions. First, the water molecules and then the full system were progressively equilibrated at 298 K by relaxing the force-capping threshold from 100 to 10 000 kT/Å. The equilibrated system's volume was relaxed to convergence at 1 atm pressure within an NPT ensemble. The integration time step was set at 0.5 fs with Berendsen coupling set at 0.1 ps (thermostat) and 1.0 ps (NPT barostat) as time constants. Production runs were carried out within an NVT ensemble with a 50 ns simulation time frame, which is considered adequate for full sampling of monosaccharide conformers.<sup>39</sup>

**Conformational Analysis.** Effective conformational analysis of the MD trajectories leads to the extraction of only statistically significant structures for further computations. The current approach is computationally more efficient than random sampling (at regular intervals across the entire trajectory) because with the former approach an optimal number of structures are used in further QM/MM calculations.

The key dihedral monitored for  $\beta$ -D-glucose and  $\beta$ -D-galactose is the  $\omega$  torsion angle, O5–C5–C6–O6. This dihedral can exist in three configurations: gg, gt, and tg, with values of  $[-120,0]$ ,  $[0,120]$ , and  $[-180,-120] \cup [120,180]$ , respectively. Snapshots representative of the most populous conformers were obtained by filtering MD frames based on dihedral values closest to the corresponding average. All snapshots were then cropped into spherical systems of 10 Å radius, based on the monosaccharide's center of mass, which we have previously shown to be the optimal radius for systems of this size.<sup>27</sup>

**ROA Calculations.** In order to effectively model the explicitly solvated systems, the QM/MM method was utilized. The two layer ONIOM method, as implemented in GAUSSIAN09, was chosen with the electronic embedding scheme, which incorporates the MM partial charges in the QM Hamiltonian. The initial geometries were taken from the selected MD snapshots with the monosaccharides making up the high/QM layer, computed at the B3LYP/6-31G\* level of theory. The explicit water molecules represented the low/MM layer, computed with the AMBER (parm96) force field using TIP3P parameters.

Two optimization methods were employed for the QM/MM calculations, OPTALL and OPTSOLUTE. The former allows full relaxation of all atoms in the system during the optimization, while the latter freezes the explicit water molecules so they retain the MD geometry and only the

monosaccharide is optimized. OPTALL calculations are computationally expensive because systems of this size with explicit solvent molecules have relatively flat potential energy surfaces and as such the convergence to the optimized geometry is slow. OPTSOLUTE is much less computationally intensive but offers poorer spectral prediction when comparing simulations to experiment.<sup>26,27</sup>

The optimized snapshots were used as starting points for the frequency-dependent ROA tensor calculations. The fully analytical two-step procedure was used<sup>10</sup> with harmonic frequency calculations at the B3LYP/6-31G\* level of theory and the ROA tensors calculated at the B3LYP/rDPS level of theory, as recommended by Cheeseman and Frisch.<sup>40</sup> The simulated incident laser wavelength for all ROA tensor calculations was set at the experimental value of 532 nm.

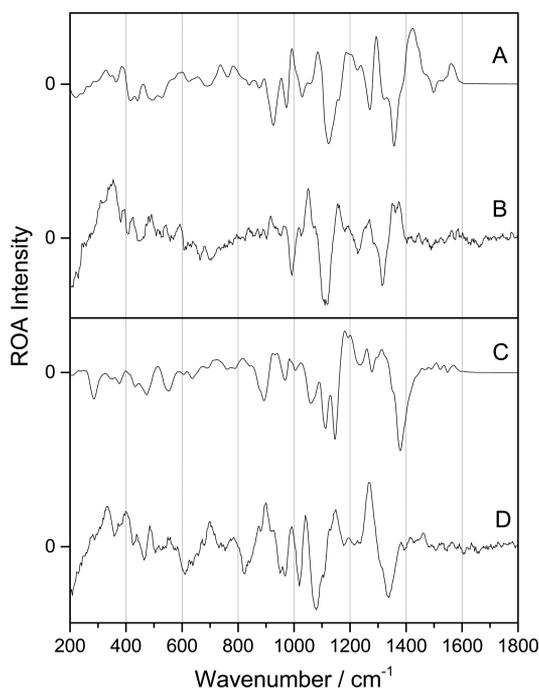
Far from resonance ROA intensities in the scattered circular polarization backscattered configuration (SCP-180) were obtained from the combination of appropriate invariants. Boltzmann and  $\nu^4$  factors were included to allow comparison against experimental spectra. All spectra were generated using a Lorentzian band shape with a peak half-width of 10  $\text{cm}^{-1}$ . As absolute ROA intensities were not measured experimentally the calculated ROA intensities are compared in arbitrary units.

The aim of this study is to directly compare the ROA spectra of  $\beta$ -D-glucose and  $\beta$ -D-galactose and hence it is desirable to keep parameters of the individual simulations as similar as possible. Thus, the snapshot selection was based on two conformers, gg and gt, for both systems. Six snapshots were selected for each conformer in the manner outlined above. To generate the final computed spectrum the six calculated spectra for each conformer were averaged using the line shape form, that is, the curves were averaged instead of averaging ROA intensities at discrete frequency values. The average spectra of each conformer were then weighted at the experimentally determined ratio of 60:40 gg/gt to give the final spectrum comprising 12 individual systems.

**Experimental Measurements.** D-glucose and D-galactose were purchased from Sigma-Aldrich and used without further purification. The ROA spectra were measured in water, at ambient temperatures using the previously described ChiralRAMAN instrument from Biotools Inc.<sup>41</sup> This instrument utilizes the scattered circular polarization (SCP) measurement strategy. The ROA spectrum is presented as circular intensity differences ( $I_R - I_L$ ) with  $I_R$  and  $I_L$  denoting the Raman intensities measured with right- and left-circular polarization, respectively. The corresponding Raman spectra (in Supporting Information) are simply the sum of the polarized intensities. The sample concentrations were  $\sim 150$  mM for each sample with experimental conditions of laser wavelength 532 nm, laser power at sample 500 mW, spectral resolution 7  $\text{cm}^{-1}$ , and an acquisition time of 18 h. Cosmic ray spikes were removed from the ROA spectra by means of a median filter, after which the spectra were smoothed using a second-order Savitzky-Golay filter.

## RESULTS AND DISCUSSION

The calculated spectra for  $\beta$ -D-glucose and  $\beta$ -D-galactose are shown in Figure 2. Only the spectra calculated using the OPTALL approach and the experimental spectra are shown. Spectra calculated using OPTSOLUTE can be found in the Supporting Information along with spectra of the isolated monosaccharide molecules in the gas phase and using PCM = water at the same level of theory. Figures S1 and S2 show what



**Figure 2.** ROA spectra of  $\beta$ -D-glucose (A, OPTALL; B, experimental) and  $\beta$ -D-galactose (C, OPTALL; D, experimental).

has already been well established in previous work, namely that gas phase and implicitly solvated calculations poorly recover the ROA spectra of carbohydrates and that OPTSOLUTE performs significantly better than the isolated molecules, but is not as accurate as OPTALL when compared to experiment.<sup>26–28</sup> The experimental and calculated Raman spectra for  $\beta$ -D-glucose and  $\beta$ -D-galactose are also shown in the [Supporting Information](#) in Figures S3 and S4, respectively. The CID values for the main ROA features measured were of the order of  $10^{-4}$ . As the focus of this work is on the local chirality sensitivity of ROA the Raman data will not be further discussed.

**Figure 2** panels A and B show the OPTALL and experimental spectra, respectively, of  $\beta$ -D-glucose. The agreement between the two spectra can generally be considered as very good. The calculated spectra are not scaled in terms of the frequency and thus appear shifted to higher wavenumbers by approximately  $30\text{ cm}^{-1}$  in comparison to the experimental spectra. The strong positive band in panel A at  $1420\text{ cm}^{-1}$  with a shoulder at a higher wavenumber corresponds to the small doublet centered at  $1370\text{ cm}^{-1}$  in panel B. The band profiles are in very good agreement with experiment from  $1100\text{ cm}^{-1}$  and above with corresponding signs but there are some minor discrepancies with intensity. Most notably the sharp positive band at  $\sim 1300\text{ cm}^{-1}$  in panel A, which is of higher intensity than the positive band at  $\sim 1200\text{ cm}^{-1}$  but is of lower intensity in panel B. The bands between  $\sim 900\text{--}1000\text{ cm}^{-1}$ , exhibiting a  $-ve, -ve, +ve$  profile are also a discrepancy, as their intensity is greatly overestimated in the calculated spectra. The low wavenumber region, between  $200\text{--}700\text{ cm}^{-1}$ , shows some differences in intensity also, although this region is notoriously difficult to model accurately due to the librational nature of the vibrations and the strong solvent interactions. This can also be noted in [Figures S1 and S2](#), as the gas phase and PCM simulations give rise to very poor agreement with experiment in this region. Even though the intensities are lower in panel A in [Figure 2](#) than in panel B, the band profiles are comparable.

Panels C and D in [Figure 2](#) show the OPTALL and experimental spectra, respectively, for  $\beta$ -D-galactose. Once again the agreement between experiment and computation is very good. The band profile from  $1050\text{ cm}^{-1}$  and above follows the correct signs and corresponding wavenumber values for both computation and experiment with the computed spectra shifted to slightly higher wavenumbers by approximately  $40\text{ cm}^{-1}$ . There are similar discrepancies regarding intensity values as those observed for  $\beta$ -D-glucose. The strong positive band at  $\sim 1300\text{ cm}^{-1}$  is higher in intensity than the positive band at  $\sim 1150\text{ cm}^{-1}$  in panel D, while the opposite occurs for the corresponding bands in panel C. The negative doublet at  $\sim 1100\text{ cm}^{-1}$  in panel C corresponds to the strong negative band with a shoulder at higher wavenumber at  $\sim 1070\text{ cm}^{-1}$ , the difference is likely due to overlapping bands in the experiment. In the lower wavenumber regions the intensities do not always correspond between experiment and theory but the band profiles are reproduced well.

Despite some discrepancies regarding ROA intensity values, the experimental spectra of  $\beta$ -D-glucose and  $\beta$ -D-galactose are reproduced well in both cases. A possible explanation why the comparison is not perfect is that even though the computed results are very good, calculations of this nature are inherently approximations and as such perfect agreement with experiment can hardly ever be obtained. It should also be noted that for the above simulations only the  $\beta$  anomer was considered, while in experimental systems the  $\alpha$  anomer would likely be present in non-negligible quantities. Although this study regards exploring spectral differences arising from a change in the stereochemistry at the C4 atom between glucose and galactose, and as such only the more stable  $\beta$  anomers were chosen. As the agreement between experiment and theory is still very good with the exclusion of the  $\alpha$  anomers, analysis of the vibrational nature of observed ROA bands by examining just the  $\beta$  anomers is deemed acceptable.

ROA band assignments for carbohydrates are essential to enable advancement of this approach for larger oligosaccharides and polysaccharides. This is in sharp contrast to studies on peptides and proteins where there are distinct markers bands for secondary structure analysis. Barron and co-workers carried out extensive experimental studies on monosaccharides in the early 90s and were able to assign regions of the spectra that correspond to anomeric configurations, ring substituents, and methylene or COH deformations.<sup>42,43</sup> Recently we have reported the first ROA band assignments for monosaccharides from combined MD QM/MM calculations, using D-glucuronic acid, N-acetyl-D-glucosamine and  $\beta$ -D-xylose as test cases.<sup>26,27</sup> The same ROA band assignment was carried out for both  $\beta$ -D-glucose and  $\beta$ -D-galactose.

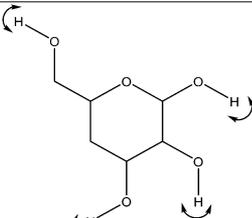
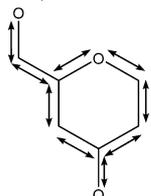
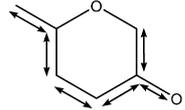
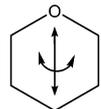
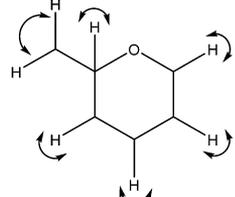
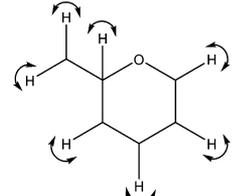
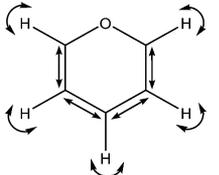
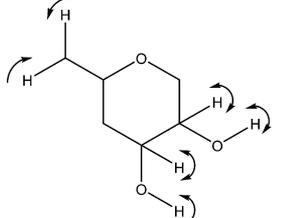
[Tables 1 and 2](#) show the ROA band assignments for  $\beta$ -D-glucose and  $\beta$ -D-galactose, respectively, based on the method outlined in reference.<sup>31</sup> The main features were determined from the OPTALL spectra in panel A in [Figure 1](#) and assignments were obtained at those calculated frequencies for each of the individual structures used to weight the final spectrum. The analysis was confirmed using atomistic displacement data for each of the noted vibrations. However, this analysis does not include low wavenumber vibrations as strong solvent interactions along with their distinct librational nature results in delocalized and nondescriptive assignments. It should be noted that only the most prominent and recurrent vibrations are reported and in many cases a significant number of atoms within the molecules are involved in diffuse vibrations.

Table 1. ROA Band Assignments for  $\beta$ -D-Glucose Based on the Calculated Bands within the OPTALL Scheme<sup>a</sup>

Wavenumber (cm <sup>-1</sup> )	Assignments <sup>b</sup>	Visualisation of mode
973 (957)	O-H wags H-C6-H twist	
992, 1029 (968, 992)	O-H wags	
1084 (1051)	Coupled C-C with C-O stretches	
1123, 1194 (1113, 1156)	Ring deformations	
1238 (1190)	C-H wags	
1271 (1229)	C-H wags	
1293, 1322, 1356 (1266, 1295, 1316)	C-H wags H-C6-H wag	
1424 (1365)	C-C stretches C-H wags	
1499 (1490)	Coupled O-H and C-H wags	
1561 (1560)	H-C6-H scissor O6-H wag	

<sup>a</sup>Values in parentheses are the corresponding wavenumbers for the experimental spectrum. <sup>b</sup>Numbering convention used for assignment and visualization is the same as shown in Figure 1.

Table 2. ROA Band Assignments for  $\beta$ -D-Galactose Based on the Calculated Bands within the OPTALL Scheme<sup>a</sup>

Wavenumber (cm <sup>-1</sup> )	Assignments <sup>b</sup>	Visualisation of mode
983, 1024 (932, 992)	O-H wags	
1059, 1112 (1020, 1079)	Coupled C-C with C-O stretches	
1146 (1107)	C2-O2 stretch C-C stretches	
1190 (1149)	Ring deformations	
1259 (1194)	C-H wags H-C6-H wag	
1313 (1268)	C-H wags	
1379 (1337)	C-C stretches C-H wags	
1508 (1466)	H-C6-H scissor Coupled O-H and C-H wags	

<sup>a</sup>Values in parentheses are the corresponding wavenumbers for the experimental spectrum. <sup>b</sup>Numbering convention used for assignment and visualization is the same as shown in Figure 1.

In recent work on  $\beta$ -D-xylose, general rules were proposed for the diagnostic capability of ROA bands calculated using the combined MD QM/MM approach. Notably, O-H wags appear in the range of 920–1020 cm<sup>-1</sup>, C-C and C-O

stretches appear in the range 1020–1240 cm<sup>-1</sup>, C-H wags appear in the range 1240–1370 cm<sup>-1</sup>, and combined C-C stretches C-H wags in the range 1370–1460 cm<sup>-1</sup>.<sup>27</sup> ROA band assignments for  $\beta$ -D-glucose and  $\beta$ -D-galactose both

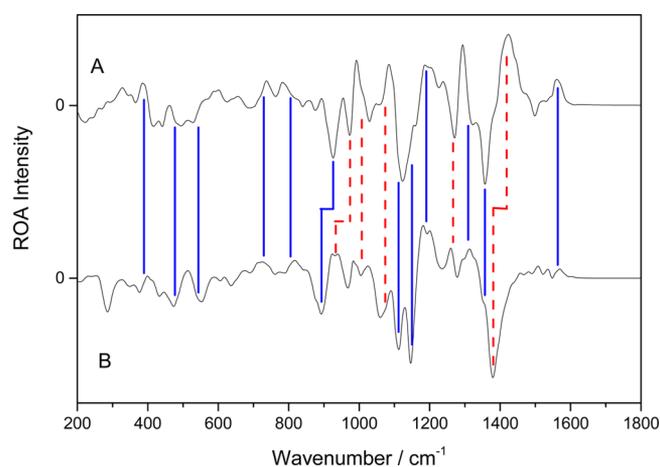
correspond to these rules, despite some bands being just outside the reported ranges.

It is possible to elucidate several further features consistent across multiple monosaccharides. Strong ring deformations occur approximately between  $1100\text{--}1200\text{ cm}^{-1}$  in the region containing C–C and C–O stretches. These vibrational modes not only occur for  $\beta\text{-D-glucose}$  and  $\beta\text{-D-galactose}$  but also for  $\beta\text{-D-xylose}$ <sup>27</sup> and  $N\text{-acetyl-D-glucosamine}$ .<sup>26</sup> Assignments of the bands recorded at the higher wavenumbers within the experimental range also show consistency. Relatively weak positive bands in the approximate region of  $1500\text{--}1560\text{ cm}^{-1}$  arise from multiple assignments. Monosaccharides containing the  $\text{--CH}_2\text{OH}$  moiety at the C5 position show strong  $\text{CH}_2$  scissoring on the C6 atom, as demonstrated by assignments at  $1561$  and  $1508\text{ cm}^{-1}$ , in Tables 1 and 2, respectively. A similar band was observed in  $N\text{-acetyl-D-glucosamine}$  in ref 26, although no assignment was obtained for this band due to its low intensity. In this same region, coupled O–H and C–H wags based around adjacent carbon and oxygen atoms can also be observed for  $\beta\text{-D-galactose}$ ,  $\beta\text{-D-xylose}$ ,<sup>27</sup> and  $\beta\text{-D-glucose}$ , although for the latter molecule this vibration contains some contribution from the  $\text{CH}_2$  scissoring. In Tables 1 and 2, this type of vibration is only shown at set atom positions, C6–O6 in  $\beta\text{-D-glucose}$  and C2–O2 and C3–O3 in  $\beta\text{-D-galactose}$ , but across the multiple snapshots analyzed this vibration was observed over several of the C–O moieties. Therefore, the vibrations recorded under the heading “visualisation of mode” in Tables 1 and 2 are the most recurrent. It should be reiterated that the ROA band assignments outlined above and in Tables 1 and 2 relate to the calculated frequencies using the OPTALL scheme. As shown above, these calculated frequencies are shifted by  $30\text{--}40\text{ cm}^{-1}$  for both monosaccharides and this shift should be considered when using these assignments with experimental ROA data.

As further monosaccharide cases are studied, a more complete picture of the vibrational profile of carbohydrates will come to light. The analysis of the present two molecules has expanded the rules we have set in previous work,<sup>26,27</sup> and will allow more in depth future analysis. The band assignments presented in Tables 1 and 2 can be used to analyze the differences observed in the ROA spectra between the two epimers. Figure 3 shows a comparison between the OPTALL spectra for  $\beta\text{-D-galactose}$  and  $\beta\text{-D-glucose}$ .

Within Figure 3, solid lines show spectral regions where bands are in agreement in sign, and dashed lines show regions where bands are of opposite sign. Wen et al.<sup>32</sup> carried out experimental studies on several monosaccharides, including multiple examples of epimers, and found that remarkable differences were observed between epimeric pairs. Dashed lines on panels A and B in Figure 3 are denoting similar observations. However, only occurrences where the sign of intensity changes were analyzed rather than intensity differences in same signed bands. Therefore, computed bands displaying large differences in intensity but of the same sign are not included in this analysis. For example, the positive band at approximately  $1300\text{ cm}^{-1}$ , which is of much higher intensity in  $\beta\text{-D-glucose}$  than in  $\beta\text{-D-galactose}$ , will not be compared with the corresponding band (i.e., between the two carbohydrates).

The most striking difference between panels A and B in Figure 3 is the strong band around  $1400\text{ cm}^{-1}$ , which is positive for  $\beta\text{-D-glucose}$  (A) and negative for  $\beta\text{-D-galactose}$  (B). This band occurs at a slightly lower wavenumber in panel B and the negative band at  $1350\text{ cm}^{-1}$  in panel A is observed as a shoulder



**Figure 3.** Comparison of the OPTALL spectra of  $\beta\text{-D-glucose}$  (A) and  $\beta\text{-D-galactose}$  (B). Solid blue lines show agreement between the two spectra and dashed red lines show opposite signs between the two spectra.

of the larger band. ROA band assignment shows that bands in this region correspond to C–C stretches combined with C–H wags, which is observed for both monosaccharides. Although the band assignments include multiple vibrations that contribute to the final spectrum from the spectra for each of the snapshots, therefore final assignments are averaged in nature and do not exhibit the full level of sensitivity possible from ROA. However, it is possible to observe trends from vibrational analysis of the individual snapshots. The positive band at  $\sim 1400\text{ cm}^{-1}$  in panel A consistently shows correlated C–H wags for those protons in the axial positions on C1, C2, C3, C4, and C5. However, the negative band at  $\sim 1400\text{ cm}^{-1}$  shows correlated C–H wags for the axial protons on C1, C2 and C3, but as the proton at C4 is in the equatorial position in  $\beta\text{-D-galactose}$ , this correlation of C–H wags is lost, and as such C4–H and C5–H generate ring vibrations that are independent of the other axial protons. This is seen consistently across the multiple snapshots averaged for both  $\beta\text{-D-galactose}$  and  $\beta\text{-D-glucose}$ .

The region  $900\text{--}1050\text{ cm}^{-1}$  also shows significant differences between the spectra of the two monosaccharides, as can be seen from Figure 3. The ROA band assignment shows that bands in this region mainly consist of O–H wags, and similar observations made above for the band at  $1400\text{ cm}^{-1}$  can be found here. Within  $\beta\text{-D-glucose}$ , there are correlated O–H wags from the hydroxyl groups in the equatorial position on the sugar ring. Still, in  $\beta\text{-D-galactose}$  the O–H wag of the O4 oxygen does not contribute in the visualization of modes in Table 2. This does not mean that there is no O4–H wag but still indicates that it does not occur as consistently across all 12 snapshots and appears with a weaker vibrational signature. The remaining bands that differ in sign between panels A and B at  $\sim 1270\text{ cm}^{-1}$  do not display similar conclusive differences between vibrational modes, as can be seen for the bands at  $\sim 1400\text{ cm}^{-1}$  and  $900\text{--}1050\text{ cm}^{-1}$ . The bands at  $\sim 1270\text{ cm}^{-1}$  arise from C–H wags, so a similar pattern where correlated wags are broken by changing the C4 proton to the equatorial position is possible here. It is also possible that the positive band at  $\sim 1270\text{ cm}^{-1}$  in panel B corresponds to the positive shoulder at  $\sim 1250\text{ cm}^{-1}$  in panel A and not the negative band at  $\sim 1270\text{ cm}^{-1}$ . However, vibrational modes in the spectral region  $1240\text{--}1370\text{ cm}^{-1}$  are very similar and it is therefore not

possible to confirm whether the negative band at  $\sim 1270\text{ cm}^{-1}$  in panel B corresponds to one of the same sign or one of opposite sign.

Regarding the bands showing opposite signs between  $\beta$ -D-galactose and  $\beta$ -D-glucose (dashed in Figure 3), the differences in the normal mode vibrations likely contribute to the observed changes in the ROA spectra measured for the two epimers. Unfortunately, it is not yet possible to definitively confirm that these differences are the sole cause for the spectral changes, considering the lack of understanding of ROA spectral features' origin.

## CONCLUSIONS

Utilizing the now established combined MD QM/MM approach for the calculation of ROA spectra of carbohydrates, we have successfully predicted the experimental spectra of  $\beta$ -D-galactose and  $\beta$ -D-glucose. Good agreement between experiment and computation has once again been found. ROA band assignments for both monosaccharides have been obtained, and the assigned vibrational modes are in agreement with those reported in our recent work.<sup>27</sup> Comparison of the current ROA band assignments with those of previously studied monosaccharides has also allowed for the observation of further recurrent carbohydrate ROA bands. These include assignments of ring deformations between  $1100\text{--}1200\text{ cm}^{-1}$  and methylene scissoring and correlated C–H and O–H wags occurring between  $1500\text{--}1550\text{ cm}^{-1}$ .

The main scope of this work was to establish the spectral variations between epimeric pairs of monosaccharides and the differences in vibrational modes leading to these changes. It turns out that changing the stereochemistry of the C4 atom of  $\beta$ -D-glucose to form  $\beta$ -D-galactose yields a strong spectral response. We have highlighted the spectral regions where the difference in intensity between bands for the two monosaccharides led to intensities of opposite signs. The vibrational modes of these bands were further analyzed and differences noted between the two carbohydrates. While the vibrational modes were similar in nature, swapping the proton at C4 from the axial to equatorial position broke the correlation in modes containing C–H and O–H wags. The corresponding modes for  $\beta$ -D-galactose show, therefore, a weaker or nonexistent contribution from the C4 and O4 positions to the diffuse and correlated wagging vibrational modes observed for  $\beta$ -D-glucose.

The deeper understanding of the vibrational modes that make up these ROA bands, with insights such as differences in normal mode vibrations leading to opposite signs in intensities, will allow for more effective analysis of carbohydrates in the future. Coupled with recent advances in methodology, this will enable the accurate computation and analysis of disaccharides and larger carbohydrates in the very near future.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.6b00358.

Experimental and calculated Raman spectra of  $\beta$ -D-glucose and  $\beta$ -D-galactose. ROA spectra for  $\beta$ -D-glucose and  $\beta$ -D-galactose calculated in the gas phase with PCM and the OPTSOLUTE approach. ROA spectra for  $\beta$ -D-glucose and  $\beta$ -D-galactose calculated using the OPTALL approach weighted using conformer ratios determined by MD. (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: paul.popelier@manchester.ac.uk.

\*E-mail: ewan.blanch@rmit.edu.au.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

E.B. and P.L.A.P. thank the UK Engineering and Physical Sciences Research Council (EPSRC) for funding (EP/J019623/1).

## REFERENCES

- (1) Yaffe, N. R.; Almond, A.; Blanch, E. W. A new route to carbohydrate secondary and tertiary structure using Raman spectroscopy and Raman optical activity. *J. Am. Chem. Soc.* **2010**, *132*, 10654–10655.
- (2) Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. Solution structure and dynamics of biomolecules from Raman optical activity. *Prog. Biophys. Mol. Biol.* **2000**, *73*, 1–49.
- (3) Zhu, F.; Tranter, G. E.; Isaacs, N. W.; Hecht, L.; Barron, L. D. Delineation of protein structure classes from multivariate analysis of protein Raman optical activity data. *J. Mol. Biol.* **2006**, *363*, 19–26.
- (4) Yamamoto, S. Conformational analyses of peptides and proteins by vibrational Raman optical activity. *Anal. Bioanal. Chem.* **2012**, *403*, 2203–2212.
- (5) He, Y.; Bo, W.; Dukor, R. K.; Nafie, L. A. Determination of absolute configuration of chiral molecules using vibrational optical activity: a review. *Appl. Spectrosc.* **2011**, *65*, 699–723.
- (6) Barron, L. D.; Buckingham, A. D. Rayleigh and Raman scattering from optically active molecules. *Mol. Phys.* **1971**, *20*, 1111–1119.
- (7) Barron, L. D.; Bogaard, M. P.; Buckingham, A. D. Raman scattering of circularly polarized light by optically active molecules. *J. Am. Chem. Soc.* **1973**, *95*, 603–605.
- (8) Barron, L. D.; Hecht, L.; McColl, I. H.; Blanch, E. W. Raman optical activity comes of age. *Mol. Phys.* **2004**, *102*, 731–744.
- (9) Haesler, J.; Schindelholz, I.; Riguet, E.; Bochet, C. G.; Hug, W. Absolute configuration of chirally deuterated neopentane. *Nature* **2007**, *446*, 526–529.
- (10) Ruud, K.; Thorvaldsen, A. J. Theoretical approaches to the calculation of Raman optical activity spectra. *Chirality* **2009**, *21*, E54–67.
- (11) Pecul, M. New applications and challenges for computational ROA spectroscopy. *Chirality* **2009**, *21*, E98–104.
- (12) Pecul, M.; Ruud, K. Ab initio calculation of vibrational Raman optical activity. *Int. J. Quantum Chem.* **2005**, *104*, 816–829.
- (13) Yamamoto, S.; Furukawa, T.; Bour, P.; Ozaki, Y. Solvated states of poly-L-alanine  $\alpha$ -helix explored by Raman optical activity. *J. Phys. Chem. A* **2014**, *118*, 3655–3662.
- (14) Zhu, F.; Kapitan, J.; Tranter, G. E.; Pudney, P. D. A.; Isaacs, N. W.; Hecht, L.; Barron, L. D. Residual structure in disordered peptides and unfolded proteins from multivariate analysis and ab initio simulation of Raman optical activity data. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 823–833.
- (15) Parchansky, V.; Kapitan, J.; Kaminsky, J.; Sebestik, J.; Bour, P. Ramachandran Plot for Alanine Dipeptide as Determined from Raman Optical Activity. *J. Phys. Chem. Lett.* **2013**, *4*, 2763–2768.
- (16) Kaminski, M.; Kudelski, A.; Pecul, M. Vibrational optical activity of cysteine in aqueous solution: a comparison of theoretical and experimental spectra. *J. Phys. Chem. B* **2012**, *116*, 4976–4990.
- (17) Jacob, C. R.; Lubber, S.; Reiher, M. Understanding the signatures of secondary-structure elements in proteins with Raman optical activity spectroscopy. *Chem. - Eur. J.* **2009**, *15*, 13491–13508.
- (18) Kapitan, J.; Zhu, F.; Hecht, L.; Gardiner, J.; Seebach, D.; Barron, L. D. Solution structures of beta peptides from Raman optical activity. *Angew. Chem., Int. Ed.* **2008**, *47*, 6392–6394.

- (19) Mukhopadhyay, P.; Zuber, G.; Beratan, D. N. Characterizing aqueous solution conformations of a peptide backbone using Raman optical activity computations. *Biophys. J.* **2008**, *95*, 5574–5586.
- (20) Macleod, N. A.; Johannessen, C.; Hecht, L.; Barron, L. D.; Simons, J. P. From the gas phase to aqueous solution: Vibrational spectroscopy, Raman optical activity and conformational structure of carbohydrates. *Int. J. Mass Spectrom.* **2006**, *253*, 193–200.
- (21) Luber, S.; Reiher. Calculated Raman optical activity spectra of 1,6-anhydro-beta-D-glucopyranose. *M. J. Phys. Chem. A* **2009**, *113*, 8268–8277.
- (22) Kaminsky, J.; Kapitan, J.; Baumruk, V.; Bednarova, L.; Bour, P. Interpretation of Raman and Raman optical activity spectra of a flexible sugar derivative, the gluconic acid anion. *J. Phys. Chem. A* **2009**, *113*, 3594–3601.
- (23) Sebek, J.; Kapitan, J.; Sebestik, J.; Baumruk, V.; Bour, P. L-alanyl-L-alanine conformational changes induced by pH as monitored by the Raman optical activity spectra. *J. Phys. Chem. A* **2009**, *113*, 7760–7768.
- (24) Pecul, M.; Deillon, C.; Thorvaldsen, A. J.; Ruud, K. The aqueous Raman optical activity spectra of 4(R)-hydroxyproline: theory and experiment. *J. Raman Spectrosc.* **2010**, *41*, 1200–1210.
- (25) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (26) Mutter, S. T.; Zielinski, F.; Cheeseman, J. R.; Johannessen, C.; Popelier, P. L. A.; Blanch, E. W. Conformational dynamics of carbohydrates: Raman optical activity of D-glucuronic acid and N-acetyl-D-glucosamine using a combined molecular dynamics and quantum chemical approach. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016–6027.
- (27) Zielinski, F.; Mutter, S. T.; Johannessen, C.; Blanch, E. W.; Popelier, P. L. A. The Raman optical activity of  $\beta$ -D-xylose: where experiment and computation meet. *Phys. Chem. Chem. Phys.* **2015**, *17*, 21799–21809.
- (28) Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L. A.; Blanch, E. W. Calculation of Raman optical activity spectra of methyl- $\beta$ -D-glucose incorporating a full molecular dynamics simulation of hydration effects. *J. Am. Chem. Soc.* **2011**, *133*, 4991–4997.
- (29) Hopmann, K. H.; Ruud, K.; Pecul, M.; Kudelski, A.; Dracinsky, M.; Bour, P. Explicit versus implicit solvent modeling of Raman optical activity spectra. *J. Phys. Chem. B* **2011**, *115*, 4128–4137.
- (30) Urago, H.; Suga, T.; Hirata, T.; Kodama, H.; Unno, M. Raman optical activity of a cyclic dipeptide analyzed by quantum chemical calculations combined with molecular dynamics simulations. *J. Phys. Chem. B* **2014**, *118*, 6767–6774.
- (31) Mutter, S. T.; Zielinski, F.; Popelier, P. L. A.; Blanch, E. W. Calculation of Raman optical activity spectra for vibrational analysis. *Analyst* **2015**, *140*, 2944–2956.
- (32) Wen, Z. Q.; Barron, L. D.; Hecht, L. Vibrational Raman optical activity of monosaccharides. *J. Am. Chem. Soc.* **1993**, *115*, 285–292.
- (33) Marchessault, R. H.; Perez, S. Conformations of the hydroxymethyl group in crystalline aldohexopyranoses. *Biopolymers* **1979**, *18*, 2369–2374.
- (34) Gonzalez-Outeirino, J.; Kirschner, K. N.; Thobhani, S.; Woods, R. J. Reconciling solvent effects on rotamer populations in carbohydrates - A joint MD and NMR analysis. *Can. J. Chem.* **2006**, *84*, 569–579.
- (35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision B.01; Gaussian, Inc.: Wallingford CT, 2009.
- (36) Todorov, I. T.; Smith, W.; Trachenko, K.; Dove, M. T. DL\_POLY\_3: new dimensions in molecular dynamics simulations via massive parallelism. *J. Mater. Chem.* **2006**, *16*, 1911–1918.
- (37) Yong, C. W. *DL\_FIELD - A force field and model development tool for DL\_POLY*; Blake, R., Ed.; CSE Frontier, STFC Computational Science and Engineering, Daresbury Laboratory: Warrington, U.K.; 2010, 38.
- (38) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622.
- (39) Kirschner, K. N.; Woods, R. J. Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10541–10545.
- (40) Cheeseman, J. R.; Frisch, M. J. Basis Set Dependence of Vibrational Raman and Raman Optical Activity Intensities. *J. Chem. Theory Comput.* **2011**, *7*, 3323–3334.
- (41) Barron, L. D.; Zhu, F.; Hecht, L.; Tranter, G. E.; Isaacs, N. W. Raman optical activity: An incisive probe of molecular chirality and biomolecular structure. *J. Mol. Struct.* **2007**, *834-836*, 7–16.
- (42) Barron, L. D.; Gargaro, A. R.; Wen, Z. Q. Vibrational Raman optical activity of carbohydrates. *Carbohydr. Res.* **1991**, *210*, 39–49.
- (43) Bell, A. F.; Barron, L. D.; Hecht, L. Vibrational Raman optical activity study of d-glucose. *Carbohydr. Res.* **1994**, *257*, 11–24.