JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Calculation of Raman Optical Activity Spectra of Methyl-β-D-Glucose Incorporating a Full Molecular Dynamics Simulation of Hydration Effects

James R. Cheeseman,[†] Majeed S. Shaik,[‡] Paul L. A. Popelier,[‡] and Ewan W. Blanch*[,§]

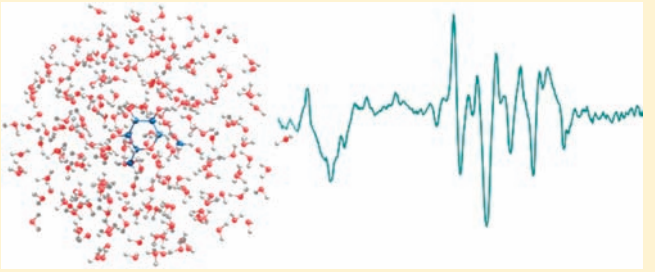
[†]Gaussian, Inc., 340 Quinnipiac Street Building no. 40, Wallingford, Connecticut, United States

[‡]Manchester Interdisciplinary Biocentre and School of Chemistry and [§]Manchester Interdisciplinary Biocentre and Faculty of Life Sciences, University of Manchester, 131 Princess Street, Manchester M1 7DN, U.K.

School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.

Supporting Information

ABSTRACT: We report calculations of the Raman and Raman optical activity (ROA) spectra of methyl-β-D-glucose utilizing density functional theory combined with molecular dynamics (MD) simulations to provide an explicit hydration environment. This is the first report of such combination of MD simulations with ROA ab initio calculations. We achieve a significant improvement in accuracy over the more commonly used gas phase and polarizable continuum model (PCM) approaches, resulting in an excellent level of agreement with the experimental spectrum. Modeling the ROA spectra of carbohydrates has until now proven a notoriously difficult challenge due to their sensitivity to the effects of hydration on the molecular vibrations involving each of the chiral centers. The details of the ROA spectrum of methyl-β-D-glucose are found to be highly sensitive to solvation effects, and these are correctly predicted for the first time including those originating from the highly sensitive low frequency vibrational modes. This work shows that a thorough consideration of the role of water is pivotal for understanding the vibrational structure of carbohydrates and presents a new and powerful tool for characterizing carbohydrate structure and conformational dynamics in solution.



INTRODUCTION

Raman optical activity (ROA) spectroscopy, which measures a small difference in Raman scattering intensities from chiral molecules using circularly polarized light,[1−3] either in the incident or scattered beams or even both, has proven to be an informative probe of molecular conformation. In recent years, the development of ROA has benefited greatly from computational modeling,[4,5] which has improved our understanding of the underlying principles behind this chiroptical technique and revealed information on vibrational motion and protein conformation. As ROA spectral calculations are highly sensitive to even small differences in molecular geometry, they can also serve as a gold standard in force field design and computational methods in general. Most of the recent published studies on computational modeling of ROA spectra have concentrated on amino acids and peptides,[6−12] reflecting the main direction of ROA experimental research, protein structure analysis. However, the structures of other classes of biological molecules are now also beginning to be explored by ROA spectroscopy, and the potential impact of such chiroptical techniques may be even greater for these other biological molecules than for proteins. For example, complex carbohydrates such as glycoproteins and peptidoglycans perform a diverse range of functions in living systems and have generally proven more difficult to structurally characterize than proteins. Although relatively few ROA studies have so far been conducted for carbohydrates, the spectra have been shown to be highly informative on carbohydrate structure. However, interpretation of the spectra is currently limited by our incomplete understanding of spectrum−structure correlations. Following the example of advances in computational modeling of peptide ROA spectra, the potential of carbohydrate ROA spectral modeling is obvious but there have been few studies of this type to date. Pecul and co-workers[13] conducted the first study using linear response SCF theory into parameters influencing ROA spectra of L-lactic acid and D-glyceraldehyde and found that the spectra were particularly sensitive to both intramolecular hydrogen bonding of hydroxyl groups and solvation. The effects of both conformational averaging and solvent interactions were further highlighted by Macleod et al.[14] in their study on D-lactose, D-glucose, and D-galactose. Luber and Reiher[15] have recently considered the effects of interactions of the carbohydrate with solvent water on the spectra of 1,6-anhydro-β-D-glucopyranose, a

Received: December 2, 2010
Published: March 14, 2011

derivative of D-glucose, through the use of both the COSMO continuum model and explicit solvation with a small number of water molecules. These authors clearly demonstrated the importance of the continued symbiosis between ROA experiments and computational modeling and that a considerable improvement in agreement with experiment could be obtained by considering even a small number of explicit water molecules.

In this paper, we report computations of both Raman and ROA spectra of methyl-β-D-glucose in order to more thoroughly establish the role of water molecules in influencing the vibrational modes of carbohydrates and to further develop the modeling of Raman and ROA spectra in general. Typically, reliable modeling of these spectra involves calculation of ROA spectra for each of the energy-minimized conformers of the target molecule and then Boltzmann averaging to approximate the experimental data. As vibrational modes of biological molecules are often coupled to those of solvent water molecules, the effect of the solvent must also be taken into account. Furthermore, it has been shown that the conformational dynamics of functional groups can significantly influence ROA bandshapes.[16] It is now common practice to approximate the influence of solvent water on ROA spectra such as through use of a polarizable continuum model (PCM), in which the solvent is treated as a structureless continuum characterized by its dielectric constant. The polarization of the solvent is represented by the solvent reaction field expressed in terms of a potential defined through an apparent charge distribution spread on the cavity surface.[17,18]

This approach has worked well for peptides but for modeling of ROA spectra of carbohydrates, which contain multiple solvated hydroxyl and other groups, previous reports in the literature suggest that a more detailed and comprehensive account of solvent effects should be included. Consequently, we have conducted the first incorporation of dynamic simulations to ROA modeling in order to more accurately model the number of contributing solvated conformers. Although our strategy is computationally more intensive, it is not only clearly more accurate but also leads to a fuller realization of the hydration of carbohydrates.

The three staggered rotamers of methyl-β-D-glucose are commonly defined with respect to the ω torsion angle (O6−C6−C5−O5; see Figure 1) as gauche−gauche (gg), gauche−trans (gt), and trans−gauche (tg), where ω is −60°, +60°, and 180°, respectively. Statistical analysis of X-ray structures[19] reveals that glucopyranosides show clear preference for only the gg and gt conformations with ratios of approximately 60:40:0 (gg:gt:tg). NMR spectroscopy combined with long time MD simulations (50−100 ns) of glucopyranosides in water have found ratios of 60% gg and 40% gt and negligible amounts of the tg rotamer.[20] Gonzalez-Outeirino et al. observed that in an aqueous environment the water disrupts stabilizing intramolecular hydrogen bonding and that H−O6 acts as a hydrogen bond donor to water. Although MD simulations (50 ns) of eight disaccharides of glucopyranose, or methyl-β-D-glucose, in water[21,22] show different ratios of the gg and gt rotamers, they also reveal negligible amounts of the tg rotamer. Kirschner and Woods[22] have demonstrated the difficulty in using theoretical gas-phase results to understand solution-phase chemistry. Populations of the gg, gt, and tg rotamers were not explained by the gas-phase rotational energy profiles, and they found that water competitively forms intermolecular hydrogen bonds with the carbohydrate and, in doing so, weakens the intramolecular hydrogen bond networks. Correct reproduction of the experimental rotamer distributions about the ω angles was obtained in their study only after explicit water was included in MD simulations, which further highlights the desirability of incorporating MD simulations into ROA computations of carbohydrate structure and dynamics.

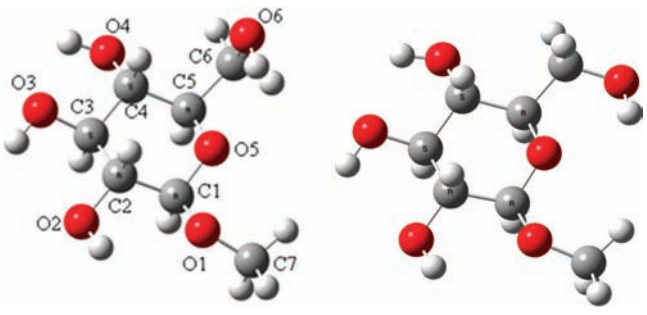


Figure 1. Ball and stick representations of methyl-β-D-glucose in the gg (left) and gt (right) conformations. The carbon and oxygen atoms of the gg rotamer are numbered according to the scheme used throughout this work, carbon atoms are shown in gray, oxygens in red, and hydrogens in white.

In the far-from-resonance theory, where the exciting laser radiation is far from the lowest allowed excited state, ROA intensity differences depend on the normal mode derivatives of three polarizability tensors, namely the electric dipole−electric dipole polarizability, the electric dipole−electric quadrupole polarizability, and the electric dipole−magnetic dipole polarizability. For more detailed descriptions of the theory of ROA and the principles underlying the calculation of ROA tensors, we direct readers to the early papers by Barron and Buckingham,[23] the recent review by the same authors,[24] and the excellent review by Ruud and Thorvaldsen.[4]

EXPERIMENTAL SECTION

Computational Methods. All ab initio calculations were performed using a development version of Gaussian.[25]

Isolated Methyl-β-D-glucose and PCM Hydration. Geometry optimizations and harmonic force field calculations on the isolated conformers of methyl-β-D-glucose (gg and gt) were computed at the B3LYP/6-31G(d) level of density functional theory. Raman and ROA tensors were computed at the HSEH1PBE/rDPS level of theory. HSEH1PBE[26] is a range-separated functional consisting of short-range exact exchange, long-range DFT exchange. The rDPS basis set[27] is the 3-21++G basis set with a semidiffuse p function (exponent 0.2) on hydrogen atoms and has been demonstrated to provide ROA intensity differences close to those obtained using the larger aug-cc-pVDZ basis set.[28] The effects of water for the isolated molecule were included using a polarizable continuum model (PCM), in which the solvent is treated as a structureless continuum characterized by its dielectric constant.[17,18]

Methyl-β-D-glucose with Explicit Water Molecules. A two-layer ONIOM[29] method was used for the QM/MM study of methyl-β-D-glucose surrounded by explicit water molecules. Initial geometries were obtained from snapshots of the MD simulation. An electronic embedding scheme,[30,31] which incorporates the partial charges of the MM region into the QM Hamiltonian, was used throughout. The QM region consisted of methyl-β-D-glucose, and the MM region consisted of all of the water molecules. The snapshots of geometries of methyl-β-D-glucose with water molecules were generated by the MM part. This part is governed by two potentials: the parm96 parameters of the AMBER force field[32] describe glucose and the TIP3P[33] parameters water. The van der Waals interaction between methyl-β-D-glucose and water was modeled using the familiar Lorentz−Berthelot mixing rule, applied to the Lennard-Jones parameters of both potentials.

For geometry optimizations and harmonic frequency calculations, the QM part of the system was computed at the B3LYP/6-31G(d) level of theory. Since the calculations of vibrational normal modes and frequencies are only valid at stationary points on the potential energy surface, the positions of the MM water molecules were either frozen (*optglc* model) or allowed to fully relax (*optall* model). For Raman and ROA tensors, the QM part of the system was computed at the HSEH1PBE/rDPS level of theory as described above.

Ab initio Computations of Raman and ROA Intensities. Frequency-dependent ROA and Raman tensors were computed using magnetic field dependent basis functions (GIAOs),[34,35] using an analytical derivative "two-step" procedure ($n + 1$ algorithm) as described by Ruud and Thorvaldsen[4] at 532 nm. Far from resonance scattered circular polarization (SCP), backscattered Raman and ROA intensities were obtained from the appropriate combinations of the Raman and ROA tensor invariants,[36,37] and include the ν^4 and Boltzmann factors (see the Supporting Information), which are necessary for comparing calculated spectra to experimental spectra. The simulated spectra were generated assuming a Lorentzian band shape with a half-width of 10 cm^{-1}. Since absolute Raman and ROA intensities were not measured experimentally, the calculated Raman and ROA intensities are therefore compared in arbitrary units and are uniformly scaled to best match the experimental spectra.

Conformational Averaging. All calculated ROA spectra for methyl-β-D-glucose are weighted and consist of 60% of the gg conformer and 40% of the gt conformer. For the isolated monosaccharide, only the lowest minimum energy conformations of the gg and gt conformers are used. Spectra for methyl-β-D-glucose surrounded by explicit water molecules were obtained as 60%:40% gg:gt weighted averages of the spectra generated from the 16 MD snapshots (taken every 25 ps) for each of the gg and gt conformers.

Molecular Dynamics Simulations. The all-atom AMBER/ff99SB force field[38,39] was used for methyl-β-D-glucose, whereas water was modeled by the TIP3P potential.[33] MD simulations were carried out using the program AMBER9[40] periodic boundary conditions and a truncated octahedron simulation cell. Electrostatic interactions are calculated with the Ewald particle-mesh method.[41] We used a cutoff of 8 Å for the real-space direct sum part of the Ewald sum and for the van der Waals interactions. The Berendsen coupling algorithm[42] for temperature (298 K and τ = 0.1 ps) and pressure (1 atm and τ = 1 ps) was applied to the simulation system.

Two different sets of simulations were carried out for methyl-β-D-glucose in water, each comprising of four stages. In both sets of simulations the first three stages are the same, but the fourth and last stage (which is the "production run") is different. The methyl-β-D-glucose is fully restrained in the first set while in the second set only the O6−C6−C5−O5 dihedral angle (Figure 1) in the glucose ring is restrained. This dihedral angle was set to either gg or gt conformation as appropriate to each simulation. This is accomplished through the use of "RESTRAINTMASK" keyword, which enables the specification of the entire methyl-β-D-glucose molecule (set 1) or just the atoms involved in the torsion angle (set 2) to be restrained throughout the simulation. The three stages common to both sets of simulations are described as follows. In the first stage, the system was gas-phase optimized, using steepest descent, with all atoms restrained except the hydrogen atoms of water and the monosaccharide. In the second stage, the system was relaxed as a *NVT* ensemble for 800 ps with all the atoms except the hydrogens of water restrained. The time integration step was set to 2 fs. In the third stage, the system was relaxed as a *NpT* ensemble for 200 ps, with all atoms except the hydrogens of water restrained. The time integration step was set to 0.5 fs. In the fourth and final stage, once equilibrated, the system was simulated as a *NpT* ensemble for 200 ps, with the time integration step being set to 0.5 fs. This strategy of preparing the production run was followed for both the first and second set of simulations. Pereira et al.[21] have suggested that simulations should be run for 30−40 ns in order to adequately sample the conformational space accessible to solvated disaccharides, while Kirschner and Woods[22] were able to reproduce the gg/gt/tg rotamer populations for methyl-α-D-glucose using a 50 ns simulation. We have only conducted MD simulations for a shorter time period of 400 ps but believe that this is sufficient as we separately simulate the gg and gt rotamers and our aim was to sample the various monosaccharide−water complexes rather than to determine their populations or the dynamics of the conformational changes.

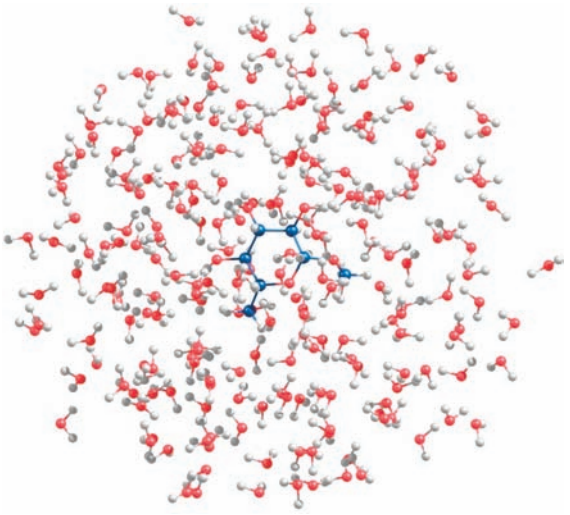


Figure 2. Representative snapshot of the explicitly hydrated methyl-β-D-glucose gg conformer with 150 water molecules.

Experimental Measurements. Methyl-β-D-glucose was purchased from Alfa Aesar (Ward Hill MA, USA) and used without further purification. Solution samples of the monosaccharide were prepared by dissolving the solid material in deionized water at a concentration of 200 mg/mL, followed by pipetting the sample into a quartz microfluorescence cell. The pH of each sample was measured to be 6.8 ± 0.2 using a digital pH meter (Hannah Instruments). Raman and ROA spectra were simultaneously recorded using a BioTools ChiralRaman spectrometer (BioTools Inc., Jupiter FL, USA) operated via Critical Link LLC software. The instrument was set up in the backscattering geometry using a Millenium Pro Nd/VO_4 laser (SpectraPhysics, UK) with an excitation wavelength of 532 nm, laser power of 700 mW at the sample, spectral resolution of 7 cm^{-1}, and spectral acquisition time of 70 min. The raw Raman and ROA spectra were plotted using Origin 8.1 Pro software (OriginLab, Northampton, MA, USA) and are presented in Figure 4.

RESULTS

Molecular Dynamics Simulations. Snapshots of the simulation were taken every 25 ps of the simulation. The snapshots contain the methyl-β-D-glucose molecule with TIP3P water within a cutoff distance of 8 Å (an example snapshot is shown in Figure 2). In total there were 16 snapshots from the final production run of each methyl-β-D-glucose conformation (gg and gt). The coordinates from these snapshots were used to generate the calculated ROA spectra.

Raman and ROA Spectra. Figure 3 shows the calculated Raman and ROA spectra of the gg and gt conformers in the gas-phase, i.e. in the absence of either explicit water molecules or the PCM. The spectra for the individual conformers shown in panels A and B are presented in proportion to their relative populations (60% gg: 40% gt). Significant differences are apparent between

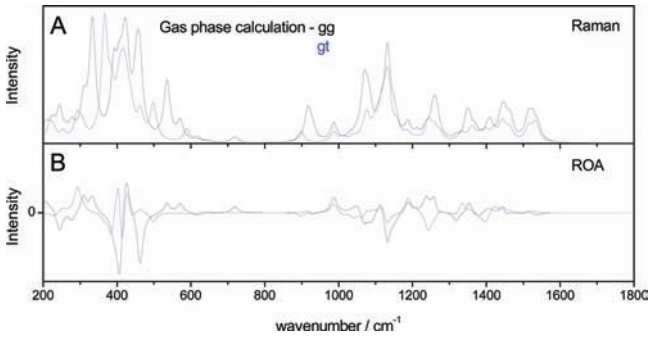


Figure 3. Calculated Raman and ROA spectra for the gg (panels A and B, black trace) and gt (panels A and B, blue trace) conformers in the gas phase of methyl-β-D-glucose.

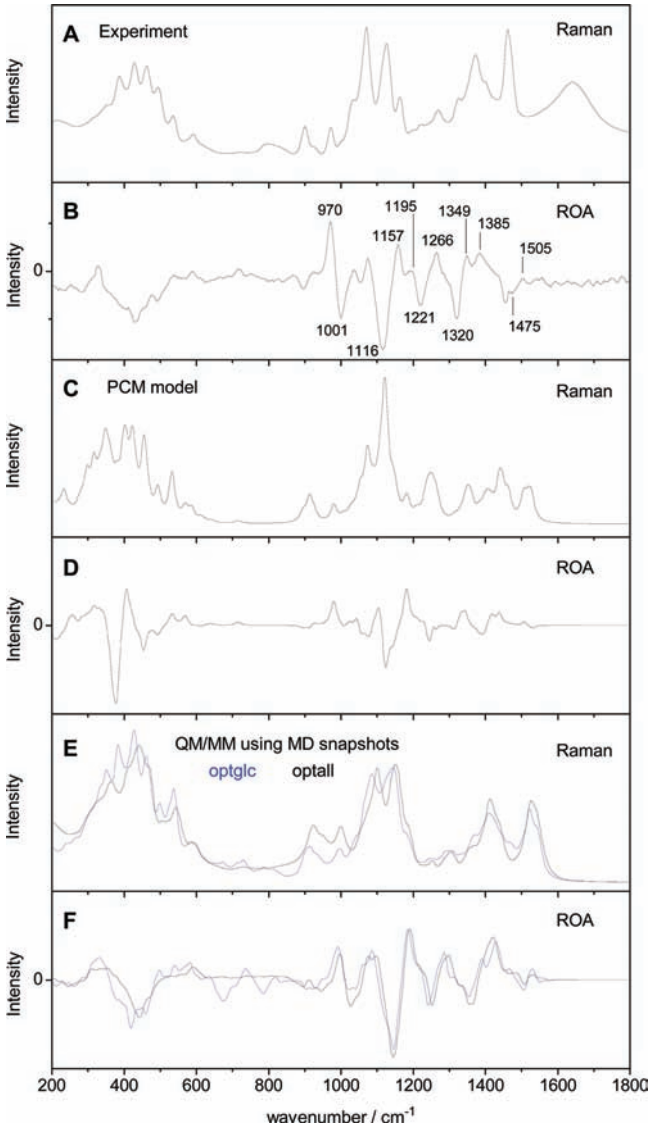


Figure 4. Experimental Raman and ROA spectra of methyl-β-D-glucose (black) (panels A and B), calculated Raman and ROA spectra using PCM (panels C and D), and optglc (blue) and optall (black) calculated Raman and ROA spectra with inclusion of MD simulations of the aqueous environment (panels E and F). All calculated spectra are 60:40 weighted averages of the gg and gt conformers.

the gas-phase spectra of these two rotamers, such as the inversion of signs of ROA features at ~400 and 460 cm^{-1} and in the range 1200−1350 cm^{-1}. Both of these ROA spectra, however, are dominated by sharp and complex band patterns below 500 cm^{-1}.

We present the measured Raman and ROA spectra of methyl-β-D-glucose in Figure 4 panels A and B, and in panels C and D are shown the Raman and ROA spectra for the weighted 60:40 population average of the gg:gt rotamers using the PCM. It is evident that the intense band envelope observed below 500 cm^{-1} for the gas phase calculations becomes simplified to some degree upon introduction of the PCM for the weighted average in the lower panel, though the sharp −ve/+ve/−ve triplet feature in this region still dominates the calculated spectrum, unlike the case of the experimental ROA spectrum. Other features in this ROA spectrum calculated using the PCM are similar to those calculated for the 60:40 weighted average of the gg and gt rotamers in the gas phase (shown in the Supporting Information), indicating that this model of the solvent does not have a significant effect on most of the vibrational modes of the monosaccharide. Above 500 cm^{-1}, we can observe that incorporation of the PCM model produces a good level of agreement with the experimental Raman and ROA spectra. However, significant differences are still apparent, notably the failure by the PCM model to predict the −ve component of the +ve/−ve couplet observed in the experimental ROA spectrum near 1000 cm^{-1}. Although the presence of other ROA spectral features above 1000 cm^{-1} are accurately predicted using the PCM model, the exact shape of the −ve band above 1100 cm^{-1} is a little different to that measured experimentally and shows an additional shoulder. The relative intensities of the ROA features above 1200 cm^{-1} also appear to be underrepresented in the PCM model.

The effect of the molecular dynamics simulations on ROA predictions is also shown in Figure 4, which presents the spectra computed for simulation snapshots from 25 to 400 ps. For the Raman and ROA spectra of the optglc simulation shown in panels E and F of Figure 4, respectively, for each snapshot the structure of methyl-β-D-glucose was separately optimized while those of all water molecules were kept frozen, while for the optall simulation for each snapshot both the structure of methyl-β-D-glucose and the positions of all water molecules were optimized and allowed to fully relax. These two predicted spectra are obviously not only an improvement on both calculations utilizing no explicit water molecules (gas phase or PCM), but are both very similar to the experimental spectrum.

The positions and intensities of the calculated Raman and ROA bands, along with their vibrational assignments (in the gas phase), are listed in Table 1. It is evident that most of these bands arise from complex vibrational modes that are coupled across several functional groups. Despite this complexity, several bands are sensitive to the type of conformer and future studies on glycans and polysaccharides should reveal whether any of these Raman or ROA bands are sensitive to the formation of higher order structure.

The agreement between computation and experiment is particularly impressive in the range 1050−1300 cm^{-1} with both of our optglc and optall models correctly predicting the number of bands, their signs, band shapes, and even relative intensities. Both of these calculations utilizing molecular dynamics simulations display the main −ve/+ve couplet, at ~1130/1160 cm^{-1}, though both overestimate the intensity of this feature by ~50% with respect to the other ROA bands in the experimental

dx.doi.org/10.1021/ja110825z | J. Am. Chem. Soc. 2011, 133, 4991–4997

Table 1. Calculated Raman and ROA Bands for Methyl-β-D-glucose in the Gas Phase

	intensity (arb. units)			
wavenumber (cm^{-1})	Raman	ROA	conformer	assignment
207	18	18	gg	CH_3 torsion, OH twist
243	26	−27	gg	OH twist
294	30	50	gt	OH twist, O(6)H twist
307	40	49	gg, gt	OH twist, ring def
332	72	31	gg	OH twist
387	78	−71	gg, gt	OH twist, ring def
410	100	−100	gg, gt	OH twist, O(6)H twist
427	99	86	gg, gt	OH twist, O(6)H twist, ring def
462	73	−83	gg, gt	OH twist, O(6)H twist, CH_2 twist, ring def
497	29	−20	gt	O(6)H twist, ring def
535	35	16	gg	OH twist, ring def
571	15	16	gg	O(6)H twist, CH_2 twist, ring def
615	6	−5	gg, gt	CH_2 twist, C(1)−O(1)−CH_3 bend, ring def
719	4	11	gg	C(5)−C(6)−O(6) bend, ring def
897	11	−5	gt	CH_2 twist, C(5)−C(6) stretch, ring def
916	21	1	gg	C(5)−C(6) stretch, C(6)−O(6) stretch, H−C(5)-C(6) bend
929	15	−2	gg	C−C stretch, C(5)-0(5) stretch, C(5)−C(6)−H bend, O(6)H twist
987	17	42	gg, gt	O(1)-CH_3 stretch, C−C stretch, C−O−H bend, C−C−H bend, OH twist
1049	17	6	gg	C(5)−O(5) stretch, C(6)−O(6) stretch, C−C stretch, C−C−H bend, C−O−H bend
1068	48	−22	gg	C(1)−O(5) stretch, O(1)-CH_3 stretch, C(5)−C(6) stretch, C−C stretch
1113	90	22	gg, gt	C(6)−O(6) stretch, C−OH stretch, C(5)−C(6) stretch, C(5)−O(5) stretch, C(1)−O(1) stretch, O−CH_3 stretch, C(5)−C(6) stretch, C−C stretch, H−C(6)-O(6) bend, C−O−H bend
1188	19	44	gg, gt	CH_3 def, O(1)-CH_3 stretch, C(1)−O(1) stretch, C−C stretch, C−C−H bend
1245	27	−7	gt	CH_2 def, C(6)−O(6)−H bend, H−C(6)-O(6) bend, C−O−H bend
1258	36	14	gg	C−C−H bend, C−O−H bend, CH_2 def
1317	9	−16	gg, gt	ring H torsions, H−C−OH bend, H−C−O bend, C−O−H bend, C−C−H bend
1336	15	15	gt	ring H torsions, CH_2 def, C−C−H bend, H−C(6)-O(6) bend
1353	25	23	gg	ring H torsions, H−C(1)-O(1) bend, H−C−O(5) bend, C−C−H bend, C−O−H bend
1384	14	−12	gg, gt	ring H torsions, H−C−OH bend, H−C−O(5) bend, C−C−H bend, C−O−H bend
1421	18	15	gt	ring H torsions, H−C−O(5) bend, H−C(5)-C(6) bend, C(6)−O(6)−H bend, C−C−H bend
1443	33	21	gg, gt	ring H torsions, H−C−OH bend, C−C−H bend, C(6)−O(6)−H bend, C−O−H bend
1516	28	4	gg, gt	CH_3 def
1541	20	−4	gg, gt	CH_3 def

spectrum. This region contains not only many of the most intense ROA bands but also several sharp couplets and so can serve as a "fingerprint" of carbohydrate identity and conformation, analogous to the way in which amide III modes in protein ROA spectra can be used as a fingerprint of protein secondary structure. These fingerprint bands include the +ve/−ve couplet located at 970/1001 cm^{-1}, the two weak positive bands between 1000 and 1100 cm^{-1}, a small +ve shoulder at ∼1195 cm^{-1}, and a −ve/+ve couplet at ∼1221/1266 cm^{-1}. All bands are well modeled by both the optglc and optall calculations. The high degree of correspondence between experiment and modeling shown for these fingerprint bands in Figure 4 demonstrates that ab initio methods are now capable of correctly predicting the solution structure of monosaccharides.

In the region above 1300 cm^{-1}, the optglc calculation is in closer agreement to the fine details of the experimental ROA spectrum than is the optall calculation, as it correctly presents both the small −ve/+ve couplet at ∼1320/1349 cm^{-1} and the small +ve peak at ∼1385 cm^{-1}. Both of these MD inclusive predictions also appear to correctly predict a small −ve/+ve feature located at ∼1475/1505 cm^{-1} in the experimental spectrum. However, this feature is only barely above the noise level in the experimental spectrum. ROA is a weak effect making measurements demanding and one must be careful when considering weak spectral features due to the possible presence of both experimental noise and minor distortions, but the presence of this weak feature in both the optglc and optall predictions, with the optall calculation also yielding the correct intensity, gives us some confidence that the couplet in the experimental spectrum at ∼1475/1505 cm^{-1} is real. We are not aware of this level of agreement in modeling of weak ROA bands in previous studies for any molecule.

The most obvious differences between the two MD-incorporated spectral predictions occur within the region of 500−900 cm^{-1}. In the experimental spectrum there are a few weak features with the optglc prediction showing the same features

dx.doi.org/10.1021/ja110825z | J. Am. Chem. Soc. 2011, 133, 4991–4997

but with much greater intensities while the optall prediction displays almost no ROA within this region. Overall, between 500 and 900 cm^{-1}, where many of the bands originate from vibrations involving the $C(6)H_2-OH$ group, the optall calculation produces a better match than the optglc calculation to the weak features displayed in the experimental spectrum. The reason for this difference in performance of the two MD-incorporated calculations with respect to different regions of the ROA spectrum is not known, though we note that the gas-phase and PCM models perform worse in this region.

Finally, the most significant illustration of the superior performance of the MD simulation of the aqueous environment and its interactions with the monosaccharide is provided by the range of 200−500 cm^{-1}. Whereas the PCM calculations poorly modeled this low wavenumber region, both of the MD inclusive calculations perform far better with the optall calculation yielding an excellent reproduction of the −ve band at ∼430 cm^{-1} and the weaker −ve band at ∼460 cm^{-1}. These low wavenumber vibrational modes and how they are moderated by molecular interactions with solvent water molecules are only correctly predicted by incorporating MD simulations of the solvent, with the optall calculations (in which the positions of all water molecules are optimized) being the most accurate.

CONCLUSION

Luber and Reiher[15] showed that explicit hydration by a small number of water molecules led to large changes in the ROA spectra of a the monosaccharide 1,6-anhydro-β-D-glucopyranose and also suggested the incorporation of MD simulations to further improve agreement with experiment. We have extended this approach to the full modeling of the aqueous environment and obtained an excellent level of agreement with experimental ROA spectra for methyl-β-D-glucose. The incorporation of full MD simulations of the aqueous environment with ROA QM/MM computations using extracted snapshots reported here has revealed the degree of interaction of the aqueous environment with individual methyl-β-D-glucose molecules and their importance in mediating the conformational dynamics of the monosaccharide. Electrostatic interactions between the QM electron density of methyl-β-D-glucose and the MM charges of the surrounding water are included in our treatment.

We emphasize that inclusion of the many explicit water molecules in the QM/MM calculations themselves adds a negligible computational cost. CPU times for the computations to obtain the vibrational frequencies and the polarizability derivatives to calculate the Raman and ROA tensors do not increase when the MM explicit water molecules are included. There is also no increase in CPU time for each individual step of the geometry optimization upon inclusion of the MM waters, though the number of steps does increase, especially for the optall calculations. There is a great need for techniques capable of resolving the structures of carbohydrates in solution as NMR is not typically applicable due to difficulties in dealing with the low sequence complexity and poor rotational dynamics of glycans. We[43] and other researchers[44,45] have previously shown that stereochemically sensitive ROA spectra can be measured for carbohydrates of any size and that ab initio calculations are the most effective means to understanding the structural information contained. However, the inability to accurately model the exquisite sensitivity of ROA features to hydration effects using PCM or limited explicit hydration had prevented realization of the potential of ROA. We have now demonstrated that this problem can be solved through incorporation of a full MD simulation of the aqueous environment, opening the door for a new level of understanding of glycobiology at the molecular level.

ASSOCIATED CONTENT

Supporting Information. Calculated Raman and ROA spectra for the 60:40 weighted average of the gg and gt rotamers in the gas phase, equations, and complete refs 25 and 40. This material is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author
e.blanch@manchester.ac.uk

ACKNOWLEDGMENT

We are grateful to Dr. Steve Liem for his help with the simulations.

REFERENCES

(1) Barron, L. D.; Bogaard, M. P.; Buckingham, A. D. *J. Am. Chem. Soc.* **1973**, *95*, 603.
(2) Nafie, L. A. *Annu. Rev. Phys. Chem.* **1997**, *48*, 357.
(3) Hug, W. In *Handbook of Vibrational Spectroscopy*; Chalmers, J.M., Griffiths, P. R., Eds.; John Wiley & Sons: Chichester, Great Britain, 2002; Vol. 1, p 745.
(4) Ruud, K.; Thorvaldsen, A. J. *Chirality* **2009**, *21*, E54–E67.
(5) Pecul, M. *Chirality* **2009**, *21*, E98–E104.
(6) Pecul, M. *Chem. Phys. Lett.* **2006**, *427*, 166–176.
(7) Jalkanen, K. J.; Degtyarenko, I. M.; Neiminen, R. M.; Cao, X.; Nafie, L. A.; Zhu, F.; Barron, L. *Theor. Chem. Acc.* **2008**, *119*, 191–210.
(8) Herrmann, C.; Ruud, K.; Reiher, M. *ChemPhysChem* **2006**, *7*, 2189–2196.
(9) Mukhopadhyay, P.; Zuber, G.; Beratan, D. N. *Biophys. J.* **2008**, *95*, 5574–5586.
(10) Kapitan, J.; Fujiang, Z.; Hecht, L.; Gardiner, J.; Seebach, D.; Barron, L. *Angew. Chem. Int. Ed.* **2008**, *47*, 6392–6394.
(11) Luber, S.; Reiher, M. *J. Phys. Chem. A* **2010**, *114*, 1057–1063.
(12) Jacob, C. R.; Luber, S.; Reiher, M. *Chem.—Eur. J.* **2009**, *15*, 13491–13508.
(13) Pecul, M.; Rizzo, A.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 11008–11016.
(14) Macleod, N. A.; Johannessen, C.; Hecht, L.; Barron, L. D.; Simons, J. P. *Int. J. Mass Spectrom.* **2006**, *253*, 193–200.
(15) Luber, S.; Reiher, M. *J. Phys. Chem. A* **2009**, *113*, 8268–8277.
(16) Kapitan, J.; Baumruk, V.; Kopecky, V. J.; Bour, P. *J. Phys. Chem. A* **2006**, *110*, 4689–4696.
(17) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.
(18) Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.
(19) Marchessault, R. H.; Perez, S. *Biopolymers* **1979**, 18.
(20) Gonzalez-Outeirino, J.; Kirschner, K. N.; Thobhani, S.; Woods, R. J. *Can. J. Chem.* **2006**, *84*, 569–579.
(21) Pereira, C. S.; Kony, D.; Baron, R.; Muller, M.; van Gunsteren, W. F.; Hunenberger, P. H. *Biophys. J.* **2006**, *90*, 4337–4344.
(22) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10541–10545.
(23) Barron, L. D.; Buckingham, A. D. *Mol. Phys.* **1971**, *20*, 1111.
(24) Barron, L.; Buckingham, A. D. *Chem. Phys. Lett.* **2010**, *492*, 199–213.

(25) Frisch, M. J. et al. *Gaussian Development Version H.07+*; Gaussian, Inc.: Wallingford, CT.
(26) Henderson, T. M.; Izmaylov, A. F.; Scalmani, G.; Scuseria, G. E. *J. Chem. Phys.* **2009**, *131*, 044108.
(27) Zuber, G.; Hug, W. *J. Phys. Chem. A* **2004**, *108*, 2108–2119.
(28) Reiher, M.; Liégeois, V.; Ruud, K. *J. Phys. Chem. A* **2005**, *109*, 7567–7574.
(29) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (Theochem)* **1999**, *462*, 1–21.
(30) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A., Jr.; Morokuma, K.; Frisch, M. J. J. *Chem. Theory Comput.* **2006**, *2*, 815–826.
(31) Bakowies, D.; Thiel, W. J. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(33) Jorgensen, W. L.; Jenson, C. *J. Comput. Chem.* **1998**, *19*, 1179–1186.
(34) London, F. *J. Phys. Radium* **1937**, *8*, 397–409.
(35) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789–807.
(36) Barron, L. *Molecular Light Scattering and Optical Activity*, 2nd ed.; Cambridge University Press: Cambridge, Great Britain, 2004.
(37) Nafie, L. A.; Che, D. In *Modern Nonlinear Optics, Part 3*; Evans, M., Kielich, S., Eds.; Wiley: New York, 1994; Vol. 85, p 105−149.
(38) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct. Bioinf.* **2006**, *65*, 712–725.
(39) Wang, J.; Cieplak, P.; Kollman, P. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(40) Case, D. A. et al. *AMBER 9*; University of California: San Francisco, 2006.
(41) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(42) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(43) Yaffe, N. R.; Almond, A.; Blanch, E. W. *J. Am. Chem. Soc.* **2010**, *132*, 10654–10655.
(44) Rudd, T. R.; Hussain, R.; Siligardi, G.; Yates, E. A. *Chem. Commun.* **2010**, *46*, 4124–4126.
(45) Johannessen, C.; Pendrill, , R.; Hecht, , L.; Widmalm, G.; Barron, L. D. Raman Optical Activity of Biomacromolecules: Structural Analysis of Sugar Moieties in Glycoproteins. *AIP Conference Proceedings XXII International Conference on Raman Spectroscopy*, 2010; paper no. 1267, pp 104−105.