

Cite this: DOI: 00.0000/xxxxxxxxxx

Simulation of Raman and Raman optical activity of saccharides in solution[†]

Vladimír Palivec,^{a‡} Vladimír Kopecký Jr.,^b Pavel Jungwirth,^a Petr Bouř,^a Jakub Kaminský,^{ac*} and Hector Martinez-Seara^{a*}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Structural studies of sugars in solutions are challenging for most of the traditional analytical techniques. Raman and Raman optical activity (ROA) spectroscopies were found extremely convenient for this purpose. However, Raman and ROA spectra of saccharides are challenging to interpret and model due to the saccharides' flexibility and polarity. In this study, we present an optimized computational protocol that enables the simulation of the spectra efficiently. Our protocol, which results in an excellent agreement with experiments, combines molecular dynamics and density functional theory calculations. It further uses a smart optimization procedure and a novel adaptable scaling function. The numerical stability and accuracy of individual computational steps are evaluated by comparing simulated and experimental spectra of D-glucose, D-glucuronic acid, *N*-acetyl-D-glucosamine, methyl β-D-glucopyranoside, methyl β-D-glucuronide, and methyl β-N-acetyl-D-glucosaminide. Overall, our Raman and ROA simulation protocol allows to routinely and reliably calculate the spectra of small saccharides and opens the door to advanced applications, such as complete 3-dimensional structural determination by direct interpretation of the experimental spectra.

1 Introduction

Carbohydrates play a pivotal role in numerous functions in living organisms. They provide mechanical support such as in the case of cellulose. They form materials for energy storage, such as starch or glycogen. Alternatively, sugars such as ribose and deoxyribose are part of the genetic material. Other sugars play a significant role in the structure and function of the glycocalyx^{1,2}, which is a 10-50 nm thick layer found on the extracellular side of the plasma membrane of several eukaryotic cells. This layer and its components rich in sugars make possible cell-cell recognition, filtration of the surrounding biomolecules, and maintain the cell integrity². Its main components are proteins and their glycosylated variants, such as proteoglycans, and glycosaminoglycans¹.

Better understanding the function of saccharides requires a method for the determination of their flexible structures in solution. However, for sugars, such a task is particularly difficult.

Saccharide molecules, especially when functionalized (e.g., glycans), are often hard to crystallize, and they provide complex NMR spectra with broad and overlapping bands³. Saccharides also usually lack UV-vis absorbing chromophores. Therefore, they cannot be easily investigated by standard UV-vis absorption or electronic circular dichroism spectroscopies⁴. There are recent advancements in vacuum-ultraviolet electronic circular dichroism spectroscopy, but the technique is not yet able to routinely analyze and assign structural features of saccharides^{5,6}. Therefore, vibrational spectroscopies, in particular Raman and Raman optical activity (ROA), have been suggested as optimal for saccharides. Both methods, especially the chiral technique ROA, are sensitive to the structure of saccharides in solutions. The ROA technique records the small but significant differences in the scattering of the right- and left-circularly polarized light, and thus it increases the information about the structure in the spectrum, compared to standard Raman scattering⁷. This is because ROA bands can be both positive and negative, which helps in their assignment, and the spectra are not only more sensitive to absolute chirality, but also to minor structure variations. The 200-1800 cm⁻¹ spectral region commonly accessible by Raman and ROA spectroscopy rich in structural information is significantly wider and easier to obtain compared to those accessible to other vibrational techniques such as infrared (IR) and vibrational circular dichroism (VCD) spectroscopies.

^a Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo náměstí 542/2, CZ166 10, Prague 6, Czech Republic; E-mail: jakub.kaminsky@uochb.cas.cz; hseara@gmail.com

^b Institute of Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 5, CZ121 16, Prague 2, Czech Republic

^c Gilead Sciences & IOCB Research Center

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 00.0000/00000000.

‡ PhD. Student of University of Chemistry and Technology Prague

Previously, Raman and ROA spectroscopies have been used, for example, to study the presence of secondary and tertiary structures of glycosaminoglycans in solution such as hyaluronan⁸ and heparan sulfate^{9,10}. In principle, this can be done by simply locating spectral changes as a function of the length of several hyaluronan moieties. However, assigning atomistic features directly from the experimental spectra is difficult. To a good approximation, measured Raman and ROA spectral shapes can be considered as weighted sums of the individual spectra of all molecular conformations adopted by the sugar. For this reason, previous Raman and ROA studies pointed out that the power of these spectroscopic techniques is significantly enhanced when coupled with computer simulations^{11–14}. Simulations not only provide the Raman and ROA spectra of a particular configuration but also allow to predict the spectra of the real conformer mixture.

Early computational studies were significantly hampered by the excessive computer time needed for spectral simulations and interpretation^{15,16}. Nevertheless, it was quickly understood that for saccharides, with many semi-free rotating hydroxyl groups, single conformer models (especially if considered only in a vacuum) cannot fully reproduce experimental data. Also, continuum (dielectric) solvent models provided only limited accuracy¹³. So far, the best practical way of simulating the spectrum of a particular sugar moiety requires a combination of molecular dynamics and quantum-chemical calculations (typically density functional theory (DFT)), including at least some explicit solvent molecules into the quantum chemical computations. Such procedure was used for example to study methyl β -D-glucopyranoside¹³, where hundreds of explicit water molecules were employed within the monosaccharide solvation model. The same approach was recently used to study mannobiose disaccharides, where the MD simulations, together with NMR spectroscopy, and DFT calculations were combined to produce Raman/ROA spectra¹⁷. Another work focused on β -xylose showed that at least a 1 nm thick layer of water molecules is needed to model correctly the bulk aqueous environment¹⁸. Other studies using explicit waters focused on subtle spectral modifications between D-glucose and D-galactose differing only in one chiral center¹⁹ and on a study of D-glucuronic acid and N-acetyl-D-glucosamine²⁰. One of the largest system studied is agarose hexamer²¹. However, in the paper only continuum solvation without explicit water molecules is used, which was shown to be insufficient to properly capture Raman/ROA spectral features¹³.

All these studies clearly hint to the necessity of including explicit waters when computing Raman and ROA spectra. Despite this important advancements, these studies neglected the anomeric equilibrium in sugars. This reason alone can explain several inconsistencies in their results like a limited agreement of the calculated spectra with experiment. Only a few other studies, however, correctly accounted for the anomeric equilibrium when using explicit waters. One example is the study of glucose/mannose mixtures,¹² where a hybrid solvation approach with the first solvation shell described quantum mechanically, together with the polarizable continuum accounting for the long range solvation effects, was used resulting in a reasonable agreement with experiments.

As of today, several protocols to model Raman and ROA spectra have been proposed with different degrees of success^{12,13}. In the present study, we compare various approaches to model Raman and ROA spectra of reducing monosaccharides and their methyl glycosides (methyl β -D-glucopyranoside [m β -glc], methyl β -D-glucuronide [m β -glcA], and methyl β -N-acetyl-D-glucosaminide [m β -glcNAc], D-glucose [glc], D-glucuronic acid [glcA], and N-acetyl-D-glucosamine [glcNAc]) and compare them to original experimental data. As a result of our comparison, we propose a quantitative and computationally efficient protocol. This protocol builds on previous works which we further optimize and improve in this current study. We also shed light into the need of taking into account the anomeration equilibrium when studying reducing sugars.

Finally, we consider the present study on monosaccharides as a first step allowing us to validate a universal methodology for computing Raman and ROA spectra, usable also for larger and more complicated systems in the future.

2 Methods

2.1 Target Monosaccharides

In this work, in order to develop a Raman and ROA modeling protocol, we benchmarked several monosaccharides. Monosaccharides are the building blocks of more complex sugars, which our protocol also aims to cover in future studies. Here, we chose the two monosaccharides that constitute hyaluronan as the major glycocalyx component. These two sugars, N-acetyl-D-glucosamine (glcNAc) and D-glucuronic acid (glcA), were complemented by D-glucose (glc), structure of which is well known and, therefore, can be used to benchmark several aspects of the methodology. In aqueous solutions all three monosaccharides exist as a mixture of their α - and β -anomeric forms. To have simpler reference systems devoid of the anomeric equilibrium, we also investigated non-reducing derivatives, i.e., the methyl β -D-glucopyranoside (m β -glc), methyl β -D-glucuronide (m β -glcA), and methyl β -N-acetyl-D-glucosaminide (m β -glcNAc). Only the pyranoses forms of the sugars were considered, as significant occurrence of other forms is not likely in aqueous solutions²². Structures of all studied molecules are depicted in Figure 1.

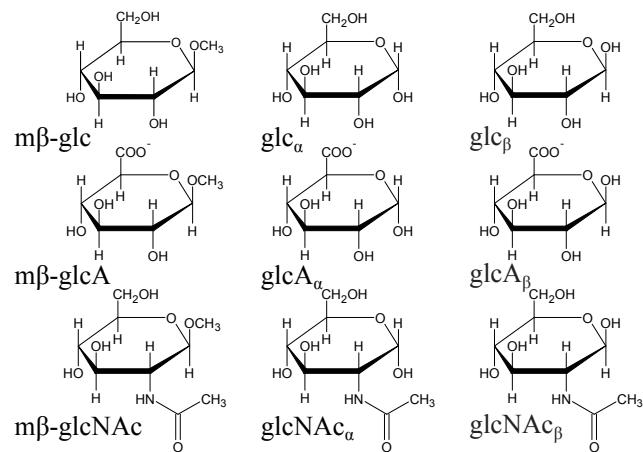


Fig. 1 Herein investigated monosaccharides.

2.2 Experimental Section

All studied reducible monosaccharides glc, glcA, and glcNAc, as well as the methyl glycosides m β -GlcA and m β -GlcNAc were purchased from Carbosynth. The concentration of sugars was ~750–1000 mmol. Backscattered Raman and scattered circular polarization (SCP) ROA spectra were recorded on a ChiralRAMAN-2X (Biotools Inc.) spectrometer equipped with a 532 nm green laser. We first removed sample impurities causing unwanted fluorescence using active carbon. Then we quenched the residual fluorescence by leaving the sample in the laser beam for an hour before measurement. Spectra were accumulated in a fused silica cell (3 mm optical path, 50 μ L sample volume) overnight (~12 h). Laser power was \approx 700 mW. The solvent signal was subtracted from the Raman spectra. Because this did not always eliminate a broad background fluorescence from residual impurities, a minor polynomial baseline correction was done manually in some cases. The spectra were finally normalized to accumulation time and we performed a relative intensity correction²³. Spectra of m β -glc were taken from Cheeseman *et al.*¹³.

2.3 Simulation of Raman and ROA Spectra

Generation of the simulated spectra consisted of four stages: 1) Generation of an ensemble of relevant structures including solvent positions using molecular dynamics (MD) simulations. 2) Calculation of the spectra using quantum chemical methods for the selected structures. 3) Minor adjustments of the calculated spectra, in particular scaling of calculated vibrational frequencies. 4) Weighting the spectra by conformer weights. A general scheme of the whole procedure is shown in Figure 2.

2.3.1 Ensemble Generation Using Molecular Dynamics

An ensemble of structures for each monosaccharide in water was generated by performing molecular dynamics (MD) simulations using the program Gromacs-2016.4²⁴. For each sugar the simulated system contained a monosaccharide, initially in the 4C_1 chair conformation, in a cubic box of $4 \times 4 \times 4$ nm³ and \sim 2150 water molecules. For the D-glucuronic acid and its methyl derivative, we kept the system electroneutral by adding a single Na $^+$ ion²⁵. In MD simulations we used the glycam-6h²⁶ and OPC3²⁷ force fields for the sugar and water, respectively. Long range electrostatic interactions were accounted for using PME²⁸ with a 1 nm cut-off for the real part. Van der Waals interactions were treated with a cut-off of 1 nm with long range dispersion corrections for energy and pressure²⁹. Both Coulomb and Van der Waals potentials were shifted to zero at the cut-off. Neighbour search was done with the Verlet list. All hydrogen containing bonds were constrained using the LINCS algorithm³⁰. Each system was first minimized using the steepest descent minimization, followed by a 100 ps heating up within a NVT protocol to 310 K, and finally equilibrated for 100 ps in a NpT ensemble (1 bar). Afterwards, we ran a production NpT simulation for 500 ns using a 2 fs time step. A Noose-Hoover thermostat^{31,32} with a coupling constant of 1 ps controlled the temperature. A Parrinello-Rahman³³ barostat using a time constant of 5 ps and compressibility of 4.5e⁻⁵ bar⁻¹ controlled the pressure. 100 uncorrelated snapshots, taken every 5 ns, were saved for future analysis.

To simplify the subsequent *ab initio* calculations, as a first approximation, we discarded around 7.5 % of all gathered snapshots where Na $^+$ ion distance to any atom of the monosaccharide was smaller than 5 Å.

2.3.2 Calculation of Raman and ROA Spectra

The Raman and ROA frequencies and intensities were calculated using the Gaussian16³⁴ suite of programs. We used the B3LYP hybrid functional, in combination with the 6-311++G** basis set, as this combination provided in a similar context excellent results in terms of quality and computer efficiency previously¹². Various solvating schemes were tried, as summarized in Table 1. For every selected MD snapshot (j), we took the monosaccharide coordinates with or without surrounding water molecules. Then, we optimized the geometry of the cropped system using various optimization schemes. The CPCM, MM, and hybrid simulation protocols employ ten optimization steps. In case of MM approach, the Cartesian coordinates were used during the optimization as the internal coordinates optimization is \sim 8 times slower. MM_w and hybrid_w optimization approach freezes cropped water molecules and the sugar molecule is optimized until gradient criteria are reached. With hybrid_fo approach we optimize whole system fully until gradient criteria are reached. Finally, we have calculated the vibrational frequencies v_{ij}^* , and corresponding Raman, $I_{Raman,ij}^{sim}$, and ROA, $I_{ROA,ij}^{sim}$, intensities, where i refers to a calculated vibrational mode of a snapshot j . Note that in this work we use interchangeably the terms vibrational frequencies v_{ij}^* and vibrational wavenumbers \tilde{v}_{ij}^* which are mutually connected as $v_{ij}^* = c \cdot \tilde{v}_{ij}^*$, where c stands for the speed of light. We used the harmonic approximation and the far from resonance coupled-perturbed Kohn-Sham theory (CPKS)^{35,36} with the 532 nm excitation frequency corresponding to the green laser used in experiment. In this work, the Raman and ROA spectra correspond to the scattered circular polarization (SCP) modulation scheme and backscattering (180°) detection.

All methods listed in Table 1 describe the monosaccharides at the DFT level of theory. When performing QM/MM calculations, we employ the ONIOM method^{37,38} with electrostatic embedding. For the force field description of the water and sugars, we used the TIP3P³⁹ and glycam-6h²⁶ charges and Lennard-Jones parameters, respectively. For other force field parameters, we use the default values as in Gaussian16, however, note that their contributions cancel in this approach. Other variations of the simulation protocols are described in Tables 1, 2, and 3 in ESI†. They address different optimization procedures, DFT functionals, basis sets, usage of QM water molecules, and convergence issues in the Raman and ROA spectra calculations.

2.3.3 Adjustments of the Calculated Spectra

The calculated vibrational frequencies, \tilde{v}_{ij}^* , are usually slightly overestimated when using hybrid DFT approaches (see section 4 in ESI†)^{35,40–45}. Also, the use of a simple harmonic approximation can lead to further deviations of the frequencies^{35,41}. To account for these deviations, we adjusted (“scaled”) the calculated vibrational frequencies. In this work, we used the following scaling function $\phi(\tilde{v}_{ij}^*)$.

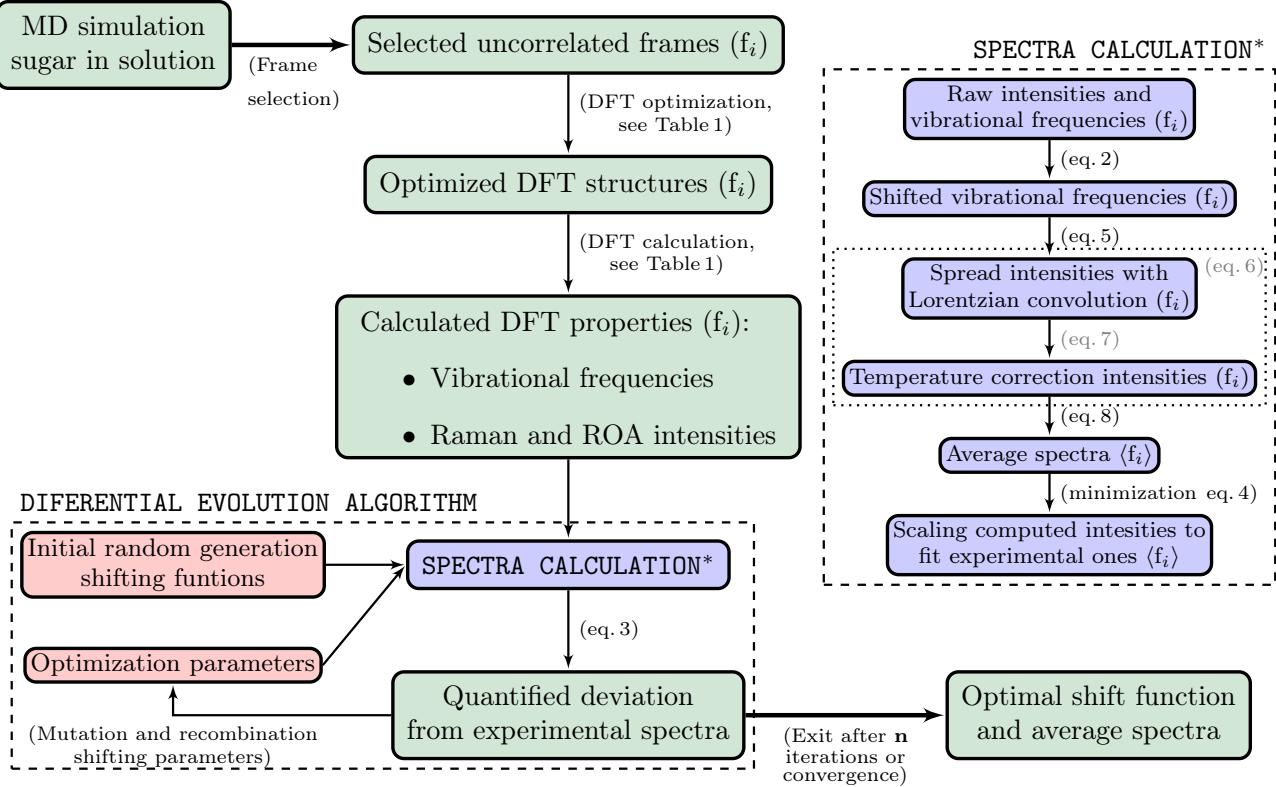


Fig. 2 Generation scheme of the simulated spectra, I_x^{sim} .

$$\phi(\tilde{v}_{ij}^*) = a \cdot \frac{1}{1 + e^{-\frac{\tilde{v}_{ij}^*+d}{c}}} + b \cdot \left(1 - \frac{1}{1 + e^{-\frac{\tilde{v}_{ij}^*+d}{c}}}\right) \quad (1)$$

The adjustable parameters a , b , c , and d represent high frequency region scaling factor (a), low frequency region scaling factor (b), c is the frequency threshold between the low and high frequency regions, and d ensures a smooth transition between these two regions. Applying the function for each of the calculated vibrational frequencies, \tilde{v}_{ij}^* , provides frequencies, \tilde{v}_{ij} , that are in better agreement with experiments.

$$\tilde{v}_{ij} = \phi(\tilde{v}_{ij}^*; a, b, c, d) \cdot \tilde{v}_{ij}^* \quad (2)$$

Note that when $a = 1$ and $b = 1$, we obtain the original simulated spectra, $\tilde{v}_{ij}^{\text{sim}}$. The parameters a , b , c , and d were optimized to fit best the experimental Raman and ROA spectra simultaneously by employing a cost function $\psi(a, b, c, d, \rho_{\text{Raman}}, \rho_{\text{ROA}})$

$$\psi = \chi_{\text{Raman}}(a, b, c, d, \rho_{\text{Raman}}) \cdot \chi_{\text{ROA}}(a, b, c, d, \rho_{\text{ROA}}). \quad (3)$$

Here

$$\chi_x(a, b, c, d, \rho_x) = \sum_{\tilde{v}=\tilde{v}_{\min}}^{\tilde{v}_{\max}} |\rho_x \cdot I_x^{\text{sim}}(a, b, c, d, \tilde{v}) - I_x^{\text{exp}}(\tilde{v})|, \quad (4)$$

serves as an overall absolute difference between experimental I_x^{exp} and ensemble average calculated spectra I_x^{sim} , using the parameter ρ_x to match the experimental and calculated intensities of a

given spectra. Note that all simulation results are shown with optimized value of ρ_x . The average magnitude of ROA to Raman intensities are experimentally related ($I_{\text{ROA}} \approx I_{\text{Raman}} \cdot 10^{-4}$), however, we optimize both spectral intensities separately. As a result, the ratio is not necessarily preserved during the optimization. More details about proposed scaling function, i.e. how it works, what are its exact effects, and how it compares to non-scaled spectra/simple scaling factor, can be found in Section 8 in ESI†.

In Equation 3, we use a product of errors which worked well in our case, however, it was later pointed out that using sum of the errors avoids possible selective optimization of either Raman or ROA spectrum. For the optimization procedure, we used a differential evolution algorithm using values of 0.7, [0.5,1], 100 as crossover probability, differential weight, and population size parameters. Such optimization ultimately allows for a better match of the calculated intensities with experiments, providing most of the structural information. The presented Raman and ROA spectra correspond to the optimized ones using a^{optimal} , b^{optimal} , c^{optimal} , d^{optimal} , $\rho_{\text{ROA}}^{\text{optimal}}$, and $\rho_{\text{Raman}}^{\text{optimal}}$.

2.3.4 Simulated Spectral Intensities

After scaling the frequencies, the simulated Raman and ROA spectrum for a given snapshot j is calculated as:

$$I_{x,j}^{\text{sim}}(\tilde{v}) = T(\tilde{v}) \cdot \sum_{i=1}^{N_{\text{vib}}} I_{x,ij}(\tilde{v}_{ij}) \cdot L(\tilde{v}; \tilde{v}_{ij}), \quad (5)$$

Table 1 Principal solvation protocols used in this work

Protocol*	Solvation Type
CPCM	CPCM continuum solvation opt=10 steps with harmonic restraints
MM	Preserved water molecules closer than 12 Å as explicit MM water molecules opt=10 steps, cartesian coordinates
MM_w	Preserved water molecules closer than 12 Å as explicit MM water molecules opt=water molecules frozen optimization in Cartesian coordinates sugar fully optimized
hybrid	Preserved water molecules closer than 3 Å as explicit MM water molecules + CPCM continuum solvation opt=10 steps
hybrid_w	Preserved water molecules closer than 3 Å as explicit MM water molecules + CPCM continuum solvation opt=water molecules frozen, sugar fully optimized
hybrid_fo	Preserved water molecules closer than 3 Å as explicit MM water molecules + CPCM continuum solvation opt=full optimization
* Additional parameters	B3LYP/6-311++G**

where

$$L(\tilde{v}; \tilde{v}_{ij}) = \frac{2}{\pi} \cdot \frac{\Gamma}{4(\tilde{v} - \tilde{v}_{ij})^2 + \Gamma^2} \quad (6)$$

is a Lorentzian function using a full width at half maximum (FWHM) $\Gamma = 7.5 \text{ cm}^{-1}$, N_{vib} is the number of calculated vibrational modes, and

$$T(\tilde{v}) = \frac{1}{\tilde{v}} \cdot \frac{1}{1 - e^{-\frac{\hbar\tilde{v}}{k_B T}}} \quad (7)$$

is a temperature correction factor, in our case applied at 310 K⁴¹. x denotes either Raman or ROA spectrum.

Raman, $I_{\text{Raman}}^{\text{sim}}(\tilde{v})$, and ROA, $I_{\text{ROA}}^{\text{sim}}(\tilde{v})$, spectra are obtained as an average over all snapshots, while the intensities are adjusted using the ρ_x parameter

$$I_x^{\text{sim}}(\tilde{v}) = \rho_x \cdot \frac{\sum_{j=1}^{N_{\text{sim}}} I_{x,j}^{\text{sim}}(\tilde{v})}{N_{\text{sim}}}, \quad (8)$$

where N_{sim} is the number of selected MD snapshots.

Final spectra were produced using 50 000 optimization cycles using eq. 3 when the spectra did not change anymore.

2.3.5 Accounting for the α/β Anomeric Ratio

In case of reducing monosaccharides (D-glucose, D-glucuronic acid, and N-acetyl-D-glucosamine) the α/β anomeric equilibrium

has to be always considered. The calculated spectra, $I_x^{\text{sim},\alpha/\beta}$, were obtained as a weighted average of α/β anomer spectra as follows

$$I_x^{\text{sim},\alpha/\beta} = (1 - x_\beta) I_x^{\text{sim},\alpha} + x_\beta I_x^{\text{sim},\beta}, \quad (9)$$

where x_β is a fraction of the β anomer, $I_x^{\text{sim},\alpha}$ is an average simulated spectrum of α anomer, and $I_x^{\text{sim},\beta}$ is an average simulated spectrum of β anomer. This equation was then used together with the Eq.3 to optimize the scaling function parameters. In this work, we calculate the spectra of reducing sugars using experimental values of x_β : D-glucose – 0.63^{22,46,47}, D-glucuronic acid – 0.61^{48,49}, and N-acetyl-D-glucosamine – 0.42⁵⁰.

2.4 Assessment of similarity of two spectra

To facilitate the comparison of tested methods (e.g., CPCM, MM, hybrid, hybrid_w, and hybrid_fo solvation models), we used an overlap integral S which is defined as

$$S = \frac{\int_{200}^{1800} I_f(\tilde{v}) I_g(\tilde{v}) d\tilde{v}}{\sqrt{\int_{200}^{1800} I_f(\tilde{v})^2 d\tilde{v} \int_{200}^{1800} I_g(\tilde{v})^2 d\tilde{v}}}, \quad (10)$$

where I_f and I_g represent two spectra that are to be compared. Upon comparing Raman spectra, the value of overlap integral is in range (0, 1), where values closer to 1 signal very similar spectra. Comparing ROA spectra yields a value in range (-1, 1), where values close to 1 signals very similar spectra, while -1 would yield exact mirror image. Note that in case of methyl- β -glucopyranoside the integration limit was set to (200,1490) cm^{-1} because of insufficient subtracted water signal appearing above that limit.

3 Results

3.1 Comparison of Simulated and Experimental Spectra

Figure 3 shows the experimental and simulated spectra of all studied monosaccharides. In the simulations, the hybrid solvation protocol from Table 1 was used. Overall, the computed spectra capture most of the experimental spectral features in all studied saccharides, that is relative Raman band intensities and for ROA also band signs, correctly. The low frequency region (200-800 cm^{-1}) is reproduced with high accuracy with the exception of D-glucuronic acid, while for the high-frequency region (800-1800 cm^{-1}), we observe larger deviations between calculated and experimental intensities, resulting in a semi-quantitative agreement only. The experimental peak in case of m β -glc close to 1620 cm^{-1} belongs to water bending and thus it was not reproduced by our calculations. Note that the spectra of glucose, glucuronic acid, and N-acetyl-D-glucosamine are coined of two independent spectra of α/β anomers. Incorporation of both anomers into the calculation is crucial and cannot be neglected as it is described in Section 3 in ESI, where the spectra of individual anomers can be found.

The spectra of methyl capped saccharides (m β -glc, m β -glcA and m β -glcNAc) are reproduced better than spectra of non-methylated sugars (glc, glcA, and glcNAc). Full optimization (hybrid_fo) can recover some of the originally non-reproduced peaks (Figure 1, 2, 3, 5, and 6 in ESI†), however, despite these minor improvements, the full optimization significantly worsened

the results as a whole. This deterioration indicates that the variability of the ensemble obtained via MD is a crucial factor in the success of our protocol. Full minimization of the extracted systems, which are just small regions of the complete phase space, reduces the number of local minima (configurations), impacting adversely the calculated spectra.

Overall, the `hybrid` computational protocol, which uses classically described water molecules, is fit to reproduce the main spectral features of the studied systems. Moreover, it provides similar quality of the results obtained when using QM described water molecules (for comparison, see Figure 15 in ESI†). The `hybrid` protocol provides reasonably accurate results, with a much lower computational cost. Computational time was saved primarily by not using DFT for the water molecules. For a monosaccharide, e.g. glucose (24 atoms), the first coordination shell contains at least 15 water molecules. By avoiding the DFT treatment of this first coordination shell, we reduced the computational time ~ 30 times. Other significant savings originate from employing the partial optimization being 3 to 8 times faster than a full optimization. In this way, we achieved one of the main objectives of our study making the computational protocol efficiently applicable also for small oligosaccharides.

3.2 Estimate of Experimental Error

As a simple approximate method of estimating the experimental error that can contribute to the disagreement between the experiment and theory, in Figure 4 we compare the experimental ROA spectra of D-glucuronic acid and N-acetyl-D-glucosamine obtained by different experimental groups. Clearly, the reproducibility of the ROA intensities is limited. Moreover, in central regions, e.g., for glcA 1050–1150 cm⁻¹, two experimental lines differ from each other more than they deviate from the calculation. Many factors contribute to such differences - different experimental setups, subtraction of baselines, quality of the products, etc. Although this is not an ideal situation, we can use the error associated with the experimental spectra as an argument that our computed spectra achieve sufficient accuracy for interpreting experiments.

3.3 Further Optimization of the Simulation Protocol

We devoted substantial attention to optimizing the spectra simulation protocol, that is to reduce the computational costs while preserving a reasonable agreement with the experimental data. We have considered multiple solvation methods, optimization procedures, functionals, and basis sets to calculate the Raman and ROA spectra.

Figure 5 summarizes the performance of the CPCM, MM, MM_w, hybrid, hybrid_w, and hybrid_fo approaches in terms of overlap integral S. Proper treatment of the aqueous solvent surrounding the monosaccharide turned out to be a critical point. Relative performances of individual solvation models (Table 1) are shown in Figures 1 to 5 in ESI†. We find that using only an implicit water solvent model can fail to reproduce experimental spectra, e.g., in case of the D-glucuronic acid (Figure 2 in ESI†). Including several solvation layers of explicit water molecules (cut-off 12 Å, MM protocol¹⁸) results in a significant improvement over

the CPCM approach. Full optimization of sugar, while the water molecules are kept frozen (`MM_w`) yields even better results, but at increased computational cost. A combination of implicit and explicit water models (`hybrid` protocol) performs even better compared to the `MM_w` protocol with the advantage of treating explicitly a significantly smaller system and being faster. Modification of `hybrid` protocol - `hybrid_w` was found to behave similarly as `MM_w`, however, still worse than the `hybrid`. The full optimization (`hybrid_fo`) was found to poorly reproduce low frequency regions at a high computational cost. The performance in relative cpu times are: cpcm : mm : MM_w : hybrid : hybrid_w : hybrid_fo = 1 : 3.7 : 7.5 : 1.7 : 3 : 14. This results into average performance in terms of overlap integral for Raman: 0.941 : 0.952 : 0.966 : 0.974 : 0.968 : 0.954 and for ROA: 0.689 : 0.761 : 0.778 : 0.805 : 0.773 : 0.735.

We conclude that the best performing approach is `hybrid` also providing results at reasonable cost.

Next, we have tested the number of optimization steps prior to Raman/ROA calculations. Figure 6 shows performance of `hybrid` approach when using 0 (MD structures), 1, 3, 5, 10, 30, 50, 100, 200, and 300(full optimization) optimization steps. Both Raman and ROA spectra needs 3-10 optimizations steps, after which the spectra deteriorates and the error goes up. This is due to the bad description of low frequencies which becomes noticeable after 30 optimization steps. The optimal number of optimization steps is 10, where the `hybrid` approach performs the best. The example of spectra evolution as a function of number of optimization step of glucuronic acid and methyl β-D-glucuronide are shown in Figure 9 and 10 in ESI†. We have further tested 'Normal mode optimization', which is commonly used in the literature⁵¹. It turned out to produce similar results as `hybrid` approach with exception of ROA spectrum of methyl β-D-glucuronide, where the frequencies near 1000 cm⁻¹ were poorly described (Figure 8 in ESI).

Next, we have tested 4 different DFT functionals (B3LYP, CAM-B3LYP, M06L, and ωB97XD). All functionals resulted in qualitatively similar results, however, the B3LYP functional performed the best as shown in Figure 7. As the B3LYP functional has been thoroughly tested for Raman and ROA spectra calculations^{12,52–54}, we have selected it for our optimized `hybrid` protocol.

The selection of a suitable basis set was also considered. Employing the `hybrid` solvation protocol we tested the cheaper 6-31G* and rDPS⁵⁵ (See Table 1 in ESI†), the intermediate 6-311++G** and the expensive aug-cc-pVTZ basis sets. Figure 8 shows that 6-31G* basis set is not large enough to reproduce spectra features, while 6-311++G** is sufficient and it performs almost the same as much more expensive aug-cc-pVTZ. Frequently used rDPS basis set^{11,13,17–20} offers slightly worse performance compared to 6-311++G** at the cost of \sim 6-31G* basis set.

3.4 Conformational Sampling

This work shows that a representative ensemble of structures is the most critical point for achieving accurate Raman and ROA

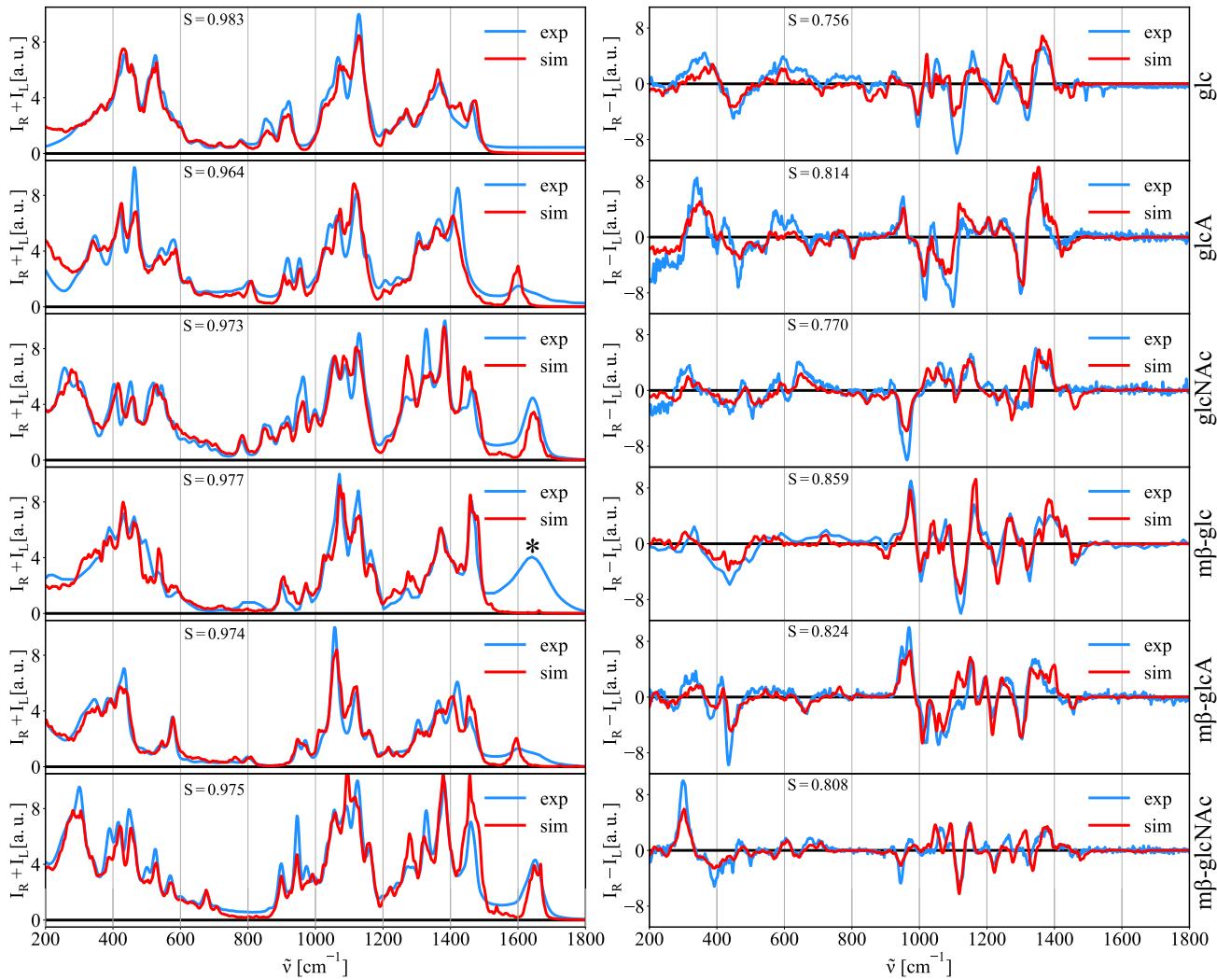


Fig. 3 Simulated and experimental Raman and ROA spectra of investigated monosaccharides. The calculated spectra were obtained using the hybrid protocol (Table 1). Optimized frequency scaling functions together with the parameters are to be found in Figures 1 to 5 in ESI†. *m β -glc: leftover water band (not connected to investigated saccharide).

spectra of saccharides. Not only do we need to properly sample the configurational space of the sugar, but we also need to map the sugar water environment reasonably well, to recover important spectral features. Two factors must be taken into account to properly describe the ensemble. The first one concerns structural nuances of the investigated sugar. Small differences in bond lengths, angles, and dihedral angles have a strong effect on the resulting spectrum. Figure 9 (top panel) shows, for example, the superposition spectra of different methyl β -D-glucuronide MD snapshots (black lines) obtained using the CPCM solvation approach (i.e., without explicit water molecules). The spectra of individual snapshots differ significantly from the average and experimental spectra, underscoring the effect of the sugar structural variability. Only sufficient averaging over a significant number of snapshots provides a computed spectrum (red line) that agrees reasonably with the experiment (blue line). Figure 16 in ESI† demonstrates that we need at least 30 snapshots to obtain a converged Raman spectrum and 80 snapshots to get a converged ROA spectrum of

a monosaccharide.

Figure 9 (middle panel) shows Raman and ROA spectra of methyl β -D-glucuronide using the hybrid approach (see Table 1). In this case, the ensemble of MD structures was generated while the geometry of the monosaccharide was frozen in an arbitrarily chosen conformation (4C_1). It is striking that despite the fixed geometry of the saccharide, the local water environment itself caused such a significant inhomogeneous broadening of the vibrational bands. Capturing both the sugar and water conformational freedom is thus essential for describing the key features of the spectra. (See the bottom panel of Figure 9).

4 Discussion

Contrary to other biomolecules conformational behavior of saccharides is much less explored. Raman and Raman optical activity (ROA) spectroscopies appear as powerful tools for gaining valuable insight into the structure of saccharides.

In this paper, we performed computations of Raman and ROA

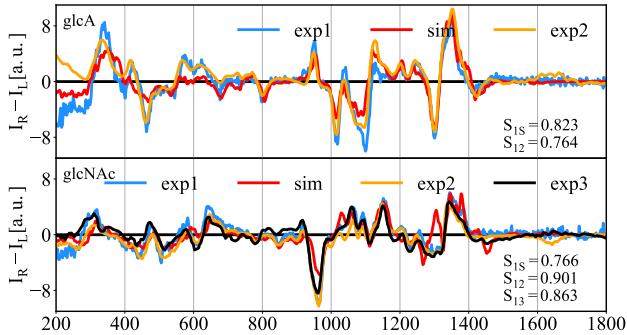


Fig. 4 Different experimental ROA spectra of D-glucuronic acid (glcA) and N-acetyl-D-glucosamine (glcNAc) available. In orange spectra from Ref. 20, in blue and black two independent spectra measured in this work (different spectrometer/sugar-provider). Simulation results obtained by the hybrid approach are in red.

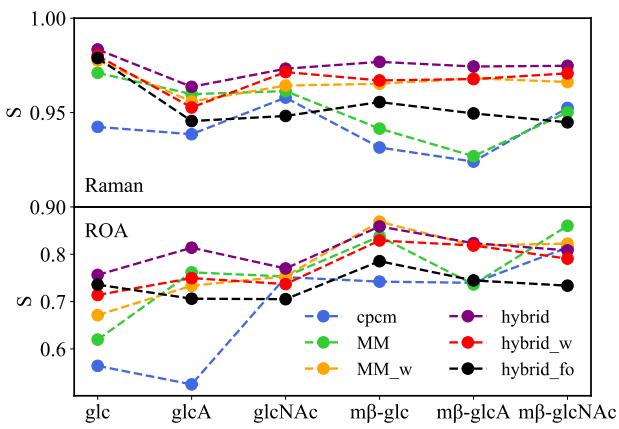


Fig. 5 Performance of employed protocols (CPCCM, MM, MM_w, hybrid, hybrid_w, and hybrid_{fo}) in terms of the overlap integral S for the studied monosaccharides species.

spectra of monosaccharides in a solution in a systematic way. We developed an optimized computational protocol to calculate the spectra for yielding a quantitative or at least a semi-quantitative agreement with experiment. This protocol is applicable also to larger structures, which will be addressed in future studies. The spectra obtained for a test set of monosaccharides — methyl β -D-glucopyranoside, methyl β -D-glucuronide, methyl β -N-acetyl-D-glucosaminide, D-glucose, D-glucuronic acid, and N-acetyl-D-glucosamine — agreed well with the experimental ones. Note that this set of monosaccharides comprises various sugar types, which suggest that the method can be applied to a large variety of sugar moieties.

An obvious but very important finding of this work is that there is not a single sugar structure that would fit the experimental spectrum, see Figure 9. Instead, we show that a representative ensemble of structures is crucial to reproduce the Raman and ROA experimental spectra. Only averaging of individual spectra thus results in a good approximation to the experimental spectra. This behavior has been to some extent observed for other flexible polyhydroxylic molecules⁵⁴. Moreover, we show that both

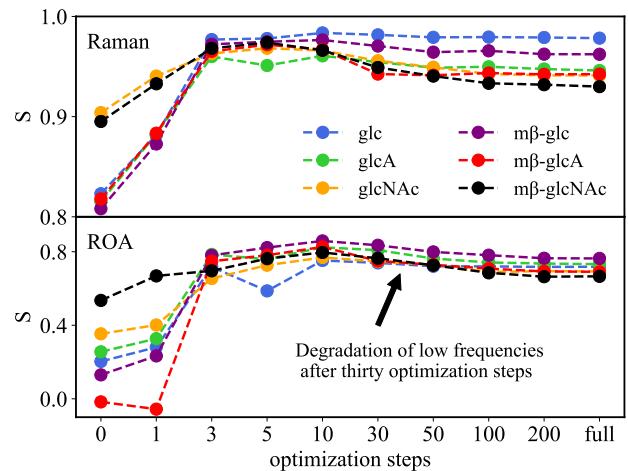


Fig. 6 Performance of the hybrid protocol when using different number of optimization steps prior to spectra calculations.

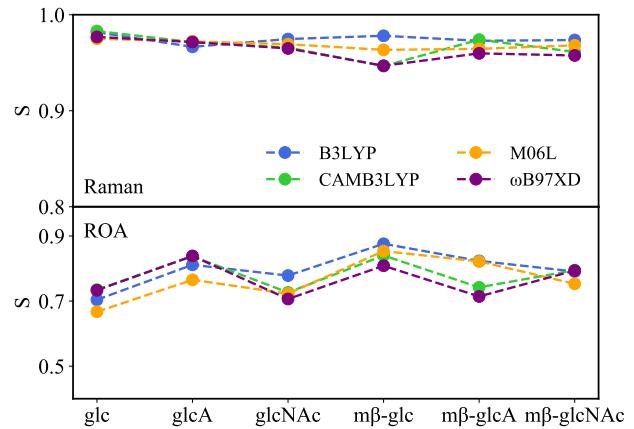


Fig. 7 Performance of the hybrid protocol when using different DFT functionals.

anomers have to be considered while aiming at simulation of reducing sugars. Failure to do so means essentially to partially simulate different system than the one that is studied experimentally.

Although we achieved good accuracy in reproducing the spectra, a question arises about the source of the remaining errors and also about the inconsistencies for D-glucose and D-glucuronic acid. Therefore, full optimization of glucose, together with using QM water molecules (Figure 15 in ESI†) was performed. Overall, the data suggests in the lower frequency region the description of water molecules mattered and QM description was superior to MM. However, the calculations were not repeated for the rest of considered molecules, since they are expensive. Another possible source of the errors is the incomplete conformational ensemble generated by molecular dynamics. We may be facing a force field problem, which goes beyond the scope of this work. Nevertheless, we could partially overcome this problem by using enhanced sampling techniques to find new configurations, followed by reconstruction of the spectra based on the experimental data.

Analyzing the main used protocols, in agreement with previous

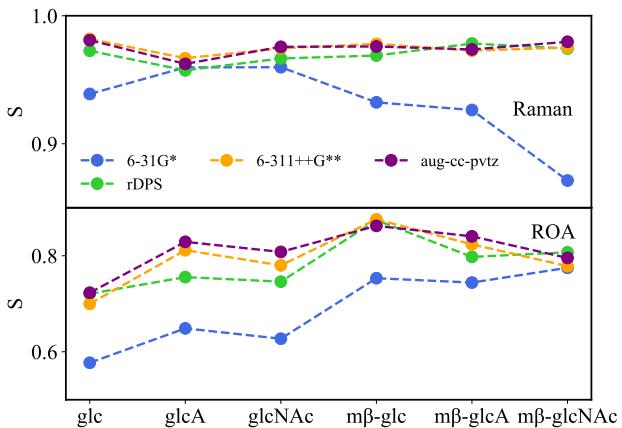


Fig. 8 Performance of the `hybrid` protocol when using different basis sets.

results¹³, we show that continuum solvation models already capture some spectral features. When adding harmonic restraints, the spectra improved, e.g., methyl β -glucopyranoside¹³. Overall, only using an implicit solvent model leads to apparent deficiencies in the calculated spectra, see e.g., ROA spectra of D-glucuronic acid in Figure 4 (top) in ESI†. In order to obtain a spectrum that can be compared quantitatively with experimental data, we need at least to include the first solvation shell explicitly. This inclusion can be done by either using the MM solvation scheme, where hundreds of MM described water molecules surround the investigated saccharide, or utilizing the `hybrid` solvation model that only includes the first coordination shell coupled with a continuum solvation model.

We further tested the `hybrid` protocol when limiting the number of QM optimization steps for each MD snapshot performed prior to calculation of Raman/ROA properties. It turned out to be a critical variable. Around 10 optimizations steps proved to be optimal while further optimization only worsened both Raman and ROA spectra. While 10 steps optimization in the `hybrid` scheme yielded the best results, when using the MM scheme it performed rather poorly. We were able to recover most of the errors performing full optimization of sugar, while keeping the water molecules frozen (`MM_w`, i.e. OPTSOLUTE approach used by “Manchester group”^{13,17}), however, at additional computational cost. Further comparison of `hybrid_w` and `hybrid_fo` showed that excessive optimization without any restraints only worsen the results. Overall, `hybrid` performs the best, while being computationally the most efficient.

In subsequent tests, we show that the impact of using different DFT functionals is small. From those tested the B3LYP performed the best. On the other hand, the choice of a basis set turned out to be important. Small basis set 6-31G* was not able to accurately reproduce the spectral features. The widely used rDPS basis set^{11,13,17–20} (see Table 1 in ESI†) provides significantly better results at the similar cost as 6-31G* basis set. Using 6-311++G** basis set provides further improvement of results at $\sim 2.5 \times$ the performance cost as compared to rDPS. Going even for larger and much more expensive aug-cc-pvtz provides only a little im-

provement. Therefore, we recommend 6-311++G** where the accuracy is crucial, while the rDPS offers best quality/cost ratio.

Unlike in most of the previous studies, we used an adaptable scaling function for vibrational frequencies $\phi(\tilde{\nu}^{\text{sim}})$. As it ultimately allows for superposition of spectra it facilitated the visual comparison of the simulated spectra. Moreover, it increases the agreement with experimental data as with its use we obtain $\sim 97\%$ overlap for Raman and $\sim 81\%$ for ROA spectra.

Note that low frequencies at the B3LYP/6-311++G** level of theory practically do not need to be scaled, while high frequencies ($> 1200 \text{ cm}^{-1}$) need to be scaled by a factor of ~ 0.98 . This can be seen in Figure 17 in ESI†, where all optimized scaling functions are shown. Optimizing the scaling parameters for each simulation is certainly desirable and for each of the protocol, a unique average scaling function exists. For the `hybrid` protocol we recommend $\phi(\tilde{\nu}^{\text{sim}}, a, b, c, d) = \phi(\tilde{\nu}^{\text{sim}}, 0.982, 1.00, 15, 1210)$, while for `rDPS` protocol (see Table 1 in ESI†) we recommend $\phi(\tilde{\nu}^{\text{sim}}, a, b, c, d) = \phi(\tilde{\nu}^{\text{sim}}, 0.957, 0.983, 40, 1437)$.

Comparing the performance in the relative cpu times with respect to `cpcm` can be found in a Table 2, where it is shown that the `hybrid` approach as it provides the best quality/cost ratio.

Lastly, we show on glucose that using QM water molecules rather than MM makes small difference at high computational cost (relative cpu times with respect to `hybrid` are `hybrid : hybrid_QM : hybrid_QM_fo = 1 : 25 : 70`; Raman overlap integral: 0.983 : 0.978 : 0.980; ROA overlap integral: 0.756 : 0.635 : 0.720).

Table 2 Comparison of performance of used protocols and their efficiency.

Protocol	rel. cpu time*	$\langle S_{\text{Raman}} \rangle$	$\langle S_{\text{ROA}} \rangle$
CPCM	1	0.941	0.698
MM	3.7	0.952	0.761
MM_w	7.5	0.966	0.778
hybrid	1.7	0.974	0.805
hybrid_w	3	0.968	0.773
hybrid_fo	14	0.954	0.735

*with respect to CPCM

5 Conclusions

We have developed and tested a computational protocol (`hybrid`, see Table 1) designed to accurately and efficiently simulate Raman and ROA spectra of saccharides. This `hybrid` protocol is based on MD simulations. A sufficient number of snapshots (~ 100) is extracted, and only water molecules closer than 3 \AA to the solute are kept. The long range solvation effects are modeled by the CPCM continuum model. The snapshots are then optimized using a 10 step optimization. The QM calculations are performed at the B3LYP/6-311++G** level of theory for the he explicit water molecules are treated at the MM level. This approach provides a good agreement with experimental data, with residual inconsistencies for Raman and ROA intensities, being accurate enough to determine conformational preferences of the sugars. Special attention is paid to making the protocol computationally efficient so that it can also be used for larger oligosaccha-

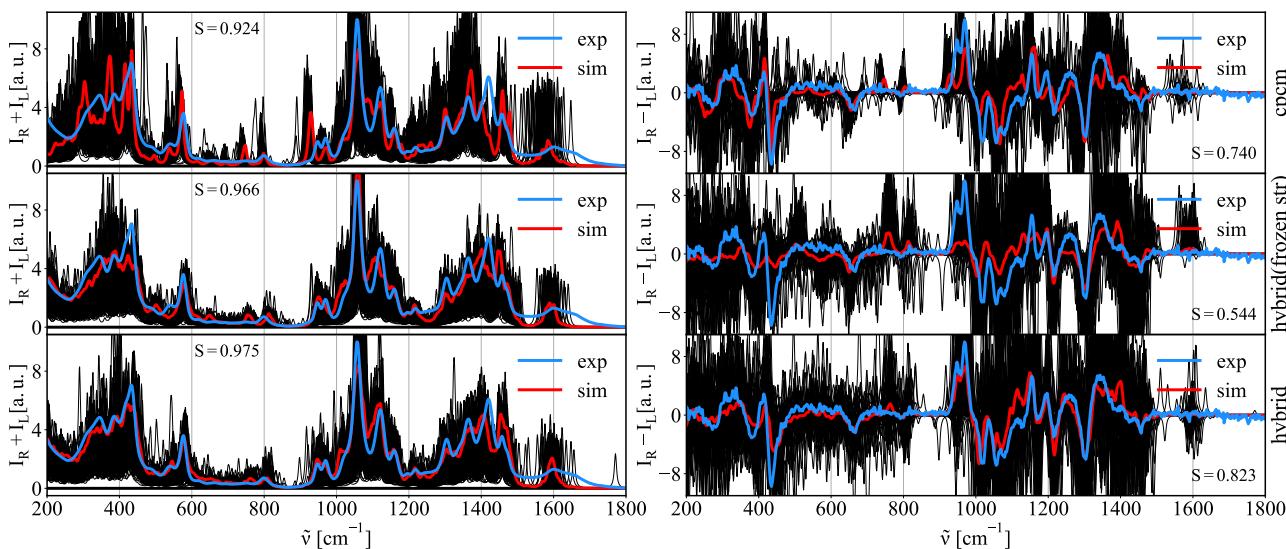


Fig. 9 Superposition of hundred Raman and ROA spectra of MD snapshots for a methyl β -D-glucuronide (black), and the corresponding average spectrum (red) when using continuum (top) and hybrid (middle/bottom) solvation protocols on the same ensemble. In the middle panel, we show the effect of water molecules on the spectra using the hybrid (frozen carbohydrate structure) protocol. Experimental spectra are plotted in blue.

rides. Simulations of the Raman and ROA intensities thus make it possible to extract detailed information about the structure of saccharides in water solutions.

Conflicts of interest

In accordance with our policy on Conflicts of interest please ensure that a conflicts of interest statement is included in your manuscript here. Please note that this statement is required for all submitted manuscripts. If no conflicts exist, please state that "There are no conflicts to declare".

Acknowledgements

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated. HMS, JK, and VP acknowledges support from the Czech Science Foundation (19-19561S). PB acknowledges support from the Ministry of Education (LTC17012). PJ acknowledges support from the European Regional Development Fund OP RDE (project ChemBioDrug no. CZ.02.1.01/0.0/0.0/16_019/0000729).

References

- J. M. Tarbell and M. Y. Pahakis, *J. Intern. Med.*, 2006, **259**, 339–350.
- J. M. Tarbell and L. M. Cancel, *J. Intern. Med.*, 2016, **280**, 97–113.
- R. R. Kapaev, K. S. Egorova and P. V. Toukach, *J. Chem. Inf. Model.*, 2014, **54**, 2594–2611.
- O. Dahlman, A. Jacobs, A. Liljenberg and A. I. Olsson, *J. Chromatogr. A*, 2000, **891**, 157–174.
- K. Matsuo, H. Namatame, M. Taniguchi and K. Gekko, *J. Phys. Chem. A*, 2012, **116**, 9996–10003.
- K. Matsuo, *Biomed. Spectrosc. Imaging*, 2017, **6**, 111–121.
- L. D. Barron, M. P. Bogaard and A. D. Buckingham, *Journal of the American Chemical Society*, 1973, **95**, 603–605.
- N. R. Yaffe, A. Almond and E. W. Blanch, *J. Am. Chem. Soc.*, 2010, **132**, 10654–10655.
- T. R. Rudd, R. Hussain, G. Siligardi and E. A. Yates, *Chem. Commun.*, 2010, **46**, 4124–4126.
- V. Profant, C. Johannessen, E. W. Blanch, P. Bouř and V. Baumruk, *Physical chemistry chemical physics*, 2019, **21**, 7367–7377.
- S. T. Mutter and E. W. Blanch, *Polysaccharides*, Springer International Publishing, 2015, pp. 1181–1218.
- A. Melcrová, J. Kessler, P. Bouř and J. Kaminský, *Phys. Chem. Chem. Phys.*, 2016, **18**, 2130–2142.
- J. R. Cheeseman, M. S. Shaik, P. L. A. Popelier and E. W. Blanch, *J. Am. Chem. Soc.*, 2011, **133**, 4991–4997.
- T. Giovannini, M. Olszówka, F. Egidi, J. R. Cheeseman, G. Scalmani and C. Cappelli, *J. Chem. Theory Comput.*, 2017, **13**, 4421–4435.
- S. Luber and M. Reiher, *J. Phys. Chem. A*, 2009, **113**, 8268–8277.
- N. A. Macleod, C. Johannessen, L. Hecht, L. D. Barron and J. P. Simons, *Int. J. Mass Spectrom.*, 2006, **253**, 193–200.
- R. Pendrill, S. T. Mutter, C. Mensch, L. D. Barron, E. W. Blanch, P. L. A. Popelier, G. Widmalm and C. Johannessen, *ChemPhysChem*, 2019, **20**, 695–705.
- F. Zielinski, S. T. Mutter, C. Johannessen, E. W. Blanch and P. L. A. Popelier, *Phys. Chem. Chem. Phys.*, 2015, **17**, 21799–809.
- S. T. Mutter, F. Zielinski, C. Johannessen, P. L. A. Popelier and E. W. Blanch, *J. Phys. Chem. A*, 2016, **120**, 1908–1916.
- S. T. Mutter, F. Zielinski, J. R. Cheeseman, C. Johannessen,

- P. L. A. Popelier and E. W. Blanch, *Phys. Chem. Chem. Phys.*, 2015, **17**, 6016–6027.
- 21 A. Rüther, A. Forget, A. Roy, C. Carballo, F. Mießmer, R. K. Dukor, L. A. Nafie, C. Johannessen, V. P. Shastri and S. Lüdeke, *Angewandte Chemie International Edition*, 2017, **56**, 4603–4607.
- 22 Y. Zhu, J. Zajicek and A. S. Serianni, *J. Org. Chem.*, 2001, **66**, 6244–6251.
- 23 V. Profant, M. Pazderková, T. Pazderka, P. Maloň and V. Baumruk, *J. Raman Spectrosc.*, 2014, **45**, 603–609.
- 24 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
- 25 D. E. Smith and L. X. Dang, *J. Chem. Phys.*, 1994, **100**, 3757–3766.
- 26 K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.*, 2008, **29**, 622–655.
- 27 S. Izadi and A. V. Onufriev, *J. Chem. Phys.*, 2016, **145**, 074501.
- 28 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 29 M. R. Shirts, D. L. Mobley, J. D. C. and V. S. Pande, *J. Phys. Chem. B*, 2007, **111**, 13052–13063.
- 30 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 31 S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511–519.
- 32 W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695–1697.
- 33 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
- 34 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian16 Revision B.01*, 2016.
- 35 L. Nafie, *Vibrational optical activity: Principles and applications*, Wiley, Chichester, 2011.
- 36 K. Ruud, T. Helgaker and P. Bouř, *J. Phys. Chem. A*, 2002, **106**, 7448–7455.
- 37 S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma and M. J. Frisch, *Theochem.*, 1999, **461–462**, 1–21.
- 38 M. Orozco, I. Marchán, I. Soteras, T. Vreven, K. Morokuma, K. V. Mikkelsen, A. Milani, M. Tommasini, M. D. Zoppo, C. Castiglioni, M. A. Aguilar, M. L. Sánchez, M. E. Martín, I. Fdez. Galván and H. Sato, *Continuum solvation models in chemical physics*, John Wiley & Sons, Ltd, Chichester, UK, 2007, pp. 499–605.
- 39 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 40 G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, K. Samanta and J. K. Ellis, *J. Chem. Phys.*, 2011, **135**, 124108.
- 41 P. L. Polavarapu, *J. Phys. Chem.*, 1990, **94**, 8106–8112.
- 42 M. K. Kesharwani, B. Brauer and J. M. L. Martin, *J. Phys. Chem. A*, 2015, **119**, 1701–1714.
- 43 M. L. Laury, M. J. Carlson and A. K. Wilson, *J. Comput. Chem.*, 2012, **33**, 2380–2387.
- 44 P. Sinha, S. E. Boesch, C. Gu, R. A. Wheeler and A. K. Wilson, *J. Phys. Chem. A*, 2004, **108**, 9213–9217.
- 45 M. Jeffrey P., D. Moran and L. Radom, *J. Phys. Chem. B*, 2007, **111**, 11683–11700.
- 46 W. A. Bubb, *Concepts Magn. Reson.*, 2003, **19A**, 1–19.
- 47 F. Franks, P. J. Lillford and G. Robinson, *J. Chem. Soc., Faraday Trans. 1*, 1989, **85**, 2417.
- 48 T. Bezabeh, O. B. Ijare, N. Albiin, U. Arnelo, B. Lindberg and I. C. P. Smith, *Magn. Reson. Mater. Phys., Biol. Med.*, 2009, **22**, 267–275.
- 49 B. Moreau, G. Lognay, C. Blecker, J. Destain, P. Gerbaux, F. Chéry, P. Rollin, M. Paquot and M. Marlier, *Biotechnol. Agron. Soc. Environ.*, 2007, **11**, 9–17.
- 50 F. C. Liu, C. R. Su, T. Y. Wu, S. G. Su, H. L. Yang, J. H. Y. Lin and T. S. Wu, *Int. J. Mol. Sci.*, 2011, **12**, 5828–5843.
- 51 P. Bouř and T. A. Keiderling, *J. Chem. Phys.*, 2002, **117**, 4126–4132.
- 52 S. Yamamoto, J. Kaminský and P. Bouř, *Anal. Chem.*, 2012, **84**, 2440–2451.
- 53 V. Parchařský, J. Kapitán, J. Kaminský, J. Šebestík and P. Bouř, *J. Phys. Chem. Lett.*, 2013, **4**, 2763–2768.
- 54 J. Kaminský, J. Kapitán, V. Baumruk, L. Bednárová and P. Bouř, *J. Phys. Chem. A*, 2009, **113**, 3594–3601.
- 55 G. Zuber and W. Hug, *The Journal of Physical Chemistry A*, 2004, **108**, 2108–2118.