Use of Raman and Raman optical activity to extract atomistic details of saccharides in aqueous solution

Vladimír Palivec,[†] Christian Johannessen,[‡] Jakub Kaminský,[†] and Hector Martinez-Seara[*,†]

†*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czech Republic*

‡*Department of Chemistry, University of Antwerp, Antwerp, Belgium*

E-mail: hseara@gmail.com

Abstract

Sugars are crucial components in biosystems and industrial applications. In aqueous environments, the natural state of short saccharides or charged glycosaminoglycans is floating and wiggling in solution. Therefore, tools to characterize their structure in a native aqueous environment are crucial but not always available. Here, we show that a combination of Raman/ROA and, on occasions, NMR experiments with Molecular Dynamics (MD) and Quantum Mechanics (QM) is a viable method to gain insights into structural features of sugars in solutions. Combining these methods provides information about accessible ring puckering conformers and their proportions. It also provides information about the conformation of the linkage between the sugar monomers, i.e., glycosidic bonds, allowing for identifying significantly accessible conformers and their relative abundance. For mixtures of sugar moieties, this method enables the deconvolution of the Raman/ROA spectra to find the actual amounts of its molecular

constituents, serving as an effective analytical technique. For example, it allows calculating anomeric ratios for reducing sugars and analyzing more complex sugar mixtures to elucidate their real content. Altogether, we show that combining Raman/ROA spectroscopies with simulations is a versatile method applicable to saccharides. It allows for accessing many features with precision comparable to other methods routinely used for this task, making it a viable alternative. Furthermore, we prove that the proposed technique can scale up by studying the complicated Raffinose trisaccharide, and therefore, we expect its wide adoption to characterize sugar structural features in solution.

1 Introduction

Carbohydrates are a key and diverse molecular group present in all living matter. It is crucial in biological functions and structure, including plant scaffold (e.g., cellulose), primary energy storage reservoir (e.g., starch/glycogen), or genetic code backbone (e.g., deoxyribose). The function and structural properties of sugars in these examples are well understood. Other roles of sugars have just started gathering attention, such as the role of branched polysaccharides attached to proteins[1] or glycosaminoglycans found in glycocalyx.[2] Glycosaminoglycans with their charged groups and hydroxyl groups are exceptionally hydrophilic and flexible, exploring a vast number of conformational states.[3] Proper understanding of glycosaminoglycans' structural properties in their functional environment, i.e., aqueous solution, is crucial to understanding their function.[4] However, correct and detailed structural determination of saccharides has proved to be challenging. Although there are many well-established techniques to extract atomistic details, most were developed for proteins or nucleic acids, and cannot be easily applied to flexible molecules like saccharides. For example, saccharides are very hard to crystallize, or often the obtained crystals are of insufficient quality for X-ray crystallographic experiments.[5] Moreover, even successful determination of the 3D structure from X-ray often yields incomplete structure due to the mobility of the saccharide moiety,[6] or the obtained structural details differ from the solution state.[7,8] Saccharides also lack good

natural chromophores hindering their study using the UV-Vis fluorescence spectroscopies. Nowadays, the methods of Nucleic magnetic resonance (NMR) are mainly used to study saccharides in solution; however, even NMR has its problems such as overlapping signals or the need for isotopically labeled compounds. All in all, saccharides are challenging to study using traditional structural techniques, which leads to an effort in exploring new structural methods for their characterization. One promising method is Raman and Raman Optical Activity(ROA)[9] spectroscopy. ROA is a Raman-based spectroscopic method where the difference in right/left circularly polarised light spectra is recorded. The advantage of Raman/ROA techniques when studying saccharides is a transparent vibrational spectral region between 75 and 3100 cm^{-1}, as compared to more traditional infrared spectroscopy (IR) coupled with vibrational circular dichroism (VCD), which offer limited information due to the strong water absorption. Moreover, Raman/ROA spectroscopies are very sensitive to conformational changes in saccharides both in the fingerprint[10] (75-1600 cm^{-1}) and CH stretching region[11] (2700-3100 cm^{-1}). Although Raman/ROA spectroscopies are readily accessible experimentally, their experimental spectra alone provide limited structural information due to their complexity and the substantial effects of the surrounding water molecules. Simulations can disentangle the crowded spectra and extract the underlying structural information. For example, the approach was used to investigate differences of two epimers, glucose and galactose,[12] and to study their vibrational properties. It also allows studying mixtures such as glucose/mannose,[13] which spectra can be decoupled to individual moieties. Another example is identifying conformers in flexible sugars such as gluconic acid that contribute to the experimental spectra.[14] It has also been applied to study sorbose/glucose in the aqueous solution in an extended vibrational range (75-3100 cm^{-1}),[11] where the effects of small conformational changes of CH_2OH group were linked to the overall shape of spectra. The method is not limited only to monosaccharides as it has been applied to the study of agarose trisaccharides in the gel phase to gain insight into its packing[15] or as in the study of differently linked mannobioses.[16] The ROA spectra were also used to characterize glucose upon its phospho-

rylation and study the influence on conformational preferences.[3] All of these studies show that Raman/ROA spectroscopies benefit from coupling to computer simulations and that there is and interest of utilizing the techniques to study saccharides.

The accuracy of simulation protocols limits the extraction of structural features, and therefore, significant effort has been devoted to developing reliable and economical computational protocols. Practically all state-of-the-art protocols nowadays consist of ensemble averaging many structures, usually tens to hundreds,[12,16–18] generated with MD simulations. The microscopic structure of the first solvation layer (3-3.6 Å cutoff) was found to be very important,[10–12,14,17,18] while the long-range polarization can be modeled either with continuum solvation methods or using a larger shell of explicit water molecules(10-12 Å cutoff).[13,18] The spectra are then calculated for solute-solvent clusters at the DFT level, usually using the B3LYP or BPW91 functionals, with a sufficiently large basis set,[10,19] e.g., 6-311++G** or aug-cc-pVTZ. Since Raman/ROA tensor invariant calculations demand diffuse basis set functions, while force field calculations generally do not, performing split calculations have been proposed to speed up the calculations. For example, the rDPS basis set[20] consists of calculating the force field part with 6-31G* basis set, while the tensor invariants with 3-21++G basis set augmented with a semi-diffuse p functions with an exponent of 0.2 on hydrogen atoms. For the final spectra calculations, the ensemble generated with MD needs to be optimized at the quantum mechanics (QM) level of theory; however, complete optimization is detrimental.[10] Partial optimization approaches such as the Normal modes optimization[21] or optimizing only the central molecule while freezing the rest of the environment[12] are used. It was recently shown that a simple ten-step optimization without any restrains is sufficient to capture the sought effects while being very simple.[10] Usually, some empirical correction of calculated vibrational frequencies is performed,[16] either as a simple one-factor scaling or using more elaborate methods.[10] Many of the presented important factors were recently thoroughly investigated on a set of six model monosaccharides, and as a result, an accurate and cost-efficient computational protocol was proposed.[10] This protocol is used in this work

to explore structural details of various saccharides in an aqueous solution.

In this paper, we explore the boundaries of using Raman/ROA spectroscopies coupled with computer simulations to extract structural details of saccharides in an aqueous solution. We show that rotation around glycosidic bonds of disaccharides leads to profound spectral changes that unambiguously identify the most abundant rotamers. Moreover, we study populations of ring conformers using methyl-β-D-glucuronic acid as a model system aiming to distinguish between various puckered conformations and their relative weights.[22] The current applicability of the simulation protocols to larger saccharides is tested on a raffinose trisaccharide. We also extract anomeric equilibrium values on a set of well-characterized monosaccharides and this approach is then extended to binary mixtures of known compositions to determine the experimental mixing ratios. Lastly, we explore the effect of intermolecular interactions on a model system of two stacked monosaccahrides. When possible, we compare our results to MD/NMR data to highlight that Raman/ROA brings additional structural information and that it can be used as a complementary technique to study saccharides in a water environment.

2 Methods

2.1 Target Saccharides

We have chosen six pyranose monosaccharides: D-glucose (Glc), D-glucuronic acid (GlcA), 2-acetamido-2-deoxy-D-glucose (aka *N*-acetyl-D-glucosamine; GlcNAc), and their methyl glycosides, methyl β-D-glucoside (MeGlc), methyl β-D-glucuronide (MeGlcA), and methyl 2-acetamido-2-deoxy-β-D-glucoside (MeGlcNAc) to study how Raman and ROA spectroscopies can be used to extract structural details of saccharides in solution (Figure 1). We have also included four glucose or mannose containing disaccharides: α-D-glucopyranosyl-(1→1)-α-D-glucopyranose (trehalose), methyl 1α-2α-mannobiose (M12), methyl 1α-3α-mannobiose (M13), and methyl 1α-6α-mannobiose (M16) to study the effects of rotation around glyco-

sidic bonds on Raman/ROA spectra (Figure 2; explicit atom definition of dihedral angles in SI). An important feature of these disaccharides is that they are non-reducing, and therefore, they do not epimerize to an α/β anomeric mixture. Lastly, we have also included a trisaccharide - raffinose (Gal(α1-6)Glc(α1-2β)Fruf) (Figure 2, bottom), to test the performance of the simulation protocol on a larger moiety.

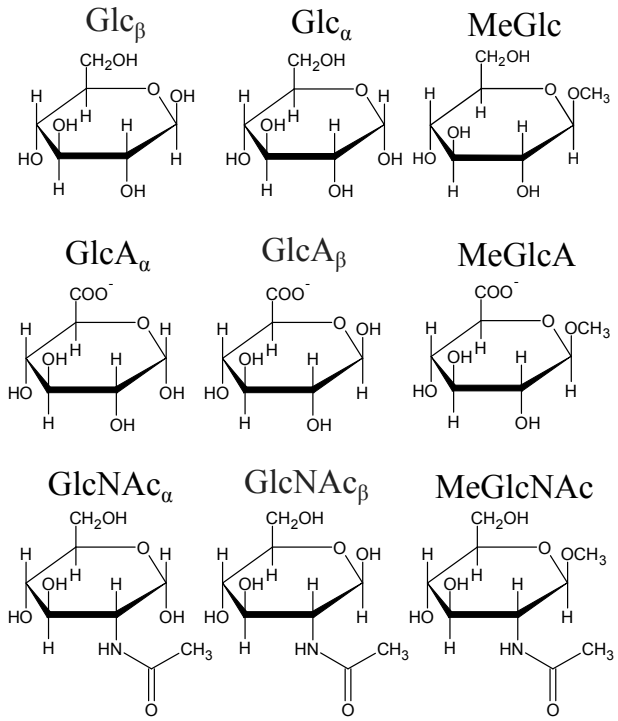

Figure 1: Investigated monosaccharides. The Figures were prepared with ChemSketch[23] and editted with Inkscape.[24]

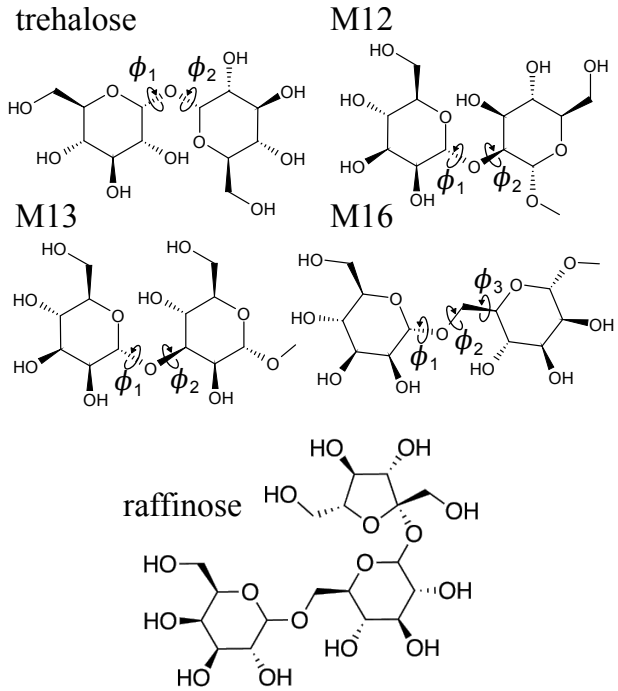

Figure 2: Investigated disaccharides: trehalose, methyl-1α-2α-mannobiose (M12), methyl-1α-3α-mannobiose (M13), and methyl-1α-6α-mannobiose (M16), together with defined glycosidic dihedral angles ϕ_1, ϕ_2, and ϕ_3 (explicit atom definitions in SI). Bottom raffinose trisaccharide.

2.2 Simulation details

Simulation of Raman/ROA spectra and connection to experimental data consist of several steps. 1) A representative ensemble of structures is generated using molecular dynamics simulations. 2) Raman/ROA spectra are calculated for each of the structures using quantum chemical methods, obtaining ensemble-averaged spectra. 3) We score the similarity of simulated spectra with experiment, and when applicable, we find the best fit of simulated data to the Raman/ROA spectra. 4) When possible, we compare our calculations to NMR data. Underneath, we explain each of these steps.

Ensemble generation. Molecular dynamics (MD) simulations were used to generate representative ensembles used to simulate spectra of studied saccharides. All MD simulations were performed using the Gromacs-2018.4[25] simulation package patched with Plumed 2.5.[26] The systems of interest were simulated at ambient conditions (p = 1 bar and T = 300 K)

using Glycam-6h[27] (sugars) and OPC3[28] (water) models. The rest of the simulation parameters are rather standard and are thoroughly described in SI. We performed three types of calculations. 'Unbiased MD' simulations from which representative ensemble of structures were obtained, for which subsequently Raman/ROA spectra was calculated. 'Free energy' simulations which used well-tempered metadynamics MD[29] to calculate the free energy profile along studied coordinate such as glycosidic dihedral angles or puckering coordinates. We use those to find local minima on the free energy profiles for which we performed a third type of a simulation. These are 'biased MD' simulations that restrained the system around selected local minima using harmonic position restrains for which ensemble average spectra were calculated. The ensemble averaged spectra for each minima were used to best fit experimental data therefore estimating their contributing weights to the overall spectra. Representative ensembles were obtained by uniformly extracting 250 structures from the 'biased MD' and 'unbiased MD' simulations. When studying Binary Mixtures and Crowding Effects 500 structures were used. Table 1 summarizes all performed simulations in this work. For further details see methods in SI.

Calculation of Raman/ROA Spectra. Raman/ROA spectra of simulated ensembles were calculated using a previously developed `hybrid` simulation approach.[10] In brief, this simulation protocol is a QM/MM based method using the Gaussian16 program package[30] that consists of extracting snapshots from the MD trajectory, where the central molecule together with surrounding water molecules (3 Å cutoff) are kept. The water molecules are treated at the MM level of theory, while the solute is calculated at the B3LYP/6-311++G** level of theory using the ONIOM method with electrostatic embedding.[31,32] TIP3P[33] and Glycam-6h[27] force field parameters were used to describe the MM part of the ONIOM method. The long-range solvation effects are taken care of by using the COSMO implicit solvation model. After preparation, the system is optimized in ten steps unrestrained optimization. After optimization, we use the far-from-resonance coupled-perturbed Kohn-Sham theory[34,35] at 532 nm (SCP180) and harmonic approximation to calculate vibrational fre-

Table 1: Summary of MD simulations performed in this work.

Conformation of Glycosidic Bonds (trehalose, M12, M13, and M16)	
1 unbiased MD	500 ns
2a free energy	metadynamics MD in $\phi_1/\phi_2/\phi_3$ glycosidic dihedral angles
2b biased MD	$md1/md2/md3/md4/md5/md6$ conformer regions 200 ns
Probing Puckering Conformations (MeGlcA)	
1 unbiased MD	500 ns
2a free energy	metadynamics MD in q_x, q_y, and q_z puckering coordinates
2b biased MD	${}^1C_4/\ {}^4C_1/{}^OS_2/{}^1S_3$ puckering conformer regions 50 ns
α/β anomeric equilibrium constant (Glc, GlcA, and GlcNAc)	
unbiased MD	500 ns for each of the α/β anomer
Raffinose Trisaccharide	
unbiased MD	500 ns
Binary Mixtures	
unbiased MD	500 ns for pure MeGlc and MeGlcNAc
Crowding Effects	
unbiased MD	500 ns for a single MeGlc
unbiased MD	500 ns for two MeGlc in close proximity [a]

[a] using a flat bottom potential with a COM-COM distance < 0.5 nm

quencies and Raman/ROA intensities. All calculated vibrational frequencies are rescaled using a frequency scaling function $\phi(\tilde{\nu}, a, b, c, d) = \phi(\tilde{\nu}, 0.982, 1.00, 15, 1210)$.[10] Calculated intensities were corrected for the temperature factor at 300 K and convoluted using Lorentzian with a full width at half maximum (FWHM) $\Gamma = 7.5$ cm^{-1}.

To asses the quality of simulated spectra we used the overlap integral S defined in Eq 1.

$$S = \frac{\int_{\tilde{\nu}_{min}}^{\tilde{\nu}_{max}} I_{sim}(\tilde{\nu}) I_{exp}(\tilde{\nu}) d\tilde{\nu}}{\sqrt{\int_{\tilde{\nu}_{min}}^{\tilde{\nu}_{max}} {I_{sim}}^2(\tilde{\nu}) d\tilde{\nu} \int_{\tilde{\nu}_{min}}^{\tilde{\nu}_{max}} {I_{exp}}^2(\tilde{\nu}) d\tilde{\nu}}}, \tag{1}$$

where I_{sim} and I_{exp} represent the simulated and experimental spectra that are to be compared. For more details, refer to former work[10] and SI.

The Best Fit to Experimental Raman/ROA Spectra. The experimental spectra are calculated as a cumulative average of all spectra produced by any accessible state of the system during the experiment. We strive to find relevance of each of selected states and

their contribution to the final spectra. For this, computationally, we first calculate ensemble average spectra for each state ($I^{sim}_{i,Raman/ROA}$) such as α/β anomer forms or various molecular conformations. We seek to find their contribution to the overall simulated spectra,

$$I^{sim}_{Raman/ROA}(\vec{A}) = \sum_{i=0}^{spectra} A_i \cdot I^{sim}_{i,Raman/ROA}, \tag{2}$$

A_i is the weight given to a where specific state i and $\vec{A}$ is the vector of weights. The set of weights A_i is calculated based on the similarity of simulated ($I^{sim}_{Raman/ROA}$) and experimental spectra, i.e., by minimizing the following cost function F:

$$F(\vec{A}) = (1 - S_{Raman}(\vec{A}))^2 + (1 - S_{ROA}(\vec{A}))^2, \tag{3}$$

where $S_{Raman}(\vec{A})$ and $S_{ROA}(\vec{A})$ are overlap integrals defined in Eq.1. The robustness of found best fit solution, i.e., the uncertainty of obtained set of weights $\vec{A}$, is determined by numerically finding weights that fulfill $S_{Raman} + S_{ROA} > 0.99(S^{bestfit}_{Raman} + S^{bestfit}_{ROA}))$. The spread of the weights values $\{A^{min}_i, A^{max}_i\}$ is reported as the uncertainty of the best fit solution.

Calculating conformer weights using NMR ^3J spin-spin coupling constants. The ensemble average ^3J coupling constants were determined using Karplus equation[36] for glycosidic bonds of M12 and M13 disaccharides using raw MD data. We also used QM calculations for glycosidic bonds of M12 and M13 disaccharides and puckering of MeGlcA. The QM derived coupling constants were calculated using optimized snapshots used for Raman/ROA calculations stripped of all water molecules at mPW1PW91/pc-J2[37] level of theory, together with the CPCM continuum solvation. To speed up the calculations, only the Fermi contact terms were considered. As in Raman/ROA spectra calculations, we determined weights of each conformer region by minimizing G,

$$G(\vec{A}) = \sum_{i=0}^{coupling\ constants} ({}^3J^{sim}_i - {}^3J^{exp}_i)^2, \tag{4}$$

where ${}^3J_i^{sim}$ is the ensemble averaged simulation spin-spin coupling constant,

$$ {}^3J_i^{sim} = \sum_{j=0}^{conformers} A_j \cdot {}^3J_j^{sim}. \tag{5} $$

We evaluated the robustness of the obtained best fit weights by considering 0.2 Hz deviation (reported experimental error[36,38]) from the best fit solution, see SI. The spread of the weights values $\{A_i^{min}, A_i^{max}\}$ is reported as the uncertainty of the best fit obtained weights.

2.3 Experimental section

D-Raffinose pentahydrate (raffinose) was purchased from Sigma, MeGlc, MeGlcNAc, and trehalose were purchased from Carbosynth. Backscattered Raman and Scattered Circular Polarization (SCP) ROA spectra of all studied saccharide samples were recorded using the ChiralRAMAN-2X TM (Biotools Inc.) spectrometer. Stock aqueous solutions of 1 M concentration were prepared by dissolution of raffinose, MeGlc and MeGlcNAc in milliQ water. Only 0.42 M solution of trehalose was used for further measurements due to limited amount of the sample (10 mg / 70 μL). Three different mixtures (1:1, 1:3 and 3:1) of mixtures of MeGlc and MeGlcNAc were prepared from their 1 M stock solutions. The impurities causing undesirable fluorescence background were partially removed from the samples by activated carbon filtration. Residual fluorescence was quenched by leaving the sample in the laser beam for at least an hour prior to the experiment. Then the Raman and ROA spectra of 1M MeGlc, 1M raffinose, 0.42M trehalose, and three mixtures of MeGlc/MeGlcNAc were recorded. We used the 532 nm green laser with the power at the head from 700 mW (trehalose) to 900 mW (raffinose, MeGlc/MeGlcNAc mixtures). The illumination period was always set to 2.5 s. Rectangular fused silica cell with approximately 3 mm optical path and 50 μL sample volume was used for all experiments. The Raman and ROA spectra were recorded simultaneously using average acquisition time of approx. 24 h for each sample. Water signal was subtracted from the Raman spectra. When necessary a polynomial baseline

correction was performed to fix any residual background or baseline drifting. The wavenumber scale was calibrated using a neon lamp. Fluorescence standard material (SRM 2242; (National Institute of Standards and Technology, USA) was used for the ROA and Raman intensity calibration.

The rest of the Raman/ROA experimental data was obtained from the literature, i.e., Glc, GlcA, and GlcNAc, as well as the spectra of pure MeGlcA, and MeGlcNAc[10] and M12, M13, and M16.[16]

3 Results

3.1 Conformation along the Glycosidic Bonds

First, we investigate selected disaccharides to monitor changes in the Raman and ROA spectra upon the rotations around their glycosidic bonds connecting the two monosaccharide units. This linker usually possesses two degrees of freedom, the $\{\phi_1, \phi_2\}$ angles. Our workflow will be thoroughly explained using methyl-1α-3α-mannobiose as a example. The results using the same protocol for trehalose and methyl-1α-2α-mannobiose will be relegated to the SI as they are similar. We conclude this section by discussing methyl-1α-6α-mannobiose. The 1$\rightarrow$6 linked disaccharide is a rather special case as it possesses an additional degree of freedom. Note that conformational preferences of differently linked mannobioses have been studied previously under various conditions and using different methods for saccharides in their free states or as bound to lectins.[39–43] Similarly, also trehalose have been thoroughly studied predominantly using NMR and/or MD simulations.[44–46]

3.1.1 Conformation of Methyl-1α-3α-mannobiose (M13)

The left panel of Figure 3 shows the calculated free energy surface profile (FES) around $\{\phi_1, \phi_2\}$ dihedral angles for M13. The middle panel overlays the FES with the $\{\phi_1, \phi_2\}$ values obtained for 250 structures extracted from unbiased 500 ns MD simulation. The right panel

shows the structures extracted from four biased MD simulations using the harmonic position restraints around the sampled local minima in the FES ($md1/md2/md3/md4$; defined in Table 2). Compared to unbiased MD simulation, this last method allows to sample $md3$ and

Table 2: Restrain values [rad] of ϕ_1/ϕ_2 glycosidic angles used in biased MD simulations of M13 yielding $md1/md2/md3/md4$ conformers.

	$md1$	$md2$	$md3$	$md4$
M13	-0.8/0.6	-1.0/-0.6	-0.6/-2.8	0.8/0.5

$md4$ regions, and to probe their contribution to Raman/ROA spectra. According to the FES profile of M13, we identified three different $\{\phi_1, \phi_2\}$ energy minima regions. Nevertheless, since the global minimum is quite narrow but long, we performed two biased MD simulations within this region, $md1$ and $md2$, differing mostly in ϕ_2. Figure 4 (top right - MD) shows then

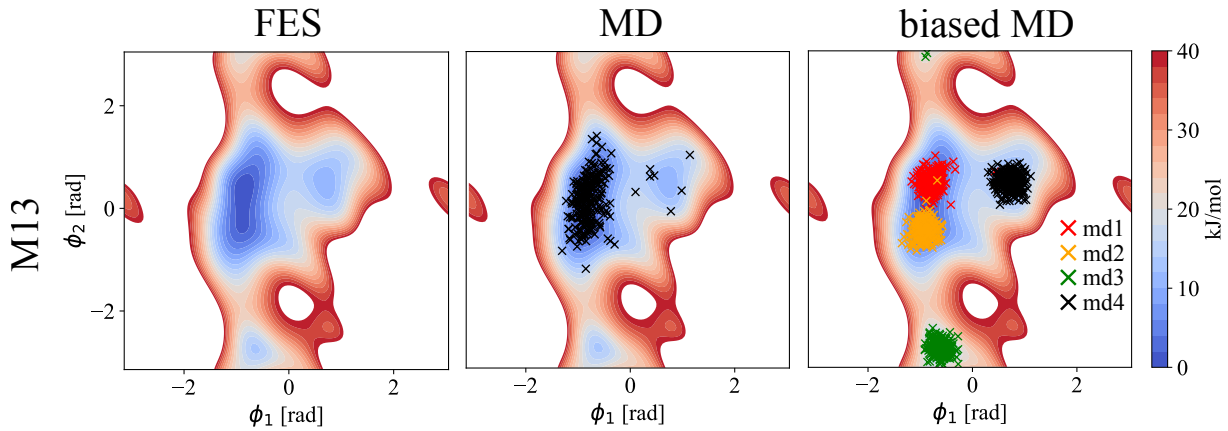

Figure 3: Left: Calculated free energy surface (FES) of M13 disaccharide in $\{\phi_1, \phi_2\}$ dihedral angles. Middle: Calculated FES, together with 250 extracted structures from unbiased 500 ns MD simulations (MD). Right: Calculated FES, together with 250 extracted structures for each biased 200 ns MD simulation (biased MD; $md1/md2/md3/md4$; restrain $\{\phi_1, \phi_2\}$ values in Table 2). White regions represent area with the free energy > 40 kJ/mol.

calculated average Raman and ROA spectra that were obtained using the unbiased MD generated structures (Figure 3, MD - black crosses) and their comparison to experimental data. The simulated spectra already agree well with experiment, indicating a good performance of used MD force field. Most of the experimental spectral bands are correctly reproduced

with comparable accuracy (S $>$ 0.9 for Raman and S $>$ 0.8 for ROA) as in our previous works with monosaccharides.[10,11] Next, we used structures extracted from the biased MD simulations, which efficiently sample local minima of the FES profile. In Figure 4 (bottom), we show the Raman/ROA spectra for each of the conformer regions ($md1/md2/md3/md4$). Subsequently, we calculated the best fit spectra combining and weighting each conformers spectra to experimental data (Figure 4, best fit). Firstly, by inspecting the spectra of a selected conformer region (local-conformer), we see that the sensitivity of the Raman spectra to changes of the glycosidic angles is rather limited as the overlap integrals are in the range of $S_{Raman} = 0.936 - 0.952$ (0.948/0.952/0.943/0.936). On the other hand, the ROA spectra are much more affected $S_{ROA} = 0.550 - 0.817$ (0.817/0.760/0.603/0.550). Most significant differences between the conformers are concentrated in spectra below 900 cm^{-1} and they happen both in the Raman and ROA spectra. Interestingly, although the Raman spectra for frequencies $>$ 900 cm^{-1} are almost identical, there are significant changes in ROA spectra. Finally, our fitting procedure (best fit) provides simulated spectra in better agreement ($S_{Raman} = 0.950$ and $S_{ROA} = 0.844$) with the experimental data than any conformer spectra. Only the $md2$ Raman spectrum alone agrees slightly better than best fit, but its ROA counterpart is far worse. This observation justifies the need of simultaneous fitting of the Raman and ROA spectra as this approach is more constrained and selective.

Although the MD simulation spectra and the best fit spectra provide almost identical accuracy (MD: $S_{Ram} = 0.951$, $S_{ROA} = 0.838$; best fit: $S_{Ram} = 0.948$, $S_{ROA} = 0.844$, they represent different conformer ratios. While the best fit data corresponds to $md1/md2/md3/md4$ = 0.48/0.31/0.07/0.12, the MD simulation gives 0.57/0.40/0.00/0.03. We observe an increase of $md3$ and $md4$ conformers, which MD misses. These differences can arise due to the equilibration issues (i.e., not long enough simulations) or because of inaccurate force field parameters. However, further analysis of obtained best fit ratios showed that there is a significant error in each conformer contribution estimation. Table 3 shows the best fit conformer ratios, together with error estimates of the fitting procedure described in the methods

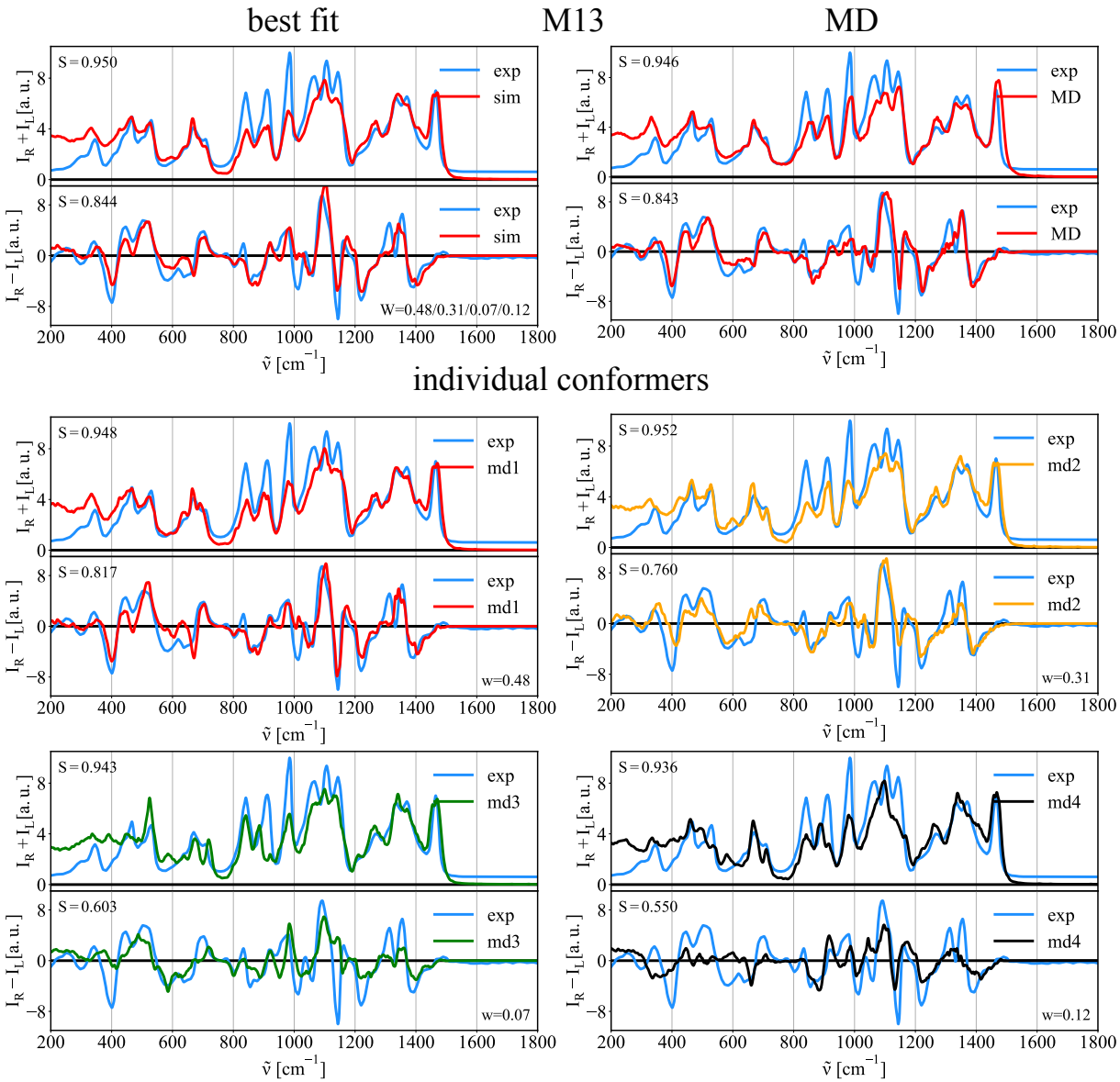

Figure 4: Comparison of experimental (exp.) and calculated spectra of M13 disaccharide. Top left - best fit: Best fit of $md1/md2/md3/md4$ Raman/ROA spectra to experimental data. Top right - MD: Simulated Raman/ROA spectra obtained using structures from unbiased MD simulation (MD). Bottom - local-conformers: Calculated ensemble averaged Raman/ROA spectra of M13 disaccharide prepared in 4 distinct conformations $md1/md2/md3/md4$ as described in Figure 3.

section. The conformer ratios are determined with a 20-30% uncertainty, making effectively the Raman/ROA spectroscopies blind to contributions below this threshold. Nevertheless, by fitting the Raman/ROA spectra, we can distinguish all major conformers that have to be

present in order to reproduce experimental data.

Table 3: Abundance of conformers of M13 disaccharide as obtained by MD, NMR, and the best fit to Raman/ROA (error bars in brackets).

M13 conformer	MD	NMR (Karplus)	NMR (QM)	Best fit
$md1$	0.57	0.61-0.73[a] (0.25-0.89)	0.43-0.55[a] (0.26-0.67)	0.48 (0.18-0.89)
$md2$	0.40	0.00-0.23[a] (0.00-0.55)	0.00-0.42[a] (0.00-0.56)	0.31 (0.00-0.67)
$md3$	0.00	0.16-0.18[a] (0.04-0.29)	0.00-0.28[a] (0.00-0.37)	0.07 (0.00-0.35)
$md4$	0.03	0.00-0.09[a] (0.00-0.21)	0.04-0.30[a] (0.00-0.41)	0.12 (0.00-0.35)

[a]no unique solution - values in this range fit experimental data

We also estimated the $md1/md2/md3/md4$ populations using the NMR $^3J_{CH}$ spin-spin coupling constants[36] and assessed the associated error. To calculate $^3J_{CH}$ from our simulations, we applied two approaches. In the first, we calculated the $^3J_{CH}$ coupling constants using the Karplus equation[36] for each frame in the bias MD simulations. In the second, we obtained the $^3J_{CH}$ coupling constants at the DFT level for the same set of frames, but after the QM optimization. Average $^3J_{CH}(sim)$ values were then calculated for each of the $md1/md2/md3/md4$ conformers, and then we performed the fitting of simulated values to experimental data,[16] together with the error estimation (see Table S6 and S7 in the SI for details of the fitting results). When using both approaches, we did not obtain a single best fit solution, but rather a range of values that fit the experimental data. This is because only two $^3J_{CH}$ coupling constants contain information about the glycosidic bond and are therefore fitted. The best fit population ranges are given in Table 3. Surprisingly, assuming 0.2 Hz deviation from the best fit solutions leads in both cases to large deviations making the NMR and Raman/ROA populations comparable, and burdened by comparable errors.

Considering the large uncertainty obtained when fitting the NMR and Raman/ROA experimental data, the MD populations of local-conformations probably reflect reality well. All methods that we used predict that the $md1$ conformer is the most abundant one (40-70%), followed with the $md2$ conformer (20-40%) with possible minor contributions of $md3$ and $md4$ conformers. Evaluating contributions of conformers with such small weights (0-

15%) is inconclusive as the uncertainty of abundances obtained by NMR/Raman/ROA are of the similar scale, and at least around 20%. Similar conclusions are drawn when studying trehalose and M13 disaccharides indicating a good performance of the Glycam force field for 1→1, 1→2, and 1→3 bonded disaccharides, at least in the mannose and glucose series. However, the additional degree of freedom in the 1→6 bonded M16 is more complicated, and it will be discussed in the next section.

3.1.2 Conformation of methyl-1α-6α-mannobiose (M16)

As compared to the previously studied disaccharides, the free energy surface (FES) of M16 is unique. The linkage of the two mannose units consists of three bonds, and therefore, it requires three unique dihedral angles $\{\phi_1, \phi_2, \phi_3\}$ to describe its conformation. Fortunately, the ϕ_3 angle population is rather confined around $\phi_3 \sim -0.84$ rad with a sharp profile with a full width at half maximum of 0.42 rad (See Figure S5 in the SI). We therefore integrated this variable obtaining a 2D free energy profile (ϕ_1 vs. ϕ_2), see Figure 5. The obtained free energy profile has several accessible minima as compared to M12, M13, or trehalose disaccharides. Therefore, instead of sampling four regions as with M13/M12/trehalose, we sampled six local minima ($md1/md2/md3/md4/md5/md6$; defined in Table 4; depicted in Figure 5-right). For

Table 4: Restrain values [rad] of ϕ_1/ϕ_2 glycosidic angles used in biased MD simulations of M16 yielding $md1/md2/md3/md4/md5/md6$ conformers.

	$md1$	$md2$	$md3$	$md4$	$md5$	$md6$
M16	3.14/3.14	0.8/-3.14	0.9/-1.7	-0.8/-3.14	-0.9/1.4	2.9/1.3

each of the minima region, we calculated corresponding Raman and ROA spectra. Figure 6 shows the Raman and ROA spectra of M16 calculated using structures from our MD simulation (MD), as well as by performing the best fit of local-conformer spectra to experimental data (best fit). As for M13, the Raman spectra of M16 present less variations compared to their ROA counterparts when comparing local-conformer spectra. Also, while several experimental bands do not manifest for some conformers (e.g., ~ 600 cm^{-1} or ~ 850 cm^{-1} bands),

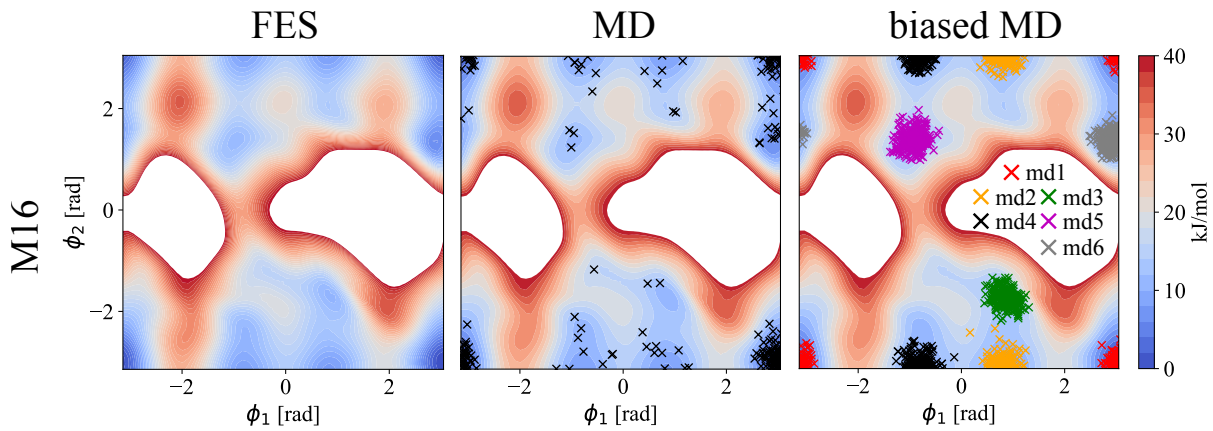

Figure 5: Left: Calculated free energy surface (FES) of M16 disaccharide in terms of the $\{\phi_1, \phi_2\}$ dihedral angles. Middle: Calculated FES, together with 250 extracted structures from unbiased 500 ns MD simulations (MD). Right: Calculated FES, together with 250 extracted structures per each biased 200 ns MD simulation (biased MD; $md1/md2/md3/md4/md5/md6$; restrain $\{\phi_1, \phi_2\}$ values in Table 4). White regions represents area with the free energy > 40 kJ/mol.

it is difficult to select particular bands as conformational probes. Again, major differences are in the lower frequency region below 1100 cm^{-1} where oscillations of the whole sugar are probed. We found that the best fit solution outperforms the MD spectra ($S_{Ram} = 0.951$ and $S_{ROA} = 0.828$ for the best fit versus $S_{Ram} = 0.936$ and $S_{ROA} = 0.700$ for MD). This likely points out a force field inaccuracy in the description of the M16 conformational behavior. The conformer ratios of MD vs. the best fit significantly differs, contrary to disaccharides with a one-atom junction, i.e., M13, M12, and trehalose. The MD populations of selected regions, local-conformer, are $md1/md2/md3/md4/md5/md6 = 0.74/0.08/0.02/0.07/0.02/0.07$, while the best fit provides populations of 0.58/0.00/0.00/0.16/0.24/0.00. We find that the best fit solution identify as significantly populated three conformers, i.e., $md1$, $md4$, and $md5$, contrary to MD simulation, where a single conformer ($md1$) dominates. However, calculated population error ranges are significant (10-20%) with all results summarized in Table 5. Calculated error ranges are similar to those previously discussed for M13 or M12 pointing to a fundamental limitation of the Raman/ROA spectral analysis in terms of estimation of the conformer populations. Straight connection to NMR data is more difficult

Table 5: Conformer abundances of M16 disaccharide obtained from the MD and the best fit to Raman/ROA experimental data. Estimated error ranges for the latter in parenthesis.

M16 conformer	MD	Best fit
$md1$	0.74	0.58 (0.36;0.78)
$md2$	0.08	0.00 (0.00;0.15)
$md3$	0.02	0.00 (0.00;0.13)
$md4$	0.07	0.17 (0.00;0.33)
$md5$	0.02	0.25 (0.00;0.48)
$md6$	0.07	0.00 (0.00;0.15)

for M16 as compared to other studied disaccharides, e.g., for M13, since experimental data are available only for the ϕ_1 dihedral angle.[47] Information about other two dihedral angles require for example $^3J_{CH}$ coupling constant across those glycosidic bonds requiring isotopically labeled compounds, which are harder to prepare. Since we do not have an access to such data, our NMR-based analysis of conformations is only limited. Nevertheless, some comparison of calculations with experiment is possible using ϕ_1 only. Since ϕ_1 is very similar for $md1$ and $md6$, $md2$ and $md3$, and $md4$ and $md5$ (see Figure 5), we can compare their sums. Analysis of available experimental $^3J_{HH}$ coupling constants using the Karplus equation can estimate the populations of $md1 + md6 = 0.49$, $md2 + md3 = 0.06$, and $md4 + md5 = 0.45$. Integrating our results show that MD provides $md1 + md6 = 0.81$, $md2 + md3 = 0.10$, and $md4 + md5 = 0.09$. The best fit to the Raman and ROA spectra results into $md1 + md6 = 0.58$, $md2 + md3 = 0.00$, and $md4 + md5 = 0.42$. The results are summarized in Table 6. Improving on the MD method, the best fit method to ROA/Ramam provides essentially the same values as NMR experiments, even after the errors considerations.

Table 6: Comparison of conformer populations of M16 disaccharide obtained using MD, NMR, and the best fit to experimental Raman/ROA data.

M16 conformer	MD	NMR[47]	Best fit
$md1$+$md6$	0.81	0.49	0.58
$md2$+$md3$	0.10	0.06	0.00
$md4$+$md5$	0.09	0.45	0.42

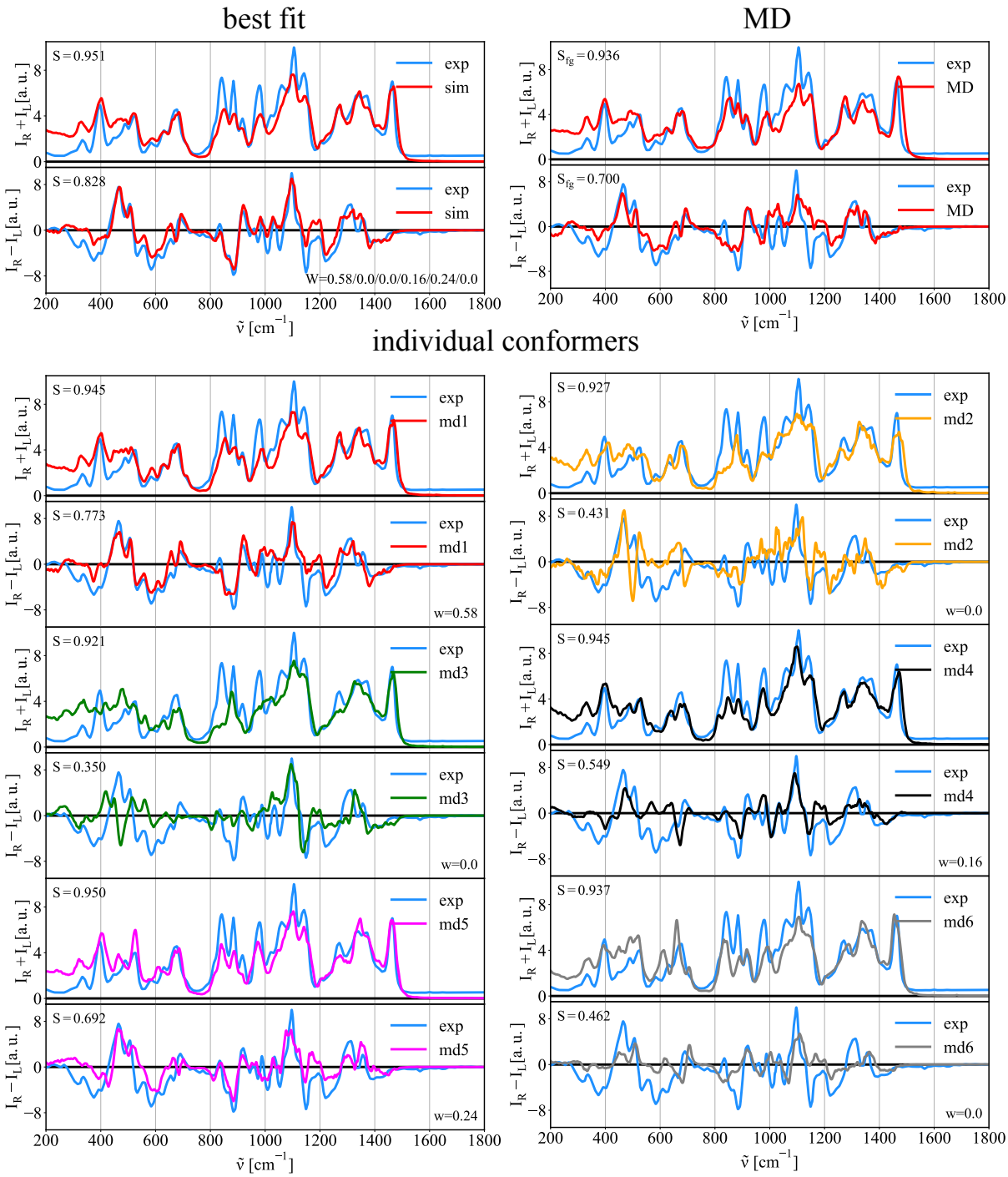

Figure 6: Comparison of experimental (exp.) and calculated spectra of M16 disaccharide. Top left - best fit: Best fit of $md1/md2/md3/md4/md5/md6$ Raman/ROA spectra to experimental data. Top right - MD: Simulated Raman/ROA spectra of disaccharide obtained using structures from unbiased MD simulation (MD). Bottom - local-conformers: Calculated ensemble averaged Raman and ROA spectra of disaccharide prepared in 6 distinct conformations $md1/md2/md3/md4/md5/md6$ as described in Figure 5.

3.2 Probing puckering conformation

The Raman and ROA spectra also allow probing of the puckering of saccharides, i.e., the ring accessible conformers. We will discuss the conformation of the pyranose ring further, but the approach is generally applicable also to saccharides with furanose rings. Figure 7 shows calculated free energy surface (FES) for methyl-β-D-glucuronic acid (MeGlcA) in the angular puckering coordinates, the MD sampled structures, and the structures gathered using the biased MD around the four selected local free energy minima($^1C_4$/ $^4C_1$/$^OS_2$/$^1S_3$; defined in Table 7). MeGlcA free energy surface obtained using the Glycam force field reveals one deep

Table 7: Restrain values [rad] of ϕ/θ puckering coordinates used in biased MD simulations of methyl-β-D-glucuronic acid yielding $^1C_4$/ $^4C_1$/$^OS_2$/$^1S_3$ conformers. $^1C_4$ and $^4C_1$ conformer were restrained only in θ puckering variable.

	$^1C_4$	$^4C_1$	$^OS_2$	$^1S_3$
MeGlcA	all/3.14	all/0	3.14/1.57	1.8/1.57

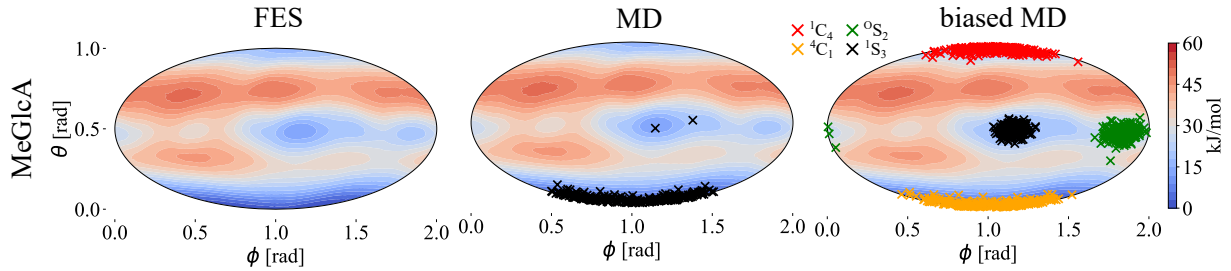

Figure 7: Left: Calculated FES of MeGlcA in ϕ/θ puckering coordinates. Middle: Calculated FES, together with 250 extracted structures from unbiased 500 ns MD simulations (MD). Right: Calculated FES, together with 250 extracted structures per each biased 200 ns MD simulation(biased MD; $^1C_4$/$^4C_1$/$^OS_2$/$^1S_3$; restrain $\{\phi, \theta\}$ values in Table 7). All plots are shown as equal area Mollweide projection.

global minimum around the south pole ($^4C_1$; $[\phi, \theta]$ of [1,0]). Mostly all structures sampled from unbiased MD simulation falls on that region as well. Still, we probe other local minima ($^1C_4$, $^OS_2$, and $^1S_3$) as to whether they contribute to the final Raman/ROA spectra which could potentially points towards a force field inaccuracy. Figure 8 shows the Raman/ROA spectra of MeGlcA obtained for the unbiased MD, and using the best fit procedure. It also

shows the spectra for each selected local-conformer each with different puckering coordinates, and their comparison with experiment. The unbiased MD spectra (both Raman and ROA) are practically identical to the best fit spectra (top right), which is composed purely from the $^4C_1$ spectra (bottom). All other conformers ($^1C_4/^OS_2/^1S_3$) do not contribute to the final spectra at all as summarized in Table 8. Again, we also tried to connect our

Table 8: Abundance of puckering conformers of MeGlcA as obtained by MD, NMR, and the best fit to Raman/ROA experimental data. Estimated error ranges for the NMR and the best fit approaches are in parenthesis.

MeGlcA conformer	MD	NMR[a]	Best fit
$^1C_4$	0.01	0.01 (0.00-0.03)	0.00 (0.00-0.08)
$^4C_1$	0.98	0.99 (0.96-0.99)	1.00 (0.91-1.00)
$^OS_2$	0.00	0.00 (0.00-0.03)	0.00 (0.00-0.08)
$^1S_3$	0.01	0.00 (0.00-0.03)	0.00 (0.00-0.08)

[a]exp data fit[38]

MD/Raman/ROA results to available experimental NMR data.[38] To do this, we calculated ensemble average $^3J_{HH}$ spin-spin coupling constants for considered $^1C_4/^4C_1/^OS_2/^1S_3$ conformers at the QM level of theory (see methods for details). Having obtained these, we performed the best fit to experimental data and then we estimated the error ranges of the fit (Table 8). Obtained populations are $^1C_4/^4C_1/^OS_2/^1S_3 = 0.01/0.99/0.00/0.00$, which is in a good agreement with both the MD and the best fit to Raman/ROA results. For the detailed values of the experimental and calculated ^3J spin-spin coupling constants see Table S8 in the SI.

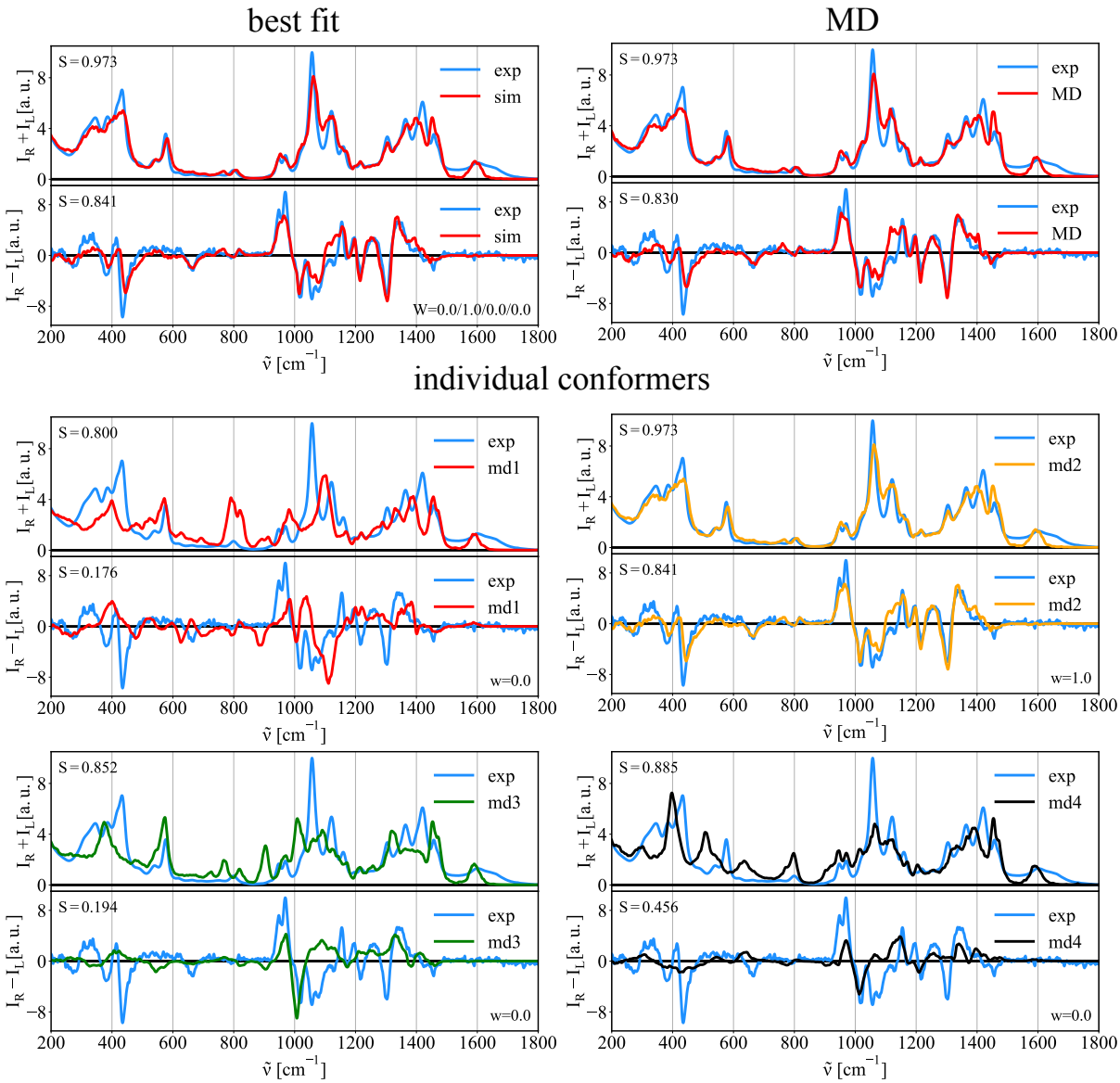

Figure 8: Comparison of experimental (exp.) and calculated spectra of MeGlcA. Top left - best fit: Best fit of $^1C_4/^4C_1/^OS_2/^1S_3$ Raman/ROA spectra to experimental data. Top right - MD: Simulated Raman/ROA spectra of the monosaccharide obtained using structures from unbiased MD simulation (MD). Bottom - local-conformers: Calculated ensemble averaged Raman and ROA spectra of the monosaccharide prepared in 4 distinct conformations $^1C_4/^4C_1/^OS_2/^1S_3$ as described in Figure 7.

3.3 Raffinose trisaccharide

Study the structure of larger saccharides in solution is a major goal of structural biology. Since the used computational protocol was developed on monosaccharides,[10] its performance in larger saccharides is not guaranteed. Therefore, we test the accuracy of the protocol here on a raffinose trisaccharide (See Fig. 2) and discuss raffinose structure based on the comparison of MD and experimental data. The MD calculations suggest that the molecule is found mainly in a few states (see Figures S6-S9 in SI), where the most variable degrees of freedom are the ϕ_1 and ϕ_2 angles between both pyranoses (see Figure S6 in SI). Using structures obtained from the unbiased MD simulation we calculated the Raman/ROA spectra given in Figure 9. A very good agreement with experiment was found considering the size and complexity of the system. The Raman spectrum is reproduced excellently with small overestimation of intensities at low frequencies around 250 cm^{-1}. These deviations, however, are also found for mannobioses, glucose, and glucuronic acid. The ROA spectrum is of slightly worse quality with the overlap integral of 0.733. Still, most features are unambiguously reproduced. Despite raffinose being a rather complex molecule, our results put into relevance the strength of our combined methodology to extract relevant structural features from sugars. We did not even need more advanced computational methods, e.g., Bias MD, to obtain reasonable outputs for this molecule.

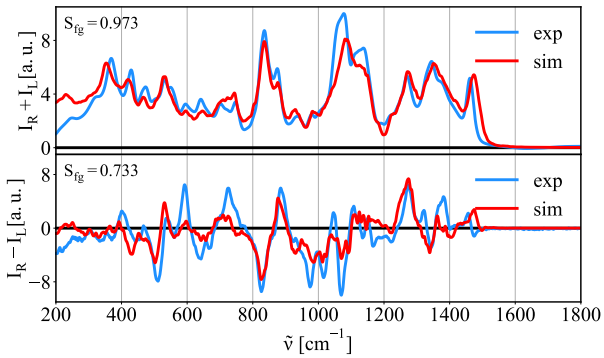

Figure 9: Comparison of calculated Raman and ROA spectra of raffinose trisaccharide with experiment.

3.4 The anomeric ratio

All reducing monosaccharides are in solution in equilibrium between their acyclic, and cyclic forms (e.g., pyranose/furanose). The reducing monosaccharides tested here (Glc, GlcNAc, and GlcA) are expected, according to NMR studies,[48] to be found overwhelmingly in the cyclic form. This cyclic form of reducing saccharides in turn exist in an anomeric equilibrium between the α, and β anomeric forms. While the contributions of the linear and furanose forms are negligible for our hexoses, the proportions of the α and β anomers are often comparable. Similarly to previous simulations, we can calculate pure spectra of the α and β anomers independently. We can then obtain the proportions of each anomer, i.e., anomeric equlibrium value, by the linear fitting of calculated α/β anomer spectra to the experimental naturally mixed spectra. Figure 10 shows simulated Raman and ROA spectra of the α/β anomers (left) and the best fit to experimental data (right).

Table 9 summarizes calculated anomeric equilibrium ratio of Glc, GlcA, and GlcNAc. We find a reasonable agreement with experimental NMR data with exception of glucuronic acid, where we see an error of 10%. Similarly as in previous cases, we show the error estimates in brackets. We can conclude that predicted anomeric ratios based on the Raman and ROA spectra are much closer to the real experimental values than traditional predictions based on Boltzmann populations obtained for *ab-initio* energies of each anomers that provide predictions with an error $> 15\%$.[13]

Table 9: Experimental and calculated anomeric ratios (β fraction) for investigated reducing sugars. Calculated data were obtained as the best fit of the Raman and ROA spectra to experimental data.

	Glc	GlcA	GlcNAc
Exp	0.62-0.65[48–50]	0.60-0.62[51,52]	0.35-0.42[53]
Sim	0.63 (0.49-0.76)	0.50 (0.33-0.68)	0.40 (0.24-0.56)

Although we are able to find a reasonable agreement with experimental data, the error of the calculation is significant (∼15%).

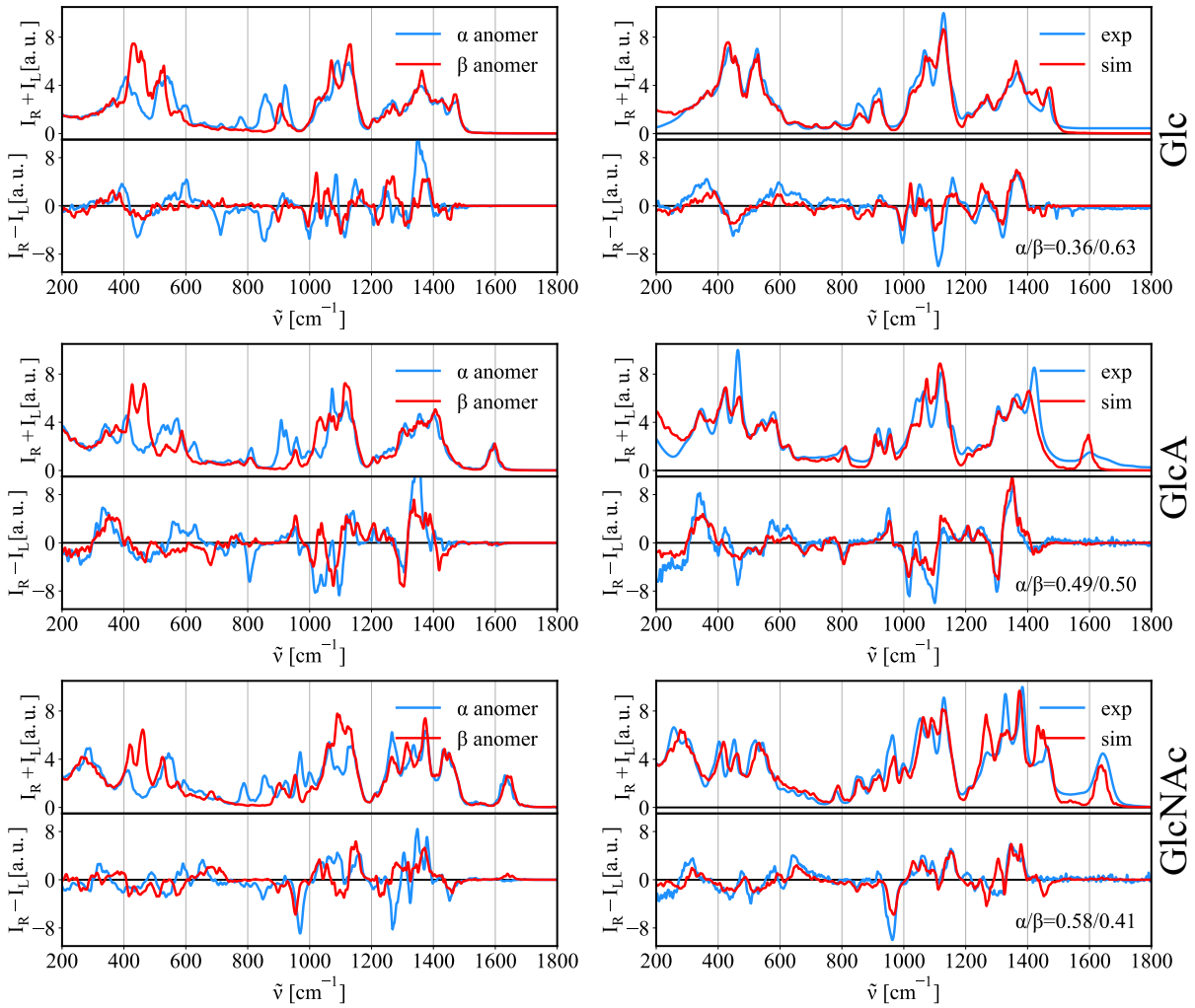

Figure 10: Left: Calculated Raman/ROA spectra for the α/β anomers (Glc, GlcA, and GlcNAc). Right: Best fit to experimental data.

3.5 Binary mixtures

We have discussed that it is possible to decompose experimental spectra to contributions of diferent chemical moieties and conformers, i.e., to calculate the anomeric equilibrium. We are not limited to the anomeric equilibrium, but any mixture can be studied in a similar manner using the same protocol. To exemplify this, we estimated the ratio of components in the experimentally prepared mixtures of methyl-β-glucose (MeGlc) and methyl-β-N-glucosamine (MeGlcNAc). Mixtures were prepared in 1:3, 1:1, and 3:1 ratios of MeGlc and MeGlcNAc. The pure substances were used as the references. Thus, the spectra of

pure saccharides were calculated and used to decompose experimentally measured spectra ($x_{MeGlc} = 0.00, 0.25, 0.5, 0.75, 1.00$). To probe the sensitivity of the methods, we have chosen chemical moieties with small structural differences. Therefore, we aim to capture relatively small differences in spectra (see simulated spectra of pure substances in section 5 in the SI). Figure 11 shows spectra of the decomposition using simulation spectra of pure substances to fit the experimentally prepared mixtures. To compare, we also used experimental spectra of pure substances for the same fit. The results are summarized in Figure 12 and Table 10.

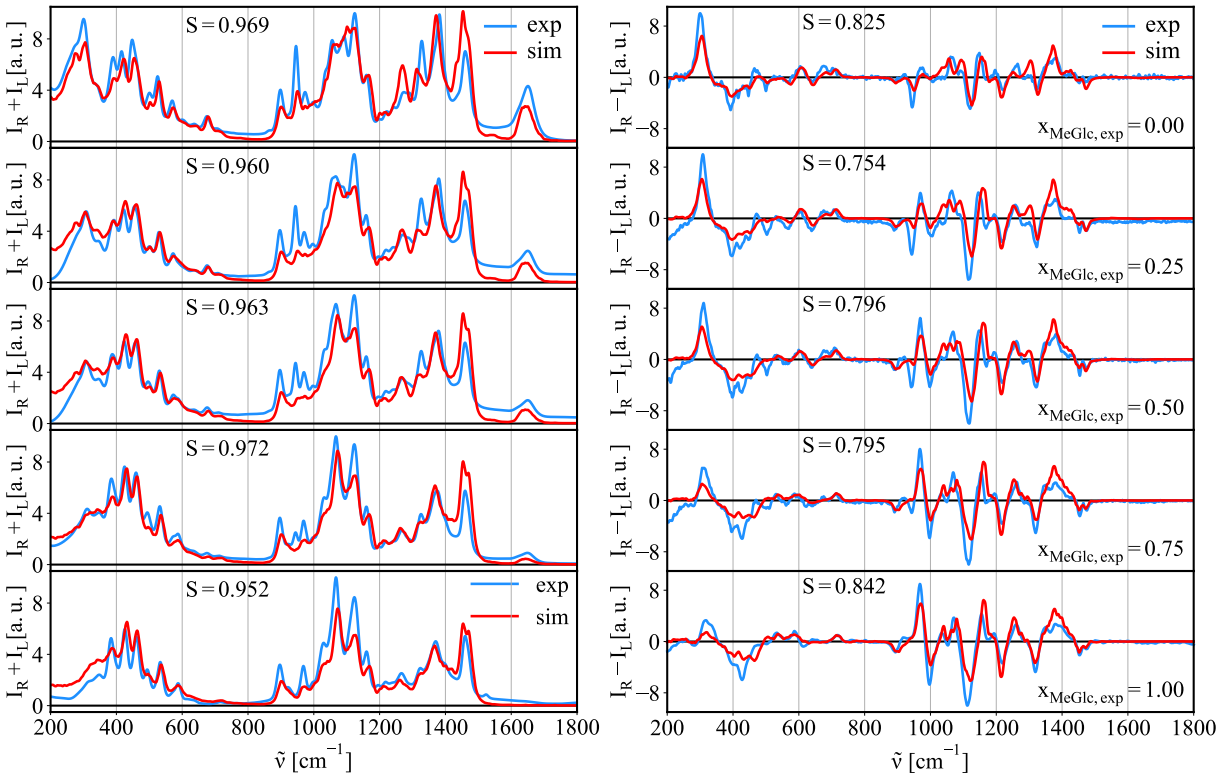

Figure 11: In blue the experimental Raman/ROA spectra of MeGlc and MeGlcNAc, and their 3:1, 1:1, and 1:3 mixtures (MeGlc:MeGlcNAc). The best fit simulation spectra to the experiment using the spectra of simulated pure substances to fit them are shown in red (see Figure 12 and Table 10 for the results of the fit).

In all cases, despite the molecular similarities, the experimental molar fraction was reproduced when using the simulated spectra of pure substances for the fit, with just an slight overestimation of MeGlc ranging between ∼0-22% (error). The determined upper boundary composition error is larger than the error reported in ref.,[13] where authors predicted the

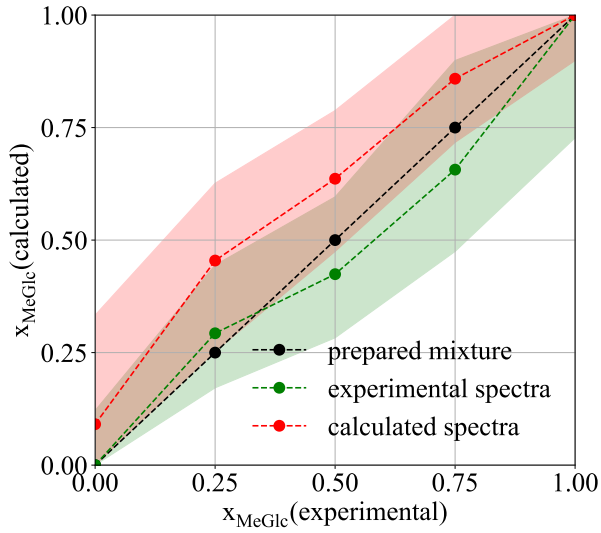

Figure 12: Summary of calculated molar fractions and estimated errors of mixtures (MeGlc:MeGlcNAc) obtained by the best fitting corresponding experimental ROA/Raman spectra of known composition (black,$x_{MeGlc} = 0.00, 0.25, 0.5, 0.75, 1.00$) using simulation(red, calculated spectra) or experimental(green, experimental spectra) spectra of pure substances to fit them.

glucose/mannose ratio based on the ROA decomposition with an error $< 15\%$. This can be attributed to more significant differences in the structure and conformational behavior of glucose and mannose than what we see here for MeGlc and MeGlcNAc. The errors are also slightly larger than those computed for the anomeric equilibrium predictions discussed above. Still, when we evaluate the uncertainty of the best fit, the molar fractions can change up to 15%. Within such a error range, our calculations cover the real concentration, see

Table 10: Summary of calculated molar fractions and estimated errors of prepared mixtures (MeGlc:MeGlcNAc; $x_{MeGlc} = 0.00, 0.25, 0.5, 0.75, 1.00$) obtained by the best fitting corresponding experimental ROA/Raman spectra using simulation(x^{sim}_{MeGlc}) or experimental(x^{exp}_{MeGlc}) spectra of pure substances to fit them.

x_{MeGlc}	0.00	0.25	0.50	0.75	1.00
x^{sim}_{MeGlc}	0.10 (0.00-0.33)	0.47 (0.25-0.63)	0.65 (0.48-0.78)	0.87 (0.72-1.00)	1.00 (0.90-1.00)
x^{exp}_{MeGlc}	0.00 (0.00-0.12)	0.29 (0.17-0.44)	0.42 (0.28-0.60)	0.66 (0.47-0.90)	1.00 (0.73-1.00)

Fig. 12. Using experimental spectra of pure substance instead of simulated ones for the fit results in smaller errors, i.e., $\sim 0 - 9\%$. However, the evaluated uncertainties of the fit are $\sim 15\%$, which makes the simulation and experimental fits well comparable.

3.6 Crowding effects

Crowded environments are common in living organisms, e.g., in their glycocalyx. This region is believed to be a relatively densely packed sugar region, where saccharides can potentially interact with each other. Here, we investigate whether such an interaction, i.e., the proximity of other sugar moiety and a contact with it, has a visible effect on the correspondent Raman/ROA spectra. Figure 13 shows a comparison of calculated Raman/ROA spectra of methyl-β-D-glucose in infinitely diluted solution (i.e., a single molecule, black), when in a close proximity of other sugar unit (stacked, red), and experimental data (blue) for comparison. The outcome is that both spectra are nearly identical in the studied frequency range. This is showing that the Raman/ROA spectroscopies cannot readily identify potential intermolecular interactions in the used spectral range unless there is a major structural change, e.g., as shown in section 3.2, which is easily visible. Figure S11 shows that including empirical dispersion interactions in the DFT calculations has no effect on the outcome of the calculations.

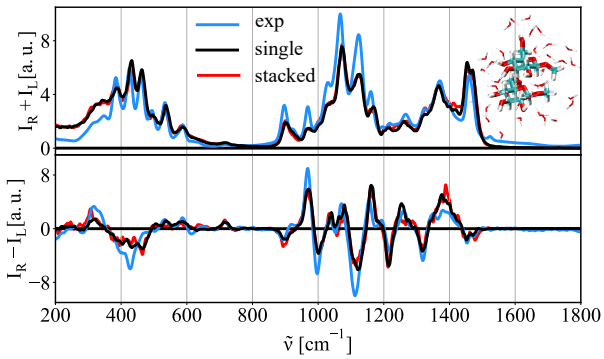

Figure 13: Calculated Raman/ROA spectra of methyl-β-glucose at infinite dilution, i.e., single molecule (black), of two interacting methyl-β-glucose sugar moieties (red, representative snapshot in the inset). In blue the experimental spectra at 1 M concentration for comparison.

4 Discussion

Conformational behaviour of saccharides in solution is still not well understood, mainly due to the lack of readily available and suitable structural techniques. In the current paper we suggest the use of Raman/ROA spectroscopies, together with computer simulations to determine structural properties of saccharides in solution. We show how the Raman and mostly ROA spectra when interpreted using appropriate computational methods resolve various structural features of sugars, such as the glycosidic linkage conformation, the ring puckering, the anomeric equilibrium, and we also reveal whether they provide any information on intermolecular interactions. Furthermore, they can be effectively used to quantitatively resolve more complex sugar mixtures.

To study the glycosidic linkages, we simulated the Raman and ROA spectra of four disaccharides (trehalose and three methyl mannobioses(1→2, 1→3, and 1→6 linked)) using ensemble generated either by unbiased MD simulations, or by advanced sampling techniques. Just briefly, simulations of M12 and M13 provided similar conformational profiles with the global minimum at $\{\phi_1, \phi_2\}$ of (-45, 35) or (-45, 22), respectively, which agrees with previous studies.[40,54,55] On the contrary, the simulation of M16 reveals that the global minimum is the extended structure with $\{\phi_1, \phi_2\}$ of (180, 180). All but the 1→6 mannobiose provide a good

agreement with experimental spectra when using unbiased MD ensemble. This indicates the reliability of the Glycam force field used to describe the 1→1, 1→2, and 1→3 glcyosidic bonds in MD simulations. On the other hand, the 1→6 linked mannobioses' spectra were reproduced with lower quality, suggesting inaccurate ensemble provided by the MD simulations. Using advanced sampling techniques we were able to significantly improve on the quality of the simulation spectra, and therefore, to better identify the true ensemble. An important lesson learned is that the Raman/ROA spectra of disaccharides change significantly with simple variations in glycosidic dihedral angles. The changes occur mostly in the ROA spectra and below 1200 cm^{-1}. However, additional small changes can be detected in the whole considered spectral range. Moreover, we found that there are no specific bands in both Raman and ROA spectra associated to particular glycosidic bond orientations on the studied molecules. This lack of correspondence was also reported on mannobioses,[16] suggesting that this a the common scenario. However, we show that the spectral changes produced by different glycosidic angles are significant enough to determine populations of conformers in a given disaccharide. Therefore, despite single bands cannot be assigned to particular conformations, coupling experimental spectra to computer simulations can yield preferred populations in solution. Our calculations also show that the error in determining the conformer populations using the Raman/ROA data is substantial, reaching up to 30%. Nevertheless, we also find a similar accuracy for predictions based on traditional NMR experiments that usually use $^3J_{CH}/^3J_{HH}$ coupling constants across glycosidic bonds. Moreover, to fully describe conformational preferences of 1→6 linked mannobiose, more coupling constant across the glycosidic link would be needed, e.g., $^3J_{CH}/^3J_{CC}$ or $^2J_{CC}$,[56,57] and not only the easily available $^3J_{HH}$. However, this is only achievable with isotopically ^{13}C labeled compounds, which are usually hard to synthesize. Our approach does not require any labelling, and using the Raman/ROA spectroscopies we are able to decompose the spectra into local-conformers (average of configurations around a conformation) even without the labeled compound. This could for example help in computational studies relying in con-

formational preferences on C-disaccharides, where the two saccharide units are joined with a methylene bridge, where little structural information is still known.[58] Overall, our data praises the Glycam force field that accurately describes the rotation around 1→1, 1→2, and 1→3 glycosidic bonds, while finding that the 1→6 linkage requires refinement.

We also tackle a basic structural feature of saccharides, their puckering, using methyl-β-D-glucuronic acid. We prepared models of the monosaccharide in various puckering conformations and subsequently calculated the Raman/ROA spectra for each puckering conformation. The variations in the Raman and ROA spectra produced by change of the puckering conformation were significant enough to assign unequivocally the $^4C_1$ puckering conformation as the dominant conformation of methyl-β-D-glucuronic acid, which is in agreement with NMR observations.

We also show that the `hybrid` calculation protocol, originally developed on monosaccharides,[10] is also applicable to larger saccharides. With the use of MD and advanced sampling methods, we explored and described the phase space of a raffinose trisaccharide. This a a challenging saccharide composed of three different monosaccharides and two likages. Still, our method can accurately reproduce experimental spectra, showing the applicability of the simulation protocol for trisaccharides at a reasonable computational cost. Still the overall accuracy is dependent on the simulation, and future improvements of the computational protocol are desirable.

As reducing sugars exist in solution as a balanced mixture of different forms, we addressed the sensitivity of the approach to discriminate such individual forms and to estimate their occurrence. Therefore, we explored the fundamental anomeric equilibrium for glucose and two glucose derivatives. Both anomers of glucose, glucuronic acid, and N-acetyl glucosamine were prepared and corresponding Raman/ROA spectra were calculated. By the best fitting to experimental spectra, we were able to successfully predict the value of the α/β anomeric equilibrium constant with an error of ∼10%. The results match with NMR-based results, which is a more commonly used method for estimating the anomeric ratio. Further analysis

showed that the best fit to the Raman/ROA data is associated with 10-15% uncertainty error, which is slightly smaller than the error found when determining populations of disaccharide conformers. This smaller error might be attributed to the fact that puckering changes the whole geometry of a molecule, leading to very distinct Raman/ROA spectra.

It was already shown that the spectral decomposition approach can be applied to estimate the composition of saccharide mixture consisting of two different diastereomers.[13] This has also been applied successfully to other sterols and complex rings mixtures in chlororoform.[59,60] Here we go even further to test the sensitivity limit of the approach, and estimate the ratio of very similar sugars such as a monosaccharide and its amine. We decompose experimental spectra of mixtures of methyl-β-glucose (MeGlc) and methyl-β-N-glucosamine (MeGlcNAc). Note that MeGlcNAc differs from MeGlc only by substitution of the hydroxyl at carbon C2, and the achiral amide group should not generate an intrinsic ROA signal. We prepared samples of pure sugars, as well as their mixtures ($x_{\mathrm{MeGlc}} = 0.25/0.50/0.75$). Then we experimentally recorded their Raman/ROA spectra and subsequently decomposed the spectra finding the ratio of molecular moieties in the target solution using either simulated or experimental spectra of the pure compounds. When using computed spectra, we were able to predict the compositions of mixtures within ∼0-22% errors. These error naturally resembles the error for the anomeric equilibrium values, as both procedures are in essence identical. When using experimental spectra of pure substances for the decomposition, we obtain better fits, i.e. $\sim 0 - 9\%$ errors. However, when we evaluated uncertainties of the bets fit values for both cases, they are equally footed. Lastly, this decomposition method is potentially applicable to other moieties, not just saccharides.

Saccharides in biology are often found in crowded environments. Therefore, we probe the sensibility of the ROA/Raman spectra to study intermolecular interactions usual in such environments. For this we calculate the spectra of a sugar moiety in the close proximity to another. When comparing a model system of a single methyl-β-glucose (MeGlc) with two MeGlc in a close proximity (average COM-COM distance 0.46 nm), we see that in-

teraction has no effect on the Raman/ROA spectra in the studied 200-1800 cm^{-1} spectral range. Therefore, any direct interaction using Raman/ROA cannot be easily seen, unless the interaction/packing results in a change of the geometry/structure of given saccharides upon the interaction. The result is in agreement with previous works in the literature, where it was concluded that the interaction regions is likely too small and that there could be some interaction specific bands present below 200 cm^{-1}.[61,62]

Overall, in our work we demonstrate the power of coupling of the Raman/ROA spectroscopies with computer simulations. This allows to extract numerous atomistic details that would be otherwise hardly obtainable. We also quantified the expected reliability of our structural predictions, which despite having substantial associated error up to 41%, this errors are comparable standard techniques based on NMR spin-spin coupling constants.

5 Conclusions

Simulation of Raman/ROA spectra of saccharides is becoming well established field. However, its applications in structure determination, in particular for large molecules, is yet largely unexplored. In this paper we show that combining ROA/Raman spectra with advanced sampling MD techniques and QM calculations one can get vast amounts of structural information about saccharides in solution. Taking advantage of the sensitivity of ROA/Raman to structural changes one can detect even small changes in rotation around the glycosidic bonds. This also applies for the puckering changes of the sugar rings. We show that the used simulation protocol can also applicable to a larger moiety, e.g., raffinose trisaccharide, at a reasonable computational cost. We can also apply the methods to mixtures. We, for example, managed to extract α/β anomeric equilibrium values of monosaccharides. But the method can be extended to any sort of mixture, even for those structurally similar. We shown that Raman/ROA spectra in the $200 - 1600$ cm^{-1} range are insensitive to direct contacts unless they produce structural changes. Lastly, the method used here allows to

estimate its inherent error when determining populations, which in this study can be as high as 41%, and comparable with other simulation and experimental methods actively used. All investigated cases show that coupling computer simulations with experimental Raman/ROA spectroscopies is a powerful structural determination tool.

Acknowledgement

HMS, JK, and VP acknowledges support from the Czech Science Foundation (19-19561S). HMS acknowledges the European Regional Development Fund OP RDE (project ChemBioDrug no. CZ.02.1.01/0.0/0.0/16_019/0000729) for computational resources. We thank Prof. Pavel Jungwirth and Prof. Petr Bouř for fruitful discussions.

5.1 Keywords

Raman, ROA, saccharides, carbohydrates, sugars, molecular dynamics, simulation, DFT, structure, aqueous solution

Supporting Information Available

Supplementary information file. The file includes a thorough description of the simulated systems and used methods. Moreover, it includes additional and supporting data for the main paper.

References

(1) Vuorio, J.; Škerlová, J.; Fábry, M.; Veverka, V.; Vattulainen, I.; Řezáčová, P.; Martinez-Seara, H. N-glycosylation can selectively block or foster different receptor–ligand binding modes. *Scientific Reports* **2021**, *11*.

(2) Tarbell, J. M.; Cancel, L. M. The glycocalyx and its significance in human medicine. *Journal of Internal Medicine* **2016**, *280*, 97–113.

(3) Tang, Y.; Cheng, F.; Feng, Z.; Jia, G.; Li, C. Stereostructural elucidation of glucose phosphorylation by Raman optical activity. *The Journal of Physical Chemistry B* **2019**, *123*, 7794–7800.

(4) Sodhi, H.; Panitch, A. Glycosaminoglycans in tissue engineering: A review. *Biomolecules* **2020**, *11*, 29.

(5) Imberty, A.; Pérez, S. Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chemical Reviews* **2000**, *100*, 4567–4588.

(6) Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. *Chemical Reviews* **2002**, *102*, 371–386.

(7) Rudolph, M. G.; Speir, J. A.; Brunmark, A.; Mattsson, N.; Jackson, M. R.; Peterson, P. A.; Teyton, L.; Wilson, I. A. The crystal structures of Kbm1 and Kbm8 reveal that subtle changes in the peptide environment impact thermostability and alloreactivity. *Immunity* **2001**, *14*, 231–242.

(8) Peréz, S.; Mouhous-Riou, N.; Nifant'ev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. Crystal and molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis x trisaccharide. *Glycobiology* **1996**, *6*, 537–542.

(9) Barron, L. D.; Bogaard, P.; Buckingham, A. D. Raman scattering of circularly polarized light by optically active molecules. *Journal of the American Chemical Society* **1973**, *95*, 603–605.

(10) Palivec, V.; Kopecký, V.; Jungwirth, P.; Bouř, P.; Kaminský, J.; Martinez-Seara, H. Simulation of Raman and Raman optical activity of saccharides in solution. *Physical Chemistry Chemical Physics* **2020**, *22*.

(11) Palivec, V.; Michal, P.; Kapitán, J.; Martinez-Seara, H.; Bouř, P. Raman optical activity of glucose and sorbose in extended wavenumber range. *ChemPhysChem* **2020**, *21*.

(12) Mutter, S. T.; Zielinski, F.; Johannessen, C.; Popelier, P. L.; Blanch, E. W. Distinguishing epimers through Raman optical activity. *Journal of Physical Chemistry A* **2016**, *120*, 1908–1916.

(13) Melcrová, A.; Kessler, J.; Bouř, P.; Kaminský, J. Simulation of Raman optical activity of multi-component monosaccharide samples. *Physical Chemistry Chemical Physics* **2016**, *18*, 2130–2142.

(14) Kaminský, J.; Kapitán, J.; Baumruk, V.; Bednárová, L.; Bouř, P. Interpretation of Raman and Raman optical activity spectra of a flexible sugar derivative, the gluconic acid anion. *Journal of Physical Chemistry A* **2009**, *113*, 3594–3601.

(15) Rüther, A.; Forget, A.; Roy, A.; Carballo, C.; Mießmer, F.; Dukor, R. K.; Nafie, L. A.; Johannessen, C.; Shastri, V. P.; Lüdeke, S. Unravelling a direct role for polysaccharide β-Strands in the higher order structure of physical hydrogels. *Angewandte Chemie International Edition* **2017**, *56*, 4603–4607.

(16) Pendrill, R.; Mutter, S. T.; Mensch, C.; Barron, L. D.; Blanch, E. W.; Popelier, P. L. A.; Widmalm, G.; Johannessen, C. Solution structure of mannobioses unravelled by means of Raman optical activity. *ChemPhysChem* **2019**, *20*, 695–705.

(17) Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L. A.; Blanch, E. W. Calculation of Raman optical activity spectra of methyl-β-D-glucose incorporating a full molecular dynamics simulation of hydration effects. *Journal of the American Chemical Society* **2011**, *133*, 4991–4997.

(18) Zielinski, F.; Mutter, S. T.; Johannessen, C.; Blanch, E. W.; Popelier, P. L. A. The Raman optical activity of β-D-xylose: where experiment and computation meet. *Physical chemistry chemical physics : PCCP* **2015**, *17*, 21799–809.

(19) Cheeseman, J. R.; Frisch, M. J. Basis set dependence of vibrational Raman and Raman optical activity intensities. *Journal of Chemical Theory and Computation* **2011**, *7*, 3323–3334.

(20) Zuber, G.; Hug, W. Rarefied Basis Sets for the calculation of Optical Tensors. 1. The Importance of gradients on hydrogen atoms for the Raman scattering tensor. *The Journal of Physical Chemistry A* **2004**, *108*, 2108–2118.

(21) Bouř, P.; Keiderling, T. A. Partial optimization of molecular geometry in normal coordinates and use as a tool for simulation of vibrational spectra. *The Journal of Chemical Physics* **2002**, *117*, 4126–4132.

(22) Cremer, D.; A. Pople, J. General definition of ring puckering coordinates. *Journal of the American Chemical Society* **1975**, *97*, 1354–1358.

(23) ACDLabs, ChemSketch 2015.2.5.

(24) Inkscape. http://www.inkscape.org/. 2020.

(25) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

(26) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banáš, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; Capelli, R.; Carloni, P.; Ceriotti, M.; Cesari, A.; Chen, H.; Chen, W.; Colizzi, F.; De, S.; De La Pierre, M.; Donadio, D.; Drobot, V.; Ensing, B.; Ferguson, A. L.; Filizola, M.; Fraser, J. S.; Fu, H.; Gasparotto, P.; Gervasio, F. L.; Giberti, F.; Gil-Ley, A.; Giorgino, T.; Heller, G. T.;

Hocky, G. M.; Iannuzzi, M.; Invernizzi, M.; Jelfs, K. E.; Jussupow, A.; Kirilin, E.; Laio, A.; Limongelli, V.; Lindorff-Larsen, K.; Löhr, T.; Marinelli, F.; Martin-Samos, L.; Masetti, M.; Meyer, R.; Michaelides, A.; Molteni, C.; Morishita, T.; Nava, M.; Paissoni, C.; Papaleo, E.; Parrinello, M.; Pfaendtner, J.; Piaggi, P.; Piccini, G. M.; Pietropaolo, A.; Pietrucci, F.; Pipolo, S.; Provasi, D.; Quigley, D.; Raiteri, P.; Raniolo, S.; Rydzewski, J.; Salvalaglio, M.; Sosso, G. C.; Spiwok, V.; Šponer, J.; Swenson, D. W.; Tiwary, P.; Valsson, O.; Vendruscolo, M.; Voth, G. A.; White, A. Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods* **2019**, *16*, 670–673.

(27) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *Journal of Computational Chemistry* **2008**, *29*, 622–655.

(28) Izadi, S.; Onufriev, A. V. Accuracy limit of rigid 3-point water models. *The Journal of Chemical Physics* **2016**, *145*, 074501.

(29) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters* **2008**, *100*, 020603.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Ren-

dell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian16 Revision B.01. 2016.

(31) Orozco, M.; Marchn, I.; Soteras, I.; Vreven, T.; Morokuma, K.; Mikkelsen, K. V.; Milani, A.; Tommasini, M.; Zoppo, M. D.; Castiglioni, C.; Aguilar, M. A.; Snchez, M. L.; Martn, M. E.; Galvn, I. F.; Sato, H. *Continuum solvation models in chemical physics*; John Wiley & Sons, Ltd: Chichester, UK, 2007; pp 499–605.

(32) Dapprich, S.; Komáromi, I.; Byun, K.; Morokuma, K.; Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *Journal of Molecular Structure: THEOCHEM* **1999**, *461-462*, 1–21.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(34) Nafie, L. *Vibrational optical activity: Principles and applications*; Wiley: Chichester, 2011.

(35) Ruud, K.; Helgaker, T.; Bouř, P. Gauge-origin independent density-functional theory calculations of vibrational Raman optical activity. *Journal of Physical Chemistry A* **2002**, *106*, 7448–7455.

(36) Säwén, E.; Massad, T.; Landersjö, C.; Damberg, P.; Widmalm, G. Population distribution of flexible molecules from maximum entropy analysis using different priors as background information: Application to the , ψ-conformational space of the α-(1→2)-linked mannose disaccharide present in N- and O-linked glycoproteins. *Organic and Biomolecular Chemistry* **2010**, *8*, 3684–3695.

(37) Jensen, F. Basis set convergence of nuclear magnetic shielding constants calculated by density functional methods. *Journal of Chemical Theory and Computation* **2008**, *4*, 719–727.

(38) Sattelle, B. M.; Hansen, S. U.; Gardiner, J.; Almond, A. Free energy landscapes of iduronic acid and related monosaccharides. *Journal of the American Chemical Society* **2010**, *132*, 13132–13134.

(39) Mikkelsen, L. M.; Hernáiz, M. J.; Martín-Pastor, M.; Skrydstrup, T.; Jiménez-Barbero, J. Conformation of Glycomimetics in the Free and Protein-Bound State: Structural and Binding Features of the C-glycosyl Analogue of the Core Trisaccharide α-D-Man-(1→3)-[α-D-Man-(1→6)]-D-Man. *Journal of the American Chemical Society* **2002**, *124*, 14940–14951.

(40) Raich, I.; Lövyová, Z.; Trnka, L.; Parkan, K.; Kessler, J.; Pohl, R.; Kaminský, J. Limitations in the description of conformational preferences of C-disaccharides: The (1→3)-C-mannobiose case. *Carbohydrate research* **2017**, *451*, 42–50.

(41) Bouckaert, J.; Hamelryck, T. W.; Wyns, L.; Loris, R. The crystal structures of Man(alpha1-3)Man(alpha1-O)Me and Man(alpha1-6)Man(alpha1-O)Me in complex with concanavalin A. *The Journal of biological chemistry* **1999**, *274*, 29188–29195.

(42) Brisson, J. R.; Carver, J. P. Solution conformation of asparagine-linked oligosaccharides: .alpha.(1-2)-, .alpha.(1-3)-, .beta.(1-2)-, and .beta.(1-4)-linked units. *Biochemistry* **2002**, *22*, 3671–3680.

(43) Kim, H.-M.; Choi, Y.-J.; Lee, J.-H.; Jeong, K.-J.; Jung, S.-H. Conformational analysis of trimannoside and bisected trimannoside using aqueous molecular dynamics simulations. *Bulletin of the Korean Chemical Society* **2009**, *30*, 2723–2728.

(44) Kan, Z.; Yan, X.; Ma, J. Conformation dynamics and polarization effect of α,α-trehalose

in a vacuum and in aqueous and salt solutions. *The journal of physical chemistry. A* **2015**, *119*, 1573–1589.

(45) Bock, K.; Defaye, J.; Driguez, H.; Bar-Guilloux, E. Conformations in solution of α,α-trehalose, α-D-glucopyranosyl α-D-mannopyranoside, and their 1-thioglycosyl analogs, and a tentative correlation of their behaviour with respect to the enzyme trehalase. *European journal of biochemistry* **1983**, *131*, 595–600.

(46) Choi, Y.; Cho, K. W.; Jeong, K.; Jung, S. Molecular dynamics simulations of trehalose as a 'dynamic reducer' for solvent water molecules in the hydration shell. *Carbohydrate Research* **2006**, *341*, 1020–1028.

(47) Olsson, U.; Säwén, E.; Stenutz, R.; Widmalm, G. Conformational flexibility and dynamics of two (1→6)-linked disaccharides related to an oligosaccharide epitope expressed on malignant tumour cells. *Chemistry - A European Journal* **2009**, *15*, 8886–8894.

(48) Zhu, Y.; Zajicek, J.; Serianni, A. S. Acyclic forms of [1-^{13}C]aldohexoses in aqueous solution: Quantitation by ^{13}C NMR and deuterium isotope effects on tautomeric equilibria. *Journal of Organic Chemistry* **2001**, *66*, 6244–6251.

(49) Bubb, W. A. NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts in Magnetic Resonance* **2003**, *19A*, 1–19.

(50) Franks, F.; Lillford, P. J.; Robinson, G. Isomeric equilibria of monosaccharides in solution. Influence of solvent and temperature. *Journal of the Chemical Society, Faraday Transactions 1* **1989**, *85*, 2417.

(51) Bezabeh, T.; Ijare, O. B.; Albiin, N.; Arnelo, U.; Lindberg, B.; Smith, I. C. P. Detection and quantification of D-glucuronic acid in human bile using ^{1}H NMR spectroscopy: relevance to the diagnosis of pancreatic cancer. *Magnetic Resonance Materials in Physics, Biology and Medicine* **2009**, *22*, 267–275.

(52) Moreau, B.; Lognay, G.; Blecker, C.; Destain, J.; Gerbaux, P.; Chéry, F.; Rollin, P.; Paquot, M.; Marlier, M. Chromatographic, spectrometric and NMR characterization of a new Set of glucuronic acid esters synthesized by lipase. *Biotechnol. Agron. Soc. Environ.* **2007**, *11*, 9–17.

(53) Liu, F. C.; Su, C. R.; Wu, T. Y.; Su, S. G.; Yang, H. L.; Lin, J. H. Y.; Wu, T. S. Efficient ^{1}H-NMR quantitation and investigation of N-Acetyl-D-glucosamine (GlcNAc) and N,N'-diacetylchitobiose $(GlcNAc)_2$ from chitin. *International Journal of Molecular Sciences* **2011**, *12*, 5828–5843.

(54) Espinosa, J.-F.; Bruix, M.; Jarreton, O.; Skrydstrup, T.; Beau, J.-M.; Jiménez-Barbero, J. Conformational differences between C- and O-glycosides: The α-C-mannobiose/α-O-mannobiose case. *Chemistry - A European Journal* **1999**, *5*, 442–448.

(55) Naidoo, K. J.; Denysyk, D.; Brady, J. W. Molecular dynamics simulations of the N-linked oligosaccharide of the lectin from Erythrina corallodendron. *Protein Engineering Design and Selection* **1997**, *10*, 1249–1261.

(56) Zhang, W.; Turney, T.; Meredith, R.; Pan, Q.; Sernau, L.; Wang, X.; Hu, X.; Woods, R. J.; Carmichael, I.; Serianni, A. S. Conformational Populations of β-(1→4) O-glycosidic linkages using redundant NMR J-couplings and circular statistics. *Journal of Physical Chemistry B* **2017**, *121*, 3042–3058.

(57) Zhang, W.; Meredith, R.; Pan, Q.; Wang, X.; Woods, R. J.; Carmichael, I.; Serianni, A. S. Use of circular statistics to model αMan-(1→2)-αMan and αMan-(1→3)-α/βMan O-glycosidic linkage conformation in ^{13}C-labeled disaccharides and high-mannose oligosaccharides. *Biochemistry* **2019**, *58*, 546–560.

(58) Alonso-Gil, S.; Parkan, K.; Kaminský, J.; Pohl, R.; Miyazaki, T.; Alonso-Gil, S.; Miyazaki, T.; Parkan, K.; Kaminský, J.; Pohl, R. *Unlocking the hydrolytic mechanism*

of GH92 α-1,2-mannosidases: computation inspires using C-glycosides as Michaelis complex mimics; 2021.

(59) Bogaerts, J.; Desmet, F.; Aerts, R.; Bultinck, P.; Herrebout, W.; Johannessen, C. A combined Raman optical activity and vibrational circular dichroism study on artemisinin-type products. *Physical Chemistry Chemical Physics* **2020**, *22*, 18014–18024.

(60) Koenis, M. A.; Tiekink, E. H.; van Raamsdonk, D. M.; Joosten, N. U.; Gooijer, S. A.; Nicu, V. P.; Visscher, L.; Buma, W. J. Analytical chemistry on many-center chiral compounds based on vibrational circular dichroism: Absolute configuration assignments and determination of contaminant levels. *Analytica Chimica Acta* **2019**, *1090*, 100–105.

(61) Yamamoto, S.; Kaminský, J.; Bouř, P. Structure and vibrational motion of insulin from Raman optical activity spectra. *Analytical Chemistry* **2012**, *84*, 2440–2451.

(62) Kaminský, J.; Horáčková, F.; Biačková, N.; Hubáčková, T.; Socha, O.; Kubelka, J. Double hydrogen bonding dimerization propensity of aqueous hydroxy acids investigated using vibrational optical activity. *The Journal of Physical Chemistry B* **2021**, *125*, 11350–11363.

Graphical TOC Entry

TOC