Reply to Reviewers - Quantitative Structural Determination of Monosaccharides in Solution from Raman and Raman Optical Activity Spectra

Vladimír Palivec,[†,‡] Vladimír Kopecký Jr.,[¶] Pavel Jungwirth,[§] Petr Bouř,[§] Jakub Kaminský,[*,§,‖] and Hector Martinez-Seara[*,§]

†*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo náměstí 542/2, CZ166 10, Prague 6, Czech Republic*

‡*PhD. Student of University of Chemistry and Technology Prague*

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

§*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, CZ166 10, Prague, Czech Republic*

‖*Gilead Sciences & IOCB Research Center*

E-mail: jakub.kaminsky@uochb.cas.cz; hseara@gmail.com

We believe that the discussion that emerged during the rebuttal together with the manuscript has a value on its own. The concise and relevant comments from the reviewers have improved, in our opinion, significantly the scientific quality of this work. Moreover, we believe that many of the questions that arose during the revision might be of interest to future readers too.

Therefore, here below you can find detailed questions that were raised by the reviewers, together with corresponding answers. The questions are denoted as RxQx, e.g. R1Q1 - question 1 from the reviewer 1. Furthermore, a blue text marks a comment from a reviewer, while a black text is the author's reply.

1 List of major changes to the text

Before we provide detailed point by point answer to the comments by the reviewers we summarize here major changes in the main text with respect to the previous version:

1. We have abandoned the use of harmonic constrains in the hybrid approach. The partial optimization approach that we now propose already ensures a limited contraction of the phase space during the minimization.

2. We have replaced our error function (δ) with the the widely used overlap integral method (S) for quality spectra assessment. The S value bears the same information as our previously proposed error function (δ) while it is easier to understand. Moreover, contrary to error function, the overlap integral method can be used to compare the spectra of different compounds.

3. We have included the spectra of individual anomers for the reducing sugars, see Figure 7 to ESI. This piece of information was previously missing and turned out to be critical for the message of this manuscript.

4. We added methyl-β-glucopyranoside monosaccharide for completion.

5. We have added performance data for the widely used rDPS basis set.

6. We have added performance data comparing different basis sets (6-31G*,6-311++G**, aug-cc-pvtz, rDPS), DFT functionals (B3LYP, M06L, CAM-B3LYP, ωB97XD), and

number of optimization steps prior to spectra calculations, for all molecules. The summary of the results can be found in the main text, see Figures 6,7, and 8.

7. To accurately answer reviewers' concerns about the optimization procedures, we added three additional modeling protocols `hybrid_w`, `hybrid_fo`, and `MM_w`. The performance of these protocols has been evaluated and summarized in the main text, see Figure 5.

8. We have added two modeling approaches using QM described water molecules, `hybrid_QM` and `hybrid_QM_fo`. The latter performs a full optimization. See new Figure 15 in ESI.

9. To clarify the advantages and nature of the proposed scaling function, we have studied additional setups where either "no scaling", or the alternative "single factor scaling" have been compared side by side with "our proposed scaling" method, see Figures 16 and 17 in ESI.

1.1 Additional modeling approaches

One of the main criticism of our work concerned the the simulation protocol. To address this issue, We have added 5 additional computational approaches summarized in Table R1. They link our work to approaches available in the literature. This comparison shows that our procedure is coherent, physically sound, and with an excellent quality/cost ratio. The additional protocols consist of slight modifications of `hybrid` approach - `hybrid_w` and `hybrid_fo`. In the `hybrid_w`, all water molecules are frozen and the central sugar is fully optimized until gradient criteria are reached. In the `hybrid_fo`, neither the sugar nor the water molecules in the surrounding 3 Å layer are unrestrained and the optimization is run until gradient convergence criteria are reached. In both simulations, all other simulation parameters remain unchanged as in the `hybrid` approach. Moreover, the `MM` variant of `hybrid_w` was also added (called `MM_w`). Additionally, two approaches (`hybrid_QM` and `hybrid_QM_fo`) were studied to test the need of explicit QM water molecules. The `hybrid_QM` is `hybrid` approach variant

using QM water molecules, while the `hybrid_QM_fo` is the `hybrid_fo` variant also using QM water molecules.

Table 1: Additional calculation protocols used in this work.

Protocol	Description
`hybrid_w`	• Water molecules closer than 3 Å as MM water molecules • CPCM • B3LYP/6-311++G**, MM water molecules • all water molecules frozen, optimization until gradient criteria reached
`hybrid_fo`	• Water molecules closer than 3 Å as MM water molecules • CPCM • B3LYP/6-311++G**, MM water molecules • optimization until gradient criteria reached
`MM_w`	• Water molecules closer than 12 Å as MM water molecules • B3LYP/6-311++G**, MM water molecules • all water molecules frozen, optimization until gradient criteria reached
`hybrid_QM`	• Water molecules closer than 3 Å as QM water molecules • CPCM • B3LYP/6-311++G**(water molecules: 6-311++G**) • 10 optimization steps
`hybrid_QM_fo`	• Water molecules closer than 3 Å as QM water molecules • CPCM • B3LYP/6-311++G**(water molecules: 6-311++G**) • optimization until gradient criteria reached

CPCM = conductor-like polarizable continuum model

2 Detailed replies to the Reviewers

In the following section we will provide answers to each of the questions raised during the revision.

2.1 Comments to Reviewer 1 (R1)

Comments to the Author 'Simulation of Raman and Raman optical activity of saccharides in solution: Reaching the experimental accuracy'

This is nice manuscript, which is not a review but has many virtues of a review article. It really defines the state-of-the-art of the interpretation of Raman and ROA spectra of carbohydrate molecules in water solution through the combination of MD and DFT calculations. Beside other virtues, presenting the comparison of various computational approaches is appreciable, since there is not unanimous consensus in the literature about the best method. Another good point is keeping in mind that experimental errors may influence comparison with calculations, and thus some experimental features to be predicted by calculations should not be overestimated. The work proceeds by benchmarking the method on a limited set of molecules (glucose, glucuronic acid, N-acetyl-glucosamine and methyl-Beta-derivatives thereof), but one may easily foresee that the method proposed here can be applied to many other sugar molecules. Having said this, I have a few comments/questions and few minor points:

We thank the reviewer for these kind words and sharing our view on the importance and potential impact of this work. In the following subsections, we address all the questions raised in the review:

2.1.1 R1Q1

While I find that for example the adjustable 'scaling' factor 'fi' in equation (2) is very useful and new, I have a feeling that the Authors have pushed their way a bit too far in introducing adjustable parameters and in a fitting procedure. Could they please report some convincing statements to prove that I had the wrong impression? To this instance, I do not understand the title of paragraph 2.6, namely "Adjustments to the Calculated Spectra"; possibly the Authors meant "Adjustments to the Experimental Spectra", or "Adjustments of the Calculated Spectra" or even "Adjustments of the Calculated Frequencies?"

The proposed scaling function (ϕ) shifts frequencies, while the ρ parameter scales all intensities equally so they match in average the experimental values. The (ϕ) scaling function is an improved form of the typical single scaling factor used commonly in the literature. Further information about the reasons behind the proposed functional form is found in the answer to question R2Q8. The intensity scaling (ρ) is necessary for Raman and ROA, where absolute intensities cannot be measured, as discussed, for example in L. D. Barron, 'Molecular light scattering and optical activity', 2004.

Finally, we agree that the name of the section was confusing so we have changed it to "Adjustments of the Calculated Spectra" as suggested by the reviewer.

2.1.2 R1Q2

Since on Page 7 the Authors state: 'Raman and Raman optical activity (ROA) spectroscopies appear as powerful tools for gaining valuable insight into the structure of saccharides when coupled with computer simulations', wouldn't it be a good idea to document somewhat the conformational properties from the MD simulations, e.g. by giving the dihedral angle behaviors as function of time? This may not only clarify the conformational panorama about the differences in the MD and DFT analysis, but could also give some hints about the question raised at page 8 on either possible incomplete ensemble or not adequate force field.

We understand perfectly the suggestion of the reviewer. In fact, we are working on a second manuscript focused on the applicability of the developed technique. The next manuscript covers in detail applications of the presented technique when dealing with structural behavior of sugars of different lengths in the solution. We also cover the performance of currently available MD force fields as suggested. The present methodological paper aims at being thorough and tests in detail all sensible variants to finally select an efficient method which has resulted into a very long paper and supplementary material already. If we would add the additional applicability data for the structural data of sugars in solution contained in the second article the length would increase significantly reducing the readability and impact of

the current manuscript. Moreover, considering the significant amount of comments raised during the manuscript revision process it seems unfeasible to reduce the size of the current manuscript to make room for the application data.

2.1.3 R1Q3

Sticking to page 8, I find that some citation about enhanced sample might be in order.

In the current manuscript we did not use any enhanced sampling technique except for the graphs provided in the reply R2Q5. Therefore a reference to these techniques in the main text seems out of place. All sugars we explore can sample all relevant minima during the 500 ns simulations we performed, see R2Q5. For other monosaccharides (e.g., iduronic acid) or larger saccharides (e.g., 16-mannobiose) enhanced sampling techniques might be required for efficient conformation sampling of all relevant equilibrium structure and gathering their relative weights. This is, however, beyond the scope of the current manuscript.

2.1.4 R1Q4

ROA (as well as VCD) started by examining first the CH-stretching region. Is the method presented by the Authors applicable also to that region? Why did the Authors not consider that experimental region? Indeed I am asking this not only because I kind-of-remember that there are numerous reports of ROA spectra of simple carbohydrates, but also because there is nice evidence that in the CH-stretching Raman spectra of most aldohexoses, the phenomenon of back-donation from oxygen lone pairs to the CH-bond has a major influence on the CH-stretching frequency values. Of course this phenomenon could interfere a lot with the interaction of carbohydrates with the solvent.

Unlike the first custom made instruments, the commercial ones, one of which we are using, cannot measure beyond 2450 cm^{-1}. We also do not think that our method is straightforwardly applicable for CH stretching, because these vibrations are strongly anharmonic and tightly packed/unresolved. Of course, their spectra would depend on the structure, too,

and we are currently looking into this by collaborating with a lab that can measure these signals, but this appears much more complex than the 200-1800 cm^{-1} region analyzed in the present manuscript. Having said this, we could argue that the majority of molecular vibrations are comprised already in our region of study. Therefore, future involvement of the CH stretching most probably would not change much the quantitative information, such as relative conformer ratios, obtained by the current analysis.

2.1.5 R1Q5

Minor points:(i) in the Abstract the sentence 'However, the Raman and ROA spectra of saccharides are challenging to interpret and model due to their flexibility and polarity' is not well constructed, since, as it is now, it seems that Raman and ROA spectra are flexible and polar. (ii) Page 2, first column: please correct: 'where hundreds of explicit water molecules were employed within in the monosaccharide solvation model' to : 'where hundreds of explicit water molecules were employed within the monosaccharide solvation sphere'. (iii) Page 6: just after equation (6) please define that x is equal to either to Raman or ROA (it is obvious, but...)

We corrected these mistakes.

2.1.6 R1Q6 (After revision)

A quite minor point: Section 2.3 apparently is composed of several sub-sections: All of them were numbered as 2.3.1: how come?

In the main text the numbering is now as it should be.

2.2 Comments to Reviewer 2 (R2)

Comments to the Author

Reviewer report for manuscript ID: CP-ART-04-2019-002429 titled "Simulation of Raman and Raman optical activity of saccharides in solution: Reaching the experimental accuracy".

The current manuscript describes the development of a computational protocol for the calculation of Raman and Raman optical activity (ROA) predicted spectra of saccharides. Applying a combined molecular dynamics (MD) and quantum mechanical molecular mechanical (QMMM) approach, the authors claim to have achieved high accuracy and efficiency with this protocol, employing a number of monosaccharides as bench marks of this. A novel scaling procedure is also introduced.

The manuscript presents a lot of computational work, with a large amount of impressively looking data, written in generally very good English (the main grammatical issue being not always using "spectrum" and "spectra" in the right tense, but this is a minor issue). Unfortunately, the underlying scientific approach is not as novel as the authors claim and furthermore very flawed, which can only lead me to recommend rejection of the manuscript. I will elaborate below.

We thank the reviewer for thoroughly reading the manuscript and providing relevant comments. Thanks to this, in the present revision, we have significantly improved, simplified and clarified the content and the proposed method for Raman and ROA spectra calculation. For example, we now use integral method (S) to evaluate the resulting spectra. We also removed the harmonic constrains of our optimization method as they are not needed.

2.2.1 R2Q1

Throughout the manuscript, it is implied that the computational protocol is new, e.g. the combination of MD and QMMM calculations has not been applied to saccharides before and hence the protocol must be compared to a number of other approaches (dielectric background up to explicit QM water). The problem is that this is not the case: This approach was initially

developed by Cheeseman et al. in 2011 (reference 11 in the manuscript) and further refined in a number of articles by the Manchester computational group, e.g. Zielinski, Mutter et al. 2015/2016 (references 15-17 in the manuscript), where several aspects of structural aspects related to Raman/ROA spectroscopy, size of the water cluster etc. are discussed. This alone makes the entire bench marking aspect of this manuscript trivial and can only be justified if the authors somehow can show their approach has an added value to the field, which is not the case at the moment, in my opinion.

We thank the reviewer for the criticism. We agree that the approach developed by Cheeseman group in 2011 constitutes an important milestone in the development of the field. Also, that the later refinements by the Manchester group, many of which are used in this work, improved the overall quality of the spectra modeling and understanding in the field. We would like to note that those refinements are in turn based on a long gradual development done by many scientists. In this work, we introduce new concepts and paths never explored before:

1. We extensively tested the hybrid QM/MM water models. The approach itself is not completely new but the exhaustive evaluation of its performance and potential applicability is. Current studies are at best limited and often inconsistent. Here we show, for example, how the QM water treatment can be circumvented when studying sugars achieving a significant speed up of the computations while maintaining with a controllable accuracy lost.

2. We proposed a physically sound scaling function that allows the optimal overlapping of the experimental and computational spectra. This function improves the agreement with experimental data, i.e. higher overlap integral, by correcting known deficiencies of the methodology, see section S5. It also helps dramatically in eyeballing problematic/missing spectral features. Our scaling function captures in a simple functional form the known need for two different scaling factors for the low/high frequency region. Our function also ensures continuity of the spectra. This topic is extensively covered

in the answer to question R1Q1/R2Q8. There, we show that our scaling function helps improving significantly the overlap integral with respect previously published data.

3. We show that a partial optimization leads to an better spectral agreement in overall. Although counterintuitive, the reason is that the final weighted spectra must include the spectra of all accessible configurations in solution, which are not necessarily sitting exactly at or close to the geometry minimum. Excessive optimization leads to an incorrect ensemble, which is reflected by worse agreement with experimental data. Limiting optimization to 10 steps provides the best agreement with experiment(see Figure 6 and Figures S8 and S9 in the ESI). This partial optimization also significantly boosts the computational performance - compare performance of `hybrid` and `hybrid_fo` in Table 2. Furthermore, at least within the harmonic approximation, this partial optimization is physically well justified, see R2Q7.

4. We thoroughly discuss the importance of considering each of the anomeric structures of a sugar moiety for Raman and ROA spectra calculation explicitly. We show that this is not optional but a must as each anomer has a significantly distinct spectrum, see R2Q12-13 for details. With our data in hand, we see that previous publications fail to address this issue properly.

5. We suggest an optimized computational protocol widely applicable not only for sugars but also for many biomolecules in solution. Our protocol reduces by at least an order of magnitude the computational costs of previously used protocols without sacrificing accuracy, see Table 2.

6. Finally, we propose an alternative method to the 'Manchester protocol' to estimate the quality of MD force fields by direct comparison of the computed spectra by using the MD ensemble and the experimental one. Partial optimization ensures a limited contraction of the final phase space used in the spectra calculation.

It is important to stress that all these new or improved points have resulted in a substantially increased quality of the spectra compared with previously published spectra; moreover, this improvement is not just a lucky coincidence but it is the result of careful planning of physically sound optimizations.

I agree that all of these aspects are new. Maybe all that is needed is a sentence in the manuscript (in the introduction) highlighting these, as in "Here, we present an improved protocol for efficient calculation of Raman and ROA spectra, taking "all the improvements" into account."? That way, the new and improved aspects stands out and takes the focus away from "we developed a protocol" to "we developed a better one".

We have clarified this point at the end of the introduction following the suggestion of the reviewer.

2.2.2 R2Q2

Furthermore, the authors claim that the protocol can be applied to bigger systems, previously not studied in this manner. I must draw attention to two recent papers from the field: In 2017, Ruther et al (Angew. Chem. Int. Ed. 2017, 56, 4674-4678) studied the effect of carboxylation of agarose polysaccharides employing a combination of experimental and computational techniques, including Raman and ROA, MD and density functional theory (DFT) calculations. In this case, QMMM was not used and neither was solvation considered as the compounds studied forms elastic gels in water. Even so, traditional MD and DFT methods were applied to agarose hexamers, showing that applying these standard methods to much larger systems than a monosaccharide is quite straight forward.

In the revised manuscript, we have referenced to this pioneer work and properly acknowledge its existence. Nevertheless, we do not see why its existence invalidates our claim. Our paper clearly shows that using explicit water molecules is crucial. The paper in question was restricted to the implicit CPMC solvent model. Also, we show that proper sampling is required to achieve experimental accuracy. In the manuscript mentioned above, the MD

simulations used are very short. The referees of our present work are questioning whether our 500 ns equilibration is sufficient, and we only consider monosaccharides. In this earlier work from 2017, the authors performed 15 ns simulations when studying polysaccharides, which is clearly insufficient for proper configurational sampling. Finally, although their method could be applied 'straightforwardly' as the reviewer says, their agreement between the simulation and experiment is inferior that obtained by our methodology.

Adding the agarose paper, but also being critical of it, is crucial here. Yes, bigger carbohydrate systems have been simulated, but at way too low sampling rate and without taking the solvent properly into consideration. Thus, the comparison is inferior to the results presented in this manuscript. The sentence accompanying the reference added by the authors says that.

2.2.3 R2Q3

More importantly, Pendrill et al published a paper in 2019 (ChemPhysChem 2019, 20, 695-705) describing the detailed analysis of computational approaches (from gas phase to QM water) applied to the prediction of the Raman and ROA spectral patterns of mannobioses disaccharides. This paper uses overlap integral analysis to quantify the comparison between experimental and predicted spectra, calculated at different levels of theory, including QMMM, with input structures based on MD production runs (combining the protocol developed in Manchester with the MD expertise of the Stockholm saccharide group). This approach is very similar (though clearly not directly copied) to the approach adopted in this manuscript and the two should have been critically compared at least. Omitting referencing the ChemPhysChem paper is unfortunate in one further aspect: One of the conclusions of this paper is the normal mode analysis of the disaccharides is much more complicated than for the monosaccharides studied previously, which means that the application of any computational protocol to larger saccharides may not be as useful as claimed in the current manuscript.

We thank the reviewer for making us aware of this new manuscript which indeed is relevant to this work. As mentioned by the reviewer, it discusses similar approaches to those proposed and improved here. We now compare some common aspects in the revision. For example, after critical comparison, we have adopted the overlap integral function and abandoned our previously proposed error function to evaluate the quality of the spectra in this revision. This function allows better comparison between systems. However, our main target — establishing a fast and accurate protocol, which is thoroughly tested and validated against experimental data — has also clear differences. For example, while Pendrill et al. use a single scaling factor, we use an adaptable scaling factor, see R1Q1/R2Q8 for implications of such change.

Finally, Pendrill et al. paper not only nicely discusses portability issues when increasing the complexity from monosaccharides to larger saccharides, but it it also hints that Raman/ROA calculations for disaccharides are possible. Our work improves the chances by streamlining the whole workflow. First, we minimize the sampling associated problems by using long molecular dynamics simulations, which can be also coupled to other advanced sampling methods. We also preserve the obtained equilibrium phase space by performing only a partial optimization that minimizes the contraction of the structural phase space during this required step (i.e., bonds and angles adjustments). Second, we produce high quality spectra specifically because of the usage of explicit water molecules in the first coordination shell coupled with an implicit water model and the use of the shifting function. Finally, since larger saccharides require larger computational resources, the tuning that achieved the best accuracy/cost is also crucial.

To back up our claims, our preliminary data on methyl (1-3) mannobiose (see Figure R1) show a significant improvement in terms of our Raman/ROA overlap integral (e.g., S=0.951/0.838 versus S=0.889/0.573 of Pendrill et al (Full QM/MM protocol)). This topic will be covered in detail in a follow up publication focus on the applications of the methodology presented in this work.

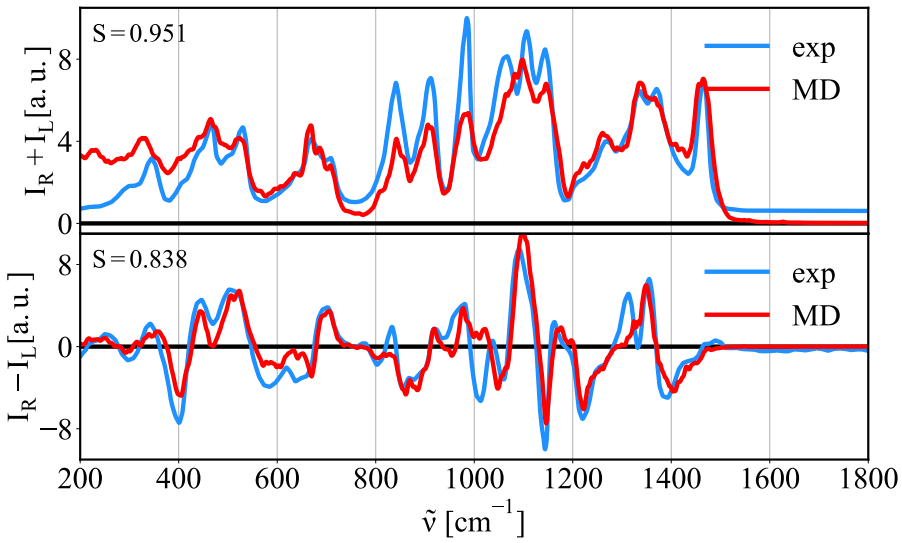

Figure 1: Calculated Raman/ROA spectra of methyl (1-3) mannobiose disaccharide with our developed hybrid protocol.

OK as is, but I actually think that the authors could be a bit more critical of the previously published papers if they wanted to, e.g. at least point out where the new and improved protocol presented here outperforms the "Manchester protocol" in the actual manuscript. But a more critical comparison could also be included in the follow up article, I suppose.

2.2.4 R2Q4

Thus, if the computational protocol presented here is not particularly novel, which aspects of the research can be seen as such? In my opinion, this would be the partial optimization scheme, the dynamic scaling function and the introduction of an error function. Unfortunately, there are big issues with all three of these.

Except for the error function that it is no longer used to evaluate the final quality of the spectra, in the reply to question R2Q1, we have carefully explained why this work is novel and how it improves the available methodology. Also, we explained there why the so-called shortcuts we took, which worry the referees, are physically sound.

Agreed

2.2.5 R2Q5

The geometry optimization approach: The 'hybrid' computational protocol is described by the authors as being very efficient, yet accurate in predicting the Raman and ROA spectra of the studied monosaccharides. The approach is based on a 500 ns production MD run, from where 100 snapshots are sampled semi-uniformly (and not randomly as claimed in the manuscript): Each snapshot is taken from a 5 ns block dispersed over the whole production run. This may lead to correct sampling of the conformational space, but I have a problem with this approach: Sattelle et al did several MD studies on monosaccharides in the early part of this decade (see, for example J. Am. Chem. Soc. 2010, 132, 13132-13134), very systematically analyzing the conformational landscape, puckering etc. Generally speaking, the conclusion of these studies were that very long production runs (micro seconds) were needed in order to find equilibrium populations, and they would even to as far as discarding the first 250 ns of a production run as this was viewed as unrepresentative of the entire run. With this in mind, the uniform sampling becomes problematic. I am convinced that a 500 ns production MD run is more than sufficient to sample the conformational surface of the compounds studied in this manuscript, provided a population analysis identifies the main conformers, after which a weighted sampling of each conformer group could be done. I am concerned that the uniform sampling over the entire production run will include a considerable portion of conformers that would not be representative in equilibrium over the experimental measurement, and hence the comparison with experiment is coincidental.

We share the concerns of the referee regarding the importance of the proper molecular sampling for the calculation of high quality Ramman/ROA spectra. In fact, the sampling problem is central in our developed method based on long simulations, large selection of uncorrelated frames, and partial optimizations. All those methodological constrains imposed in our approach ensure that the obtained spectra correspond to the equilibrium ensemble provided by the MD simulations. This alone directly allows the assessment of the quality of the used force field.

First, the reviewer is concerned about our uniform sampling over our 500 ns MD simulation (every 5 ns, taking 100 snapshots). We would like to point out that this method of gathering frames is the one used by all MD software because it results in perfect equilibrium distributions when the sampling is sufficient. Also it is easy to prove that randomly picking structures from well equilibrated simulation results in an identical ensemble. To avoid any misunderstanding, we have removed the word 'randomly' from the scheme. This change as explained above has no physical implications in the validity of the populations analyzed.

Second, using glucose as representative example, in Figure 2 we show that during our simulations we indeed sample efficiently the system to be representative of the underlying free energy profile (Black crosses). We see that the only mainly contributing structures are found in $^4C_1$ puckering conformation, which is consitent with the over ∼15 kJ/mol penalty one has to pay for populating other sugar conformations. This is also true for all the others sugars presented in this work.

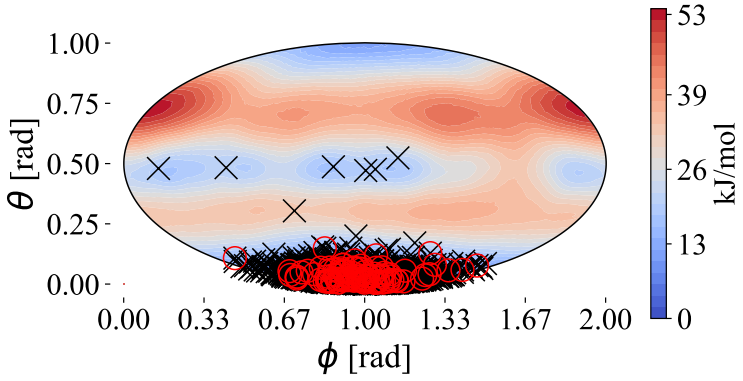

Figure 2: β-glucose: Free energy profile in ring puckering coordinates at 310 K. Black crosses indicate sampled conformers during 500 ns unbiased MD run (25000 snapshots). Red circles indicate 100 snapshots uniformly sampled every 5ns, which spectra are further shown in Figure 4 and Figure 5. Free energy profile was obtained from metadynamics MD simulation.

Third, using the same example, we see that our selection method results in frames with equal ensemble properties for the puckering coordinates which map the overall sugar spatial conformation, red circles in Figure 2. Moreover, considering finer structural details in Figure 3 we study the behavior of the exocyclic CH_2OH rotation that has 2 major minima

which are roughly equally populated. Notably, this is also true for the frame selection used for calculating the spectra. Note that this results are for β-glucose molecule, but all other studied molecules behave in a similar way. The similarities between the ensembles provided by the simulation and selection procedure are clear, meaning that we are indeed sampling the equilibrium phase space and we do not have any weighting problem.

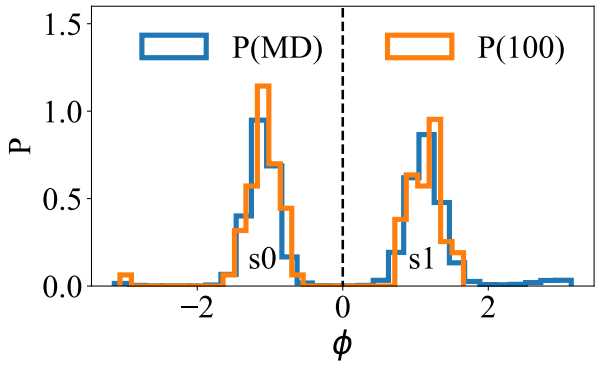

Figure 3: β-glucose: Histograms of rotation of exocyclic CH_2OH dihedral angle obtained from whole 500 ns MD simulation(blue) and from 100 uniformly sampled snapshots(orange). The dashed line separates the populations into groups s0 and s1.

Fourth, we uniformly sampled the 500 ns MD simulation obtaining 100/250/500 structures (5ns/2ns/1ns sampling rate) and calculated the average spectra. In the limit of large time scales and sufficient number of structures we should obtain a correctly sampled phase space and that is indeed what we get. Figure 4 shows that 100 structures is already enough to obtain a converged data, as adding more structures does not qualitatively change the resulting spectra. More information about the convergence of the spectra with the number of frames can be found in Figure 18 in ESI.

Moreover, our proposed method has a limited resolution. Therefore, small variations on the populations due to our selection of frames method has a negligible impact on the outcome spectra. As an example, in Figure R5, we mix the spectra of two real β-glucose conformers (s1 and s0 ensembles) in different proportions. The conformers are divided based on the exocyclic CH_2OH dihedral angle, as shown in Figure R3. Adding 10-20% of s1 conformers to the pure s0 spectra has a minor effect on the overall s0 spectra. This is true even though

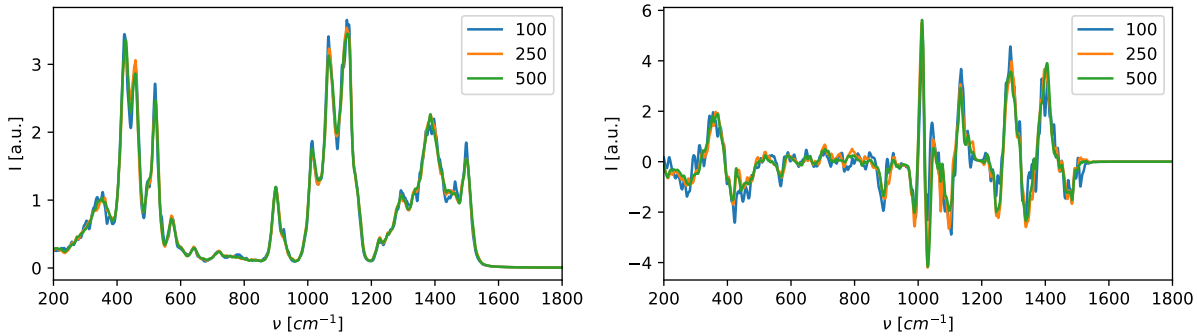

Figure 4: β-glucose: Raman(left) and ROA(right) average spectra calculated for 100/250/500 structures from 500 ns MD simulation when uniformly sampled (5 ns, 2 ns, and 1ns sampling rate).

s1 spectra possess clear distinguishable features.

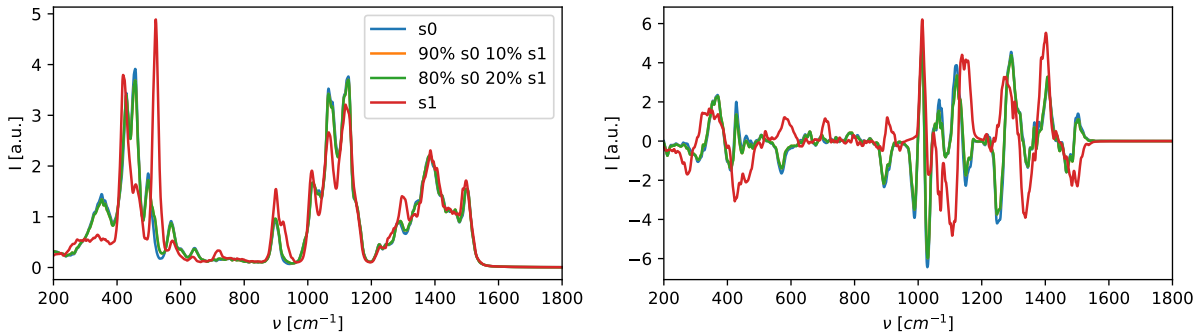

Figure 5: β-glucose: Raman(left) and ROA(right) average spectra calculated for s0 and s1 ensembles and their mix. s1 and s0 ensembles refer to clusters based on the exocyclic CH_2OH dihedral angle as defined in Figure 3

All these results make us confident that this work does no suffer from any sort of sampling issue. All these results also support the notion that the improvements of the spectra in this manuscript with respect to the ones published in the literature are the result of the well designed approach and not some sort of causality.

I think this is a very important finding and one of the big arguments for this approach. The Almond group (Satelle et al) most likely ran their MD runs too long, i.e. equilibrium was reached before the microsecond time frame, at least for some starting structures... They also sampled some structures that were very far from equilibrium. In this manuscript, all

starting structures were in the 4C1 conformation (a very sensible choice), so equilibrium should be reached before 500 ns. I think that it is an important finding that the conformer distribution appears stable quite early in the run, and that taking the ROA spectra over different time frames show little change to the spectra. This, to me, in fact validates the sampling procedure. Thus, I think at least some of the data presented in the rebuttal letter should be included in the ESI.

This rebuttal is published together with the paper, therefore its content will not be added to the ESI.

2.2.6 R2Q6

This issue is further confounded when you consider the QMMM level geometry optimization. In the hybrid protocol, the authors employ a restricted geometry optimization, in order to retain the conformational input from the MD run. This would make sense if the MD snapshots were population weighted (even though the protocol developed by the Manchester group would indicate it is not that important), but since they are not, I question the validity of the approach.

Our method of selecting snapshots, as explained carefully in R2Q5 provides the correct equilibrium ensemble within the MD quality limit. Therefore, there is no need to weighting anything. The weights are merely the relative abundance of each of the selected structures. So yes, we want to keep at all cost the phase space obtained by the selection method. However, bonds, angles, and some interatomic distances are not accurate enough in the MD data. Therefore a limited small QM optimization is required to relax such degrees of freedom before the spectra calculation. Still, motivated by the comment of the reviewer, we realized that we were overdoing it. In this revised manuscript, we have eliminated the need for harmonic restrains during the optimization process when we use explicit water molecules such as in the 'hybrid' method. The partial optimization we propose now takes care of avoiding a significant contraction of the phase space.

Good point on the weighting issue. I will comment more on restricting the number of optimization steps below, but removing the harmonic restrains is an improvement.

2.2.7 R2Q7

Combine this with the seemingly arbitrary cut off of '10 optimization steps with harmonic restrains' and the protocol is now physically unsound. Stopping the geometry optimization at specific number of steps more or less guarantees that the system is not in a steady state, hence making property tensor calculations meaningless. I do not believe that convergence towards the error function is a suitable measure for the quality of the calculations, nor is bench-marking to QGRAD or 'unrestrained optimization', when both of these also are cut off at a certain number of steps (50 and 200, respectively). Why not just let these run to convergence of the gradient? The argument of efficiency is completely moot if non-physical data is produced.

We are sure that the reviewer is aware that within the harmonic approximation the frequencies are the same for any geometry, and the same argument holds for the property tensor calculations. There are certainly many interesting theoretical aspects to this, but most were already solved. For example we find papers discussing: the normal mode optimization (J. Chem. Phys. 2002, 117, 4126-4132.; Collect. Czech. Chem. Commun. 2005, 70, 1315-1340; J. Phys. Chem. B 2012, 116, 336-342.), Raman/ROA intensity carrying modes (Chem. Phys. Chem. 2009, 10, 2049-2057), mode tracking optimization (J. Phys. Chem. A 2004, 108, 2053-2061), or the instantaneous normal mode approach wildly used in simulations of liquids (J. Phys. Chem. A 1997, 101, 16, 2921-2930). Even in the 'Manchester' approach, that the reviewer presents as reference in the field, water positions are not optimized. Therefore, we do not understand why the reviewer thinks that the approach proposed here is not physically sound. Certainly, we cannot just 'run to convergence of the gradient', without destroying the information from non-zero temperature MD simulation.

We have carefully tested all our procedures, and added even more tests during this

revision as explained below to ensure proper physically sound behavior of our approach. As we obtained a good agreement with experiment not in one case, but in all 6 cases (also others not included in this publication), we feel that our approach is sound and produce accurate data.

Nevertheless, as the reviewer suggested, we decided to make additional calculations which address the raised concerns. We use two new approaches, `hybrid_w` and `hybrid_fo` which are variants of `hybrid` approach. In `hybrid_w`, all water molecules are frozen and the central sugar is optimized until gradient criteria are reached while the rest of the procedure stays the same. In `hybrid_fo`, none of the 3 Å water molecules layer are frozen and the optimization is run until gradient criteria are reached while again the rest stays unchanged. We have also added so called OPTSOLUTE approach developed by the Manchester group - `MM_w`, which is similar to `MM` protocol with the difference that all water molecules are frozen and the sugar is fully optimized. Description of all the mentioned approaches has been added to the main text and the details can be found there.

Figure R6 shows the overall performance of all the used approaches in terms of the overlap integral. All new protocols (`hybrid_w`, `hybrid_fo`, and `MM_w`) are run until the convergence criteria are reached. All produce worse results than the 10 steps optimization `hybrid`. In particular, the full optimization `hybrid_fo` significantly worsens the results. The `hybrid_w` and `MM_w` provide better results than `hybrid_fo`, but still worse than our selected `hybrid` and at higher computational cost. Therefore, `hybrid` provides the best quality spectra at the lowest cost.

Lastly, since the reviewer does not feel comfortable about the selection of 10 steps optimization, we tested the performance of `hybrid` protocol using different number of optimization steps prior to spectra calculation. Figure R7 shows that the best results are obtained when using 5-10 optimization steps and that the spectra deteriorate upon further optimization.

I am not sure I understand the point about the harmonic approximation: The Hessian

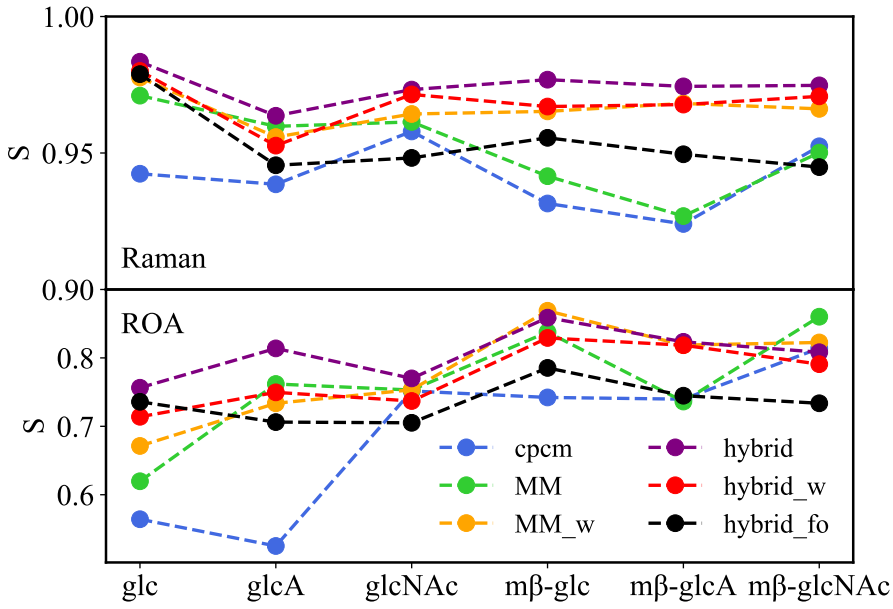

Figure 6: Comparison of all used approaches (`cpcm`, `MM`,`MM_w`, `hybrid`, `hybrid_w`, and `hybrid_fo`) using overlap integral S.

is dependent on the normal coordinates and thus geometry dependent. And so are the invariants calculated for the property tensor calculations. That said, I think the need for a limited number of steps in the optimization procedure now, with the additional testing the authors have done, has gone from being a critique point of the manuscript to one of the most important findings! First of all, good that several other combinations of hybrid schemes have been tested, including full optimization, just to show that the primary difference merely is additional time costs. But systematically showing that up to 10 optimization steps is sufficient and that the Sfg value actually drops with further "optimization" could be a potential game changer for the entire field! This is sufficiently highlighted in the new version of the manuscript, which is a massive improvement.

Our point about the harmonic approximation really was that in the strict harmonic limit the Hessian does not depend on the geometry. Such a dependence would include its first derivatives = third energy derivatives, but these are neglected in the harmonic limit. In any case, we agree with the reviewer that in reality the Hessian does depend on the

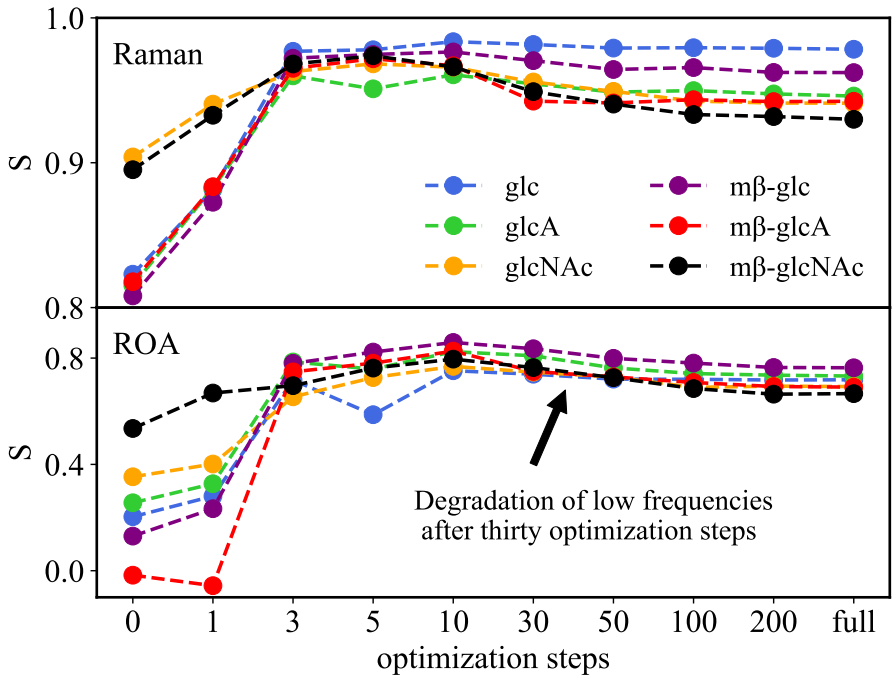

Figure 7: Performance of the `hybrid` protocol when using different number of optimization steps prior to spectra calculations.

coordinates, and that only the testing of the optimization approximations provides a proof of their usability.

Finally, we agree with the reviewer that requiring only ∼10 optimization steps can be a game changer for the computational requirements of the ROA/Raman spectra calculations.

2.2.8 R2Q8

Dynamic scaling: While this is an interesting concept, I was initially quite skeptical of using variable scaling. Furthermore, it took me about four reads to grasp what was going on mathematically, so this part of the methodology should at least be clearer. The real problem here lies in the fact that the raw simulation Raman and ROA data is, as described above, somewhat arbitrarily produced and combining this with a more complicated fitting procedure that has the purpose of minimizing an error function, is problematic. In a sense, the predicted spectra are made to fit as best as, regardless of the initial quality.

The use of scaling factors when calculating vibrational frequencies has been proposed long ago [DOI: 10.1002/cphc.201801172] and is widely used. Moreover, the use of different scaling factors for different frequency regions is not new either. For example, in ref [DOI: 10.1002/jcc.23073], the authors present different scaling factors for the low frequency ($<$1000 cm^{-1}) region and high frequency region ($>$1000 cm^{-1}). For details on this topic see ESI Section 5. In our view, the need of a more complex scaling function arises for the arbitrarity of the selection of the threshold (1000 cm^{-1}). Because of that, we propose our scaling function which is designed to find optimal position of the threshold and ensure a soft transition between the two regions. Why do we use 4 parameters in the scaling function? Because 2 parameters correspond to the high/low frequency scaling factors (like in DOI: 10.1002/jcc.23073), 1 parameter correspond to the threshold between these scaling factors, and the last parameter ensures the smooth transition between the low/high frequency regions.

To demonstrate the advantage of our scaling function, for example, Figure R8 shows the calculated spectra of methyl β-glucuronide (top) without any scaling, (middle) when using a single scaling factor [0.98 in this case], and (bottom) when using our proposed scaling function. When we do not scale, the spectra are shifted above 1200 cm^{-1} and they do not reproduce the experimental data. When we use a single scaling factor(0.98), which is the most widely used method in the literature, we can see that the region above 1200 cm^{-1} region now corresponds to the experimental data, however, the low frequency region is badly adjusted. Only when using different scaling factors for low/high frequency regions, as with our scaling function, we obtain good agreement with experimental spectra. This can be monitored using the overlap integral scoring function (S). Calculated values for unscaled/scaled(0.98)/scaling function are for Raman: 0.951, 0.933, and 0.974 and for ROA: 0.579, 0.618, and 0.823 respectively. The improvement of the obtained values points to the strength of using the proposed scaling function.

Moreover, notice that the scaling function is the same for Raman and ROA, meaning

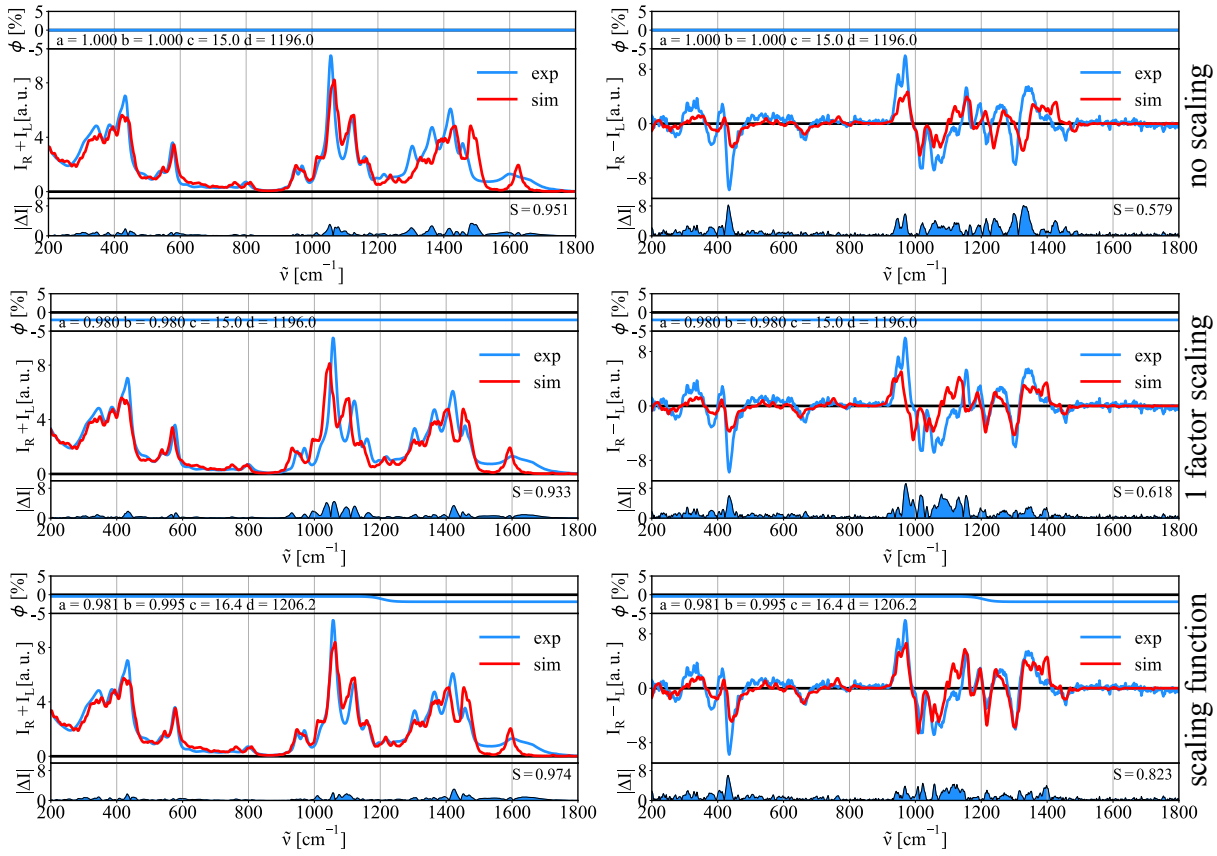

Figure 8: Raman and ROA spectra of methyl β-glucuronide calculated by hybrid approach, where calculated frequencies were not scaled (top), were scaled as found in the literature by single scaling factor 0.98(middle), and scaled by our adaptive scaling function(bottom). In each graph the scaling function is found at the top, simulation and experimental data in the middle, and differences in experiment and simulation intensities, together with error function and overlap integral at the bottom.

that we do not adjust every spectrum separately. Finally, we would like to point out that all our approaches converge to similar shaped scaling functions, regardless of the particular molecule or optimization procedure (Figure R9).

Now that I am convinced about the sampling and optimization procedure, I have less of a problem with the variable scaling. My biggest concern about this idea, in general (as the authors point out, they are not the first to suggest a more involved scaling than one linear scale factor), is that the more complicated the scaling, the easier it is for none-experts to overdo it. Considering the authors of the present manuscript, I am convinced that in this

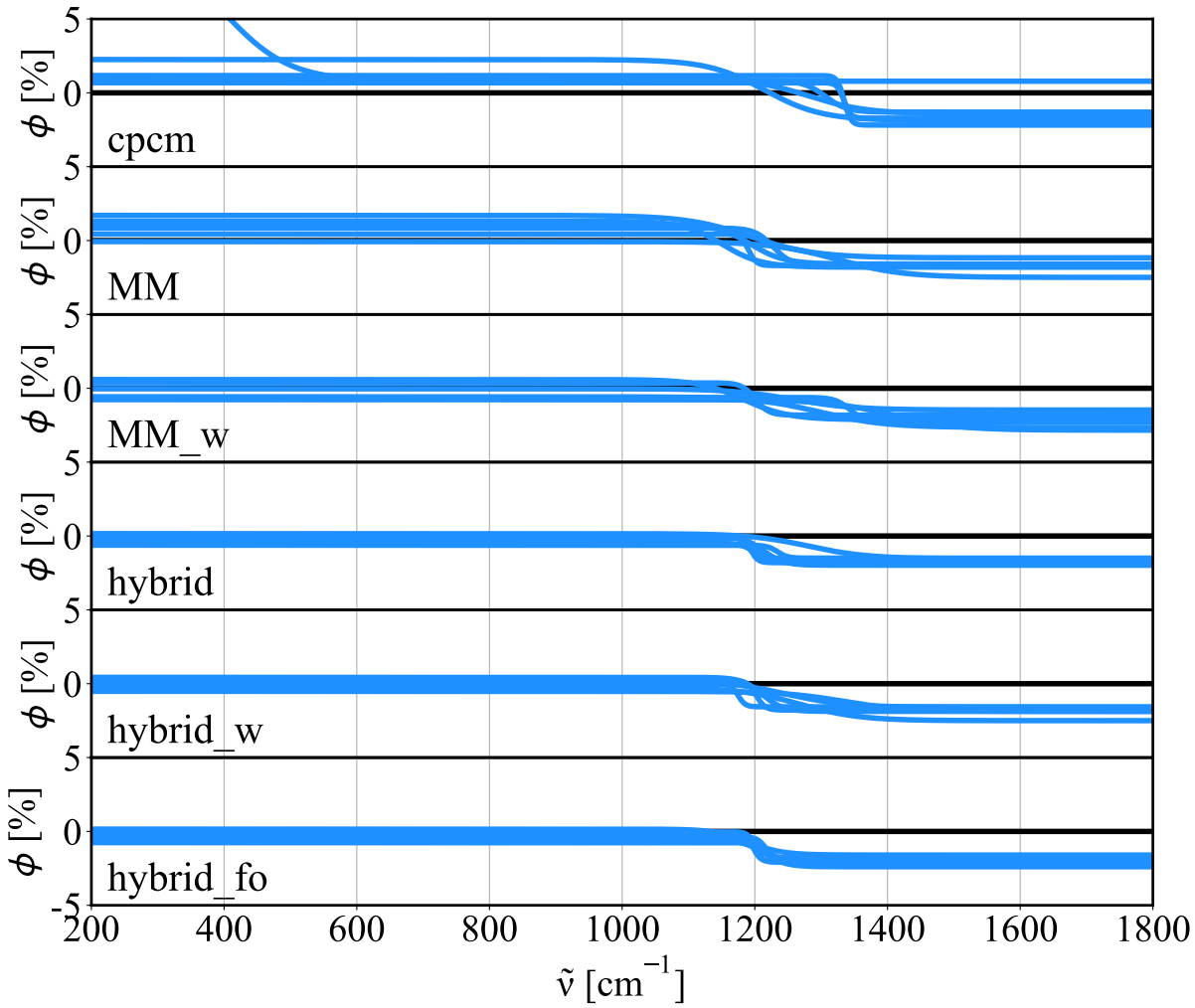

Figure 9: Scaling functions produced by different optimization approaches (1 scaling function per molecule, i.e. 6 per approach)

case, the scaling is done properly and consistently, but this may not be the case for other researchers. The question is: Is there some way to safeguard against random scaling done to fit apples and oranges? Probably not. But, as I also agree with the authors that simple empirical scaling is not always optimal, at least if they set the new standard for variable scaling, at least I know that it is done right in their papers. So, I guess I am OK with this.

One of the results of our paper is that we propose an average scaling function for a given method which is carefully tested. Moreover, in the answers to questions R2Q8 and R3Q2 we tried to show that using our proposed form of scaling cannot be used to fit apples with

oranges. Therefore, we do not worry about the misuse of the scaling function in its current form.

2.2.9 R2Q9

Some of the ideas presented here (as this is a novel concept, a bit of background to how the equations were developed would be good) are sound and straight forward, but the frequency scaling abruptly 'jumping' at the intersect (as seen in some of the figures in the SI) and the intensity fitting based on the rho factor are particularly concerning.

The scaling function is needed to partially correct known problems when using DFT for spectral calculations, as described in Section 5 in ESI. We explain in detail the reason behind the chosen functional form in the previous question (R2Q8). We agree that the 'jump' around 1200 cm^{-1} is interesting and deserves full understanding. However, currently, we only partially understand the underlying problem. Finally, we have improved the functional form description in the current manuscript.

The intensity scaling is necessary for Raman and ROA, where absolute intensities cannot be measured, as discussed, for example, in L. D. Barron, 'Molecular light scattering and optical activity' (2004). The ρ factor is only used to automatically scale calculated intensities to experimental intensities, which limits 'human errors'.

Agreed

2.2.10 R2Q10

Error function: This is a smaller issue than the ones described above, but I have a problem with using a quantitative comparison indicator when the calculated data already have been fitted to have maximum overlap with the experimental data. In this case, the error function should naturally become small.

First, we agree with the reviewer that the overlap integral is much more intuitive and easier to interpret as compared to formerly proposed error function. Therefore, we abandoned

our error function and we use overlap integral instead for determining the quality of the spectra. Second, our optimization method still uses the old error function. Therefore, now the quantification of the quality of the spectra uses a different metric than the optimization algorithm making our result more robust as it is governed by to independent measurements of quality. Finally, note that optimization methods (to obtain best overlap) for Raman/ROA spectra calculation are common in the literature (Pendrill et al., ChemPhysChem 2019, 20, 695-705).

Agreed

2.2.11 R2Q11

Also, it is unclear from the manuscript what constitutes a 'good' number. Obviously, zero would be optimal, but that is never going to happen without full CI calculations. Studying Figure 8 and 9 in the SI, 1.9 appears to be a 'large' error function and 0.6 a 'small' one for ROA. This analysis could be compared to overlap integral analysis, which is more intuitive (e.g. percent overlap) or benchmarked on something simple, such as alpha-pinene, which experimental and predicted ROA spectra are very similar.

As is stated in the answer above, we are abandoned the error function and we are using the overlap integral instead.

Agreed

2.2.12 R2Q12

Finally, I would like to comment on some more minor issues in the manuscript. Alpha-beta anomer contributions: The authors claim that the alpha-beta anomeric distribution has been neglected in previous studies predicting ROA of monosaccharides. This is not entirely true. In some studies, Such as Cheeseman et al, anomeric dependence was circumvented by studying derivatized samples, locking the anomeric carbon in one conformation.

We agree that the sugar molecules can be studied by using derivatives that lock the

anomeric carbon in one conformation, such as the methyl glycoside. However, locking chemically a sugar molecule in a configuration as Cheesman et al. did, does not allow for studying α/β anomeric equilibria and how it affects Raman/ROA spectra. Moreover, we find unfortunate that this phenomenon was in many studies simply neglected or at occasions deemed as irrelevant contrary to our results, e.g. DOI: 10.1039/c4cp05517a, DOI: 10.1021/acs.jpca.6b00358, DOI: 10.1039/c5cp02969d.

Having clarified this point, we are the first to show that without independent consideration of the α/β anomeric isomers for a reducing monosaccharide, one cannot hope to achieve spectral agreement with experiments. To make it clearer, we added each of the anomer spectra to ESI (Figure 7) and here in Figure R10. This figure shows the spectra of glucose (top)/ glucuronic acid (middle)/ N-acetyl glucosamine (bottom). For each compund, on the left we see calculated spectra of α/β anomer and on the right comparison of α/β weighted average to experimental data. This figure shows clearly without doubt the importance of this phenomena in the Raman/ROA spectra calculations of non-reduced sugars.

Lastly, we must point out that a common erroneous assumption is that the same β sugar is bought and measured. The same anomer is also calculated for comparison. Unfortunately, what happens is that after solvation the sugar quickly epimerizes until the α/β anomeric equlibrium is reached. This process is very quick taking only a few hours (DOI: j.carbpol.2011.06.056). Since the Raman/ROA spectra measurement takes around ∼10-20 hours, what is measured in fact is not a single anomer, but again the equilibrium mixture of both.

Here, I very much agree with the authors. I even think that it could be stressed more in the manuscript how important it is to include both anomers instead of just ignoring one! This is very obvious from the spectra added to the ESI.

We agree and therefore we added a following text to the main text:

"Moreover, we show that both anomers have to be considered while aiming at simulation of reducing sugars. Failure to do so means essentially to partially simulate different system

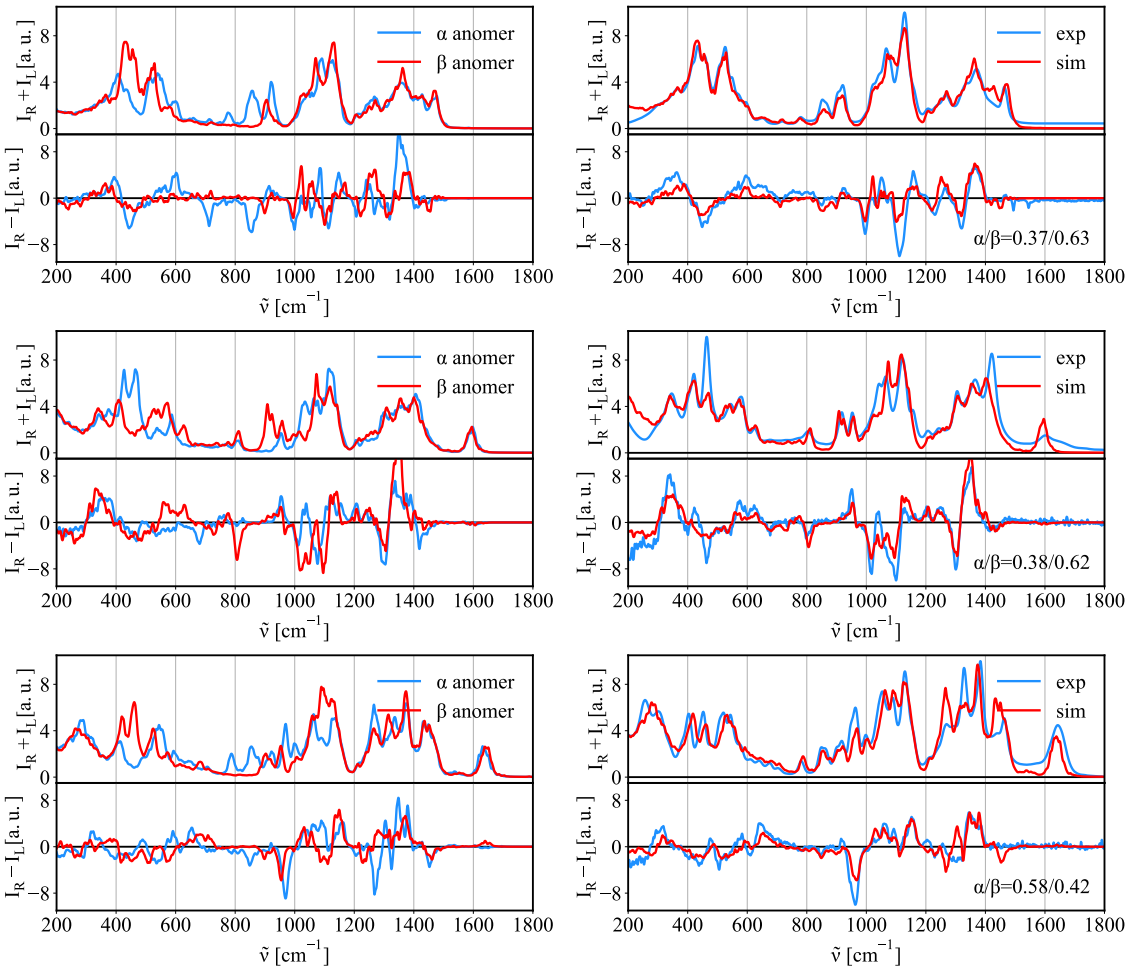

Figure 10: Calculated Raman/ROA spectra of glucose(top)/glucuronic acid(middle)/N-acetyl glucosamine(bottom) alpha/beta anomers, their weighted average(sim, glucose - 0.63, glucuronic acid - 0.42, N-acetyl glucosamine - 0.62) and comparison to exp data(exp).

than the one that is studied experimentally."

2.2.13 R2Q13

But, in the study of epimers (Mutter et al 2016, reference 16 in the manuscript), the anomeric distribution is mentioned and then discarded as 'not important' for the study. I would have liked to have seen, in this manuscript, spectra averaged over the two anomers and how this impacts the resulting ROA spectra, in order to see how big this effect really is.

We are aware of this work and we think that it is wrong as explained in the previous

question (R2Q12). In fact, this was actually one of the reasons why we started this work. We were puzzled about how such significant changes in the molecular structure could result in negligible changes in the Raman/ROA spectra while minor structural changes usually result in substantially different spectra. Taking as example Figure R10, we show that one cannot neglect the presence of different anomers in solution.

See above.

2.2.14 R2Q14

The quality of the D-glucose predicted spectra: The authors do mention that D-glucose is not as good a fit to experiment as the other samples, but the poor comparison is striking, and not discussed in nearly enough detail. D-glucose has been studied extremely extensively with a myriad of techniques, experimentally and theoretically. The anomeric distributions, conformer landscape etc. is well known and predicted ROA spectra that are in much better accordance with experiment have been published (Mutter et al 2016, reference 16 in the manuscript). Yet, the surprising discrepancies between experiment and calculations are very much brushed off in the manuscript, which is very unfortunate, as it make it appear as the authors are avoiding discussing this issue.

In the revision we have abandoned the use of the harmonic constrains during the optimization. In effect, this made the `hybrid` protocol perform significantly better for the glucose molecule (the value of overlap integral reaches similar values as in case of other molecules, see Figure 1 and Figure 5 in the main text).

Agreed. Nice.

2.2.15 R2Q15

Instrument comparison (figure 4): It is hardly surprising that ROA spectra from different instruments (and different sample providers/handlers) are different to some extent. This can be due to the alignment of the spectrometer, wear of the optical elements, dust on these,

instrument make, sample quality etc. Using experimental variability to argue for a flawed computational approach is not a valid argument for the latter.

We are afraid that the reviewer did not understand well the argument. By pointing out the experimental inaccuracy, we do not try to excuse any computational flaw. On the contrary, we use it as a warning to avoid overinterpreting the similarities between the computed and experimental shapes. Thus, we are just stressing that the accuracy of the presented method will be always limited by the accuracy of the experiments that we compare to.

It is now clear that the computational approach is sound, so ignore the "flawed" comment. That said, I still think it is dangerous to use experimental variability in this argument. Well aligned/well-kept ROA instruments should produce very similar spectra. When this is not the case, this is mainly due to alignment issues. I have seen experimental ROA data published that bore no resemblance with previously published spectra, meaning one spectrum is real and the other full of artefacts. One would not like to fit correctly samples computational data to the latter, as that should result in inferior comparison? The issue is that if the sampling is not correct, a reasonable comparison with an artefact prone experimental spectrum may lead to incorrect conclusions. I guess what I am saying is that the computational chemist have to be careful when choosing experimental data to compare with. But I guess this section can be kept in.

We agree with the reviewer.

2.2.16 R2Q16

MN0L functional: I have trawled Gaussian's home page, done extensive searches on the internet and in literature, but I cannot find a single reference to the 'MN0L' functional anywhere.

This is a mistake done on our side, we correct it. (the correct name M06L)

Agreed

2.2.17 R2Q17

Full QM protocol: I assume this is one further 'benchmark' in support of the 'hybrid' protocol, but it is hardly surprising that not much happens in the 'full QM' calculations when everything is restricted and only ten optimization steps are run... That is so far from convergence that no gain in quality is made from treating the waters as QM objects. Also, how are the distances from the solute determined in this case? Waters out to 2.5 Å is treated at QM level, while waters from 2.5 to 3 Å are MM. How does that even work when a O-H bond in water is 0.9 Å? Half an MM water?

After reading the reviewers' comments, we have decided to abandon the QM-MM water molecules approach. Despite being interesting, this approach has a certain level of selection arbitrarity, as pointed out by the reviewer. Therefore it is prone to questions and opinions. In the current manuscript, the water molecules are either treated as MM or QM waters.

Also, we agree that it should be checked whether 10 steps optimization is enough for water molecules to properly relax to see any effect. Therefore, we decided to carry out the full optimization calculations (using both QM or MM water molecules) for glucose molecule to see this effect. Figure 11 shows calculated spectra of glucose using `hybrid`, `hybrid_fo`, `hybrid_QM`, and `hybrid_QM_fo` protocols. The first two (`hybrid` and `hybrid_fo`) use MM water molecules, while the latter (`hybrid_QM` and `hybrid_QM_fo`) use QM water molecules. In case of `hybrid` and `hybrid_QM`, the optimization prior to the spectral calculations is run only for 10 steps, while the other two (`hybrid_fo` and `hybrid_QM_fo`) optimize the system until gradient criteria are reached.

Using `hybrid_fo` approach with MM water molecules gives slightly worse results as compared to 10 steps optimization using `hybrid` approach. Switching to QM described water molecules when performing 10 optimization steps (`hybrid_QM`) leads to even worse description of the high frequency region (>1000 cm^{-1}), however, the low frequency region of ROA spectrum was better described. Using full optimization with QM water molecules (`hybrid_QM_fo`) results into the improvement of the high frequency region as compared

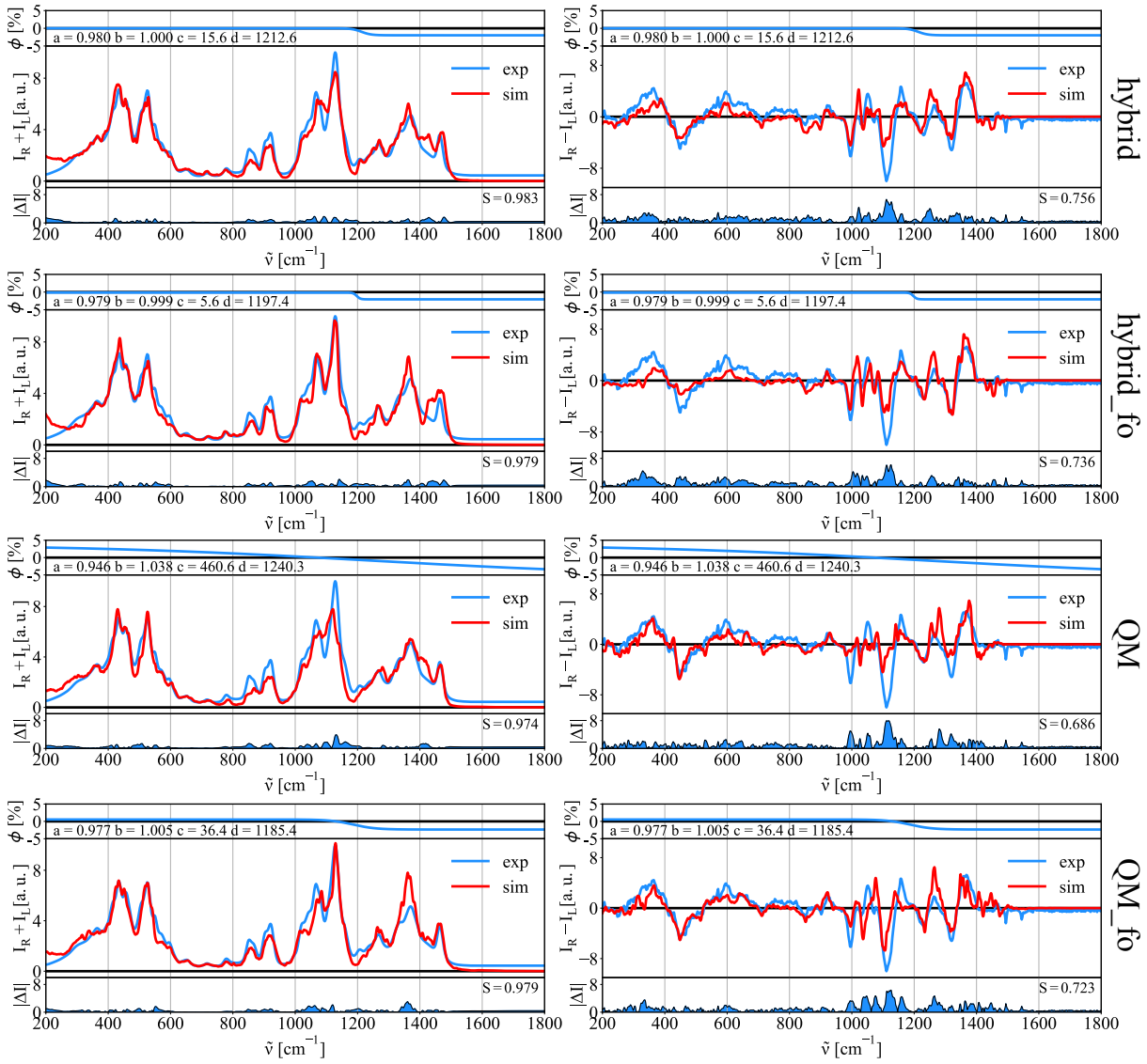

Figure 11: Simulated and experimental Raman (left) and ROA (right) spectra of glucose obtained by `hybrid`, `hybrid_fo`, `hybrid_QM`, and `hybrid_QM_fo` simulation approaches. Vibrational frequency scaling function ϕ with optimized parameters is provided with each of the simulation (above spectrum). Absolute differences between the experimental and the model spectrum are shown below the spectrum, together with the overlap integral S.

to `hybrid_QM` approach, while the low frequencies are still well reproduced. Still, the `hybrid_QM_fo` evaluated by the overlap integral performs worse that `hybrid`.

The data suggests that higher the frequency region (>1000 cm^{-1}) is slightly better reproduced when using full optimization, regardless of the theory used for the description of water molecules. In the lower frequency region the description of water molecules mattered

and QM description is superior to MM.

Overall, the `hybrid` approach seems to be better balanced. Moreover, using QM water molecules is expensive. In the relative cpu times, the time needed for each approach was: `hybrid:hybrid_fo:hybrid_QM:hybrid_QM_fo`=1:8:25:70.

Again, good work removing any doubt about the method by further testing.

2.2.18 Conclusion remarks of the first draft

In conclusion, I trust I have covered, in sufficiently detail, the major deficiencies in this manuscript, illustrating why the manuscript is unsuitable for publication in Physical Chemistry Chemical Physics or in any other journal.

We hope that with the provided explanations, additional calculations and modifications of the manuscript all the concerns of the reviewer were addressed.

2.2.19 Small issues after 2nd revision

Small things noticed in the main text: p. 1. I just noticed that the paragraph introducing ROA (second column, lower part) does not contain a single reference. Maybe refer to a few of the classical Barron papers?

A reference was added.

p. 3., col. 2 line 29: "The" before "CPCM"

p.3., col. 2 line 31: "the" before "MM"

p.3., col. 2 line 33 "approach" should be "approaches"

The mistakes were corrected.

p.4. Figure 2: It is unclear what "mutation and recombination" means.

"Mutation and recombination" are standard words for describing two critical steps of a differential evolution algorithm. Still, we understand now that it can be confusing for people not familiar with this technique. We changed a bit the figure to be more self-explanatory. Now we just talk about "Parameters optimization" inside the box, demoting the "mutation

and recombination" sentence to a secondary position (below the arrow).

p.5. col. 1 line 8 (I think. Generating the .pdf has mashed the text around a bit right here): "Section 8" should be "Section 5"

We agree. The main text looks fine.

p.7. col. 1 "the conclusion". I agree, but maybe write a sentence hypothesizing why? I am especially intrigued about the "best number of steps", e.g. more optimization steps above the 10 decreases the quality of the spectra: Does this mean that DFT level optimization leads to a poorer conformational sampling by diverting the MD geometries away from the correct distribution? If so, we should all stop doing full optimizations and only use the present approach. Obviously, I am not suggesting writing conjecture, but maybe include a sentence or two about this?

This means that for a given setup - 1 solvation layer and optimization without any restrains - an overdone optimization only worsen the results. This is truly because of diverting too much from the MD distribution, which reflect non-zero temperature. Optimizing too much in this case destroys this temperature effect.

We have to stress that for other setups the results might differ. Therefore, stopping the optimization early while using frozen water molecules might not be optimal. Using optimization in normal modes may also benefit from optimizing the system fully.

Supplementary information: Table 1 vs. table 3: Is "CPCM" and "cpcm_HR" not the same?

No, they are not. `CPCM` does not use harmonic restrains during the optimization step as compared to `CPCM_HR`.

2.3 Comments to reviewer-3 (R3)

2.3.1 R3Q1

The authors tackle a very challenging problem, namely the theoretical modelling of saccharides in solution. This is in principle an important study as this class of molecules is of high importance and relevance, and the experimental methods available for obtaining structural information about this class of molecules are limited. Furthermore, analysing the experimental ROA and Raman spectra to deduce structural information will require concomitant computational simulations.

We thank the reviewer for the comments.

The paper has been reviewed by to independent reviewers, which give different conclusions on the suitability of the paper for publication in PCCP. I unfortunately largely agrees with the most critical of these two reviewers. The goal of the project is laudable, but I feel that the authors on their way to this goal, a reliable computational protocol for simulating Raman and ROA spectra of saccharides, sacrifice a bit too much in order to get good agreement with experiment at a reasonable computational cost.

As already states above, our approach not only leads to significant speedups but it does it in a physically sound way without nearly compromising the obtained spectra.

2.3.2 R3Q2

The second reviewer has elaborated in detail on these aspects. From my point of view, there are two main aspects that I struggle with in particular: - The adjustments of the calculated spectra against the experimental ones. Although there are constraints imposed on the fit, I am concerned that the hidden errors and limitations in the computational protocols/results are hidden because of this adjustment process. For sure it would be good to see how well (or not) the computational results are without this step.

We agree that the devil may be 'hidden in details', but in this particular case we really

do not see any 'hidden' obstacles as elaborated in the answer to R2Q8. There we show how the spectra would look like without the use of scaling function and with the use of a single number scaling, e.g. 0.98 for all frequencies.

Figure 12 further shows that the scaling function cannot completely transform the simulation data to fit any experimental data.

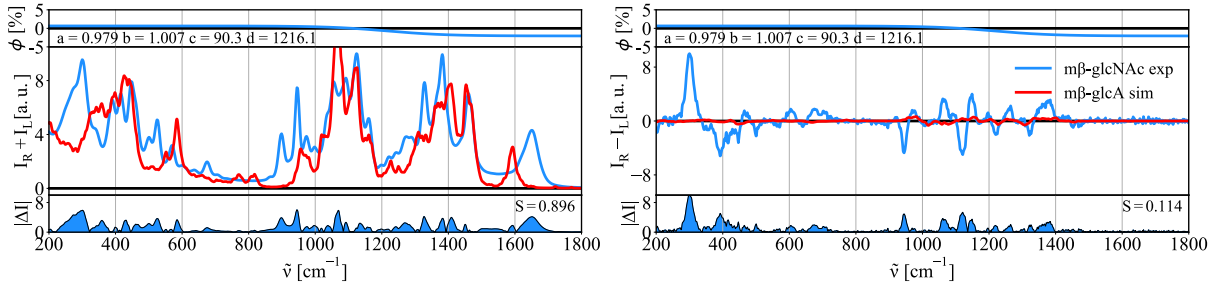

Figure 12: Example of what happens when we use scaling function when fitting simulation data of mβ-glcA to experimental data of mβ-glcNAc.

2.3.3 R3Q3

It is otherwise nice to see that the authors notes and acknowledges the need for different scaling factors in different frequency regions, as this is something we have also observed in my group yet which I believe is not well recognised in the literature.

We were also were puzzled by the need of different scaling factors for different frequency regions so we thank for the encouragement. As elaborated above (R2Q8), also some other people noticed this, but certainly we do agree that this simple fact is not generally recognised.

2.3.4 R3Q4

The use of non-optimized structures. The authors have great expertise on finding efficient means of finding optimised structures of complex intermolecular systems, but at a non-optimised structure, you will have residual forces acting on the atoms, and the quality of the normal modes will in this case not be clear. This then also links to the previous concern,

namely whether these limitations are "hidden" due to the subsequent adjustment protocol against the experimental spectra.

We understand this objection, but there is no rule forbidding other approaches. As elaborated in R2Q7 many researchers including us acquired significant experience with them. Regarding the present paper, this issue is perhaps tested most of all. There are additional tests in the revision and they all show that the quality of the description of modes of interest (>200 cm^{-1}) is not compromised by using nonoptimized structures as they already produces high quality Raman/ROA spectra. Stopping the optimization after 10 steps both save time a produces accurate data (compared to experimental data), see new Figure 6.

2.3.5 R3Q5

The second reviewer also raises a number of additional concerns that I do repeat here. Overall I find that there is a lot of potential in this paper, but that it would benefit from keeping a somewhat clearer focus on the goal of a pure computational protocol and not too heavily rely on approximations and fitting to the experimental spectra.

What is a problem here is not the ensemble obtained during the optimization but whether the tensor calculations and the second derivatives using the harmonic approximation are reliable. Considering our systematic agreement with experimental data it is unlikely that our good agreement is coincidental. However, we acknowledge that this does not have to be true always. But there are currently no accessible protocols without any flaws and the optimal performance/cost ratio of our method likely justifies this small uncertainty. We are aware of used approximations and in case of doubt the user can always go to higher level methods, e.g., anharmonic calculations.