Ion spectroscopy of heterogeneous mixtures: IRMPD and DFT analysis of anomers and conformers of monosaccharides

Loïc Barnes, Abdul-Rahman Allouche, Stéphane Chambert, Baptiste Schindler, Isabelle Compagnon

Ion spectroscopy of heterogeneous mixtures: IRMPD and DFT analysis of anomers and conformers of monosaccharides.

Loic Barnes[1], Abdul-Rahman Allouche[1], Stephane Chambert[2], Baptiste Schindler[1] and Isabelle Compagnon[1*]

(*) Corresponding author: isabelle.compagnon@univ-lyon1.fr

1. UnivLyon, Université Claude Bernard Lyon1, CNRS, Institut Lumière Matière, F-69622Villeurbanne, France
2. UnivLyon, Université Claude Bernard Lyon1, CNRS, Laboratoire de Chimie Organique et Bioorganique, INSA Lyon, F-69622Villeurbanne, France

keywords: Protonated hexosamines; IRMPD spectroscopy; glycomics

Abstract

The gas phase IR spectra of three protonated hexosamines are measured in the 3 µm range using IRMPD spectroscopy and compared with DFT simulation of the IR spectra to identify the ring conformations and the contributions of the α and β anomers to the experimental spectra. We propose an original approach for the deconvolution and quantification of heterogeneous IR spectra based on linear combination of the DFT spectra. The method is applied to the analysis of a two-component anomeric mixture and validated using NMR analysis. Finally, we further estimate the relative populations of a three-component mixture: the anomers and conformers of glucosamine.

Introduction

Ion spectroscopy has long been used for conformational analysis of small biomolecules.[1] A few years ago, its potential for analytical chemistry has emerged, in particular in the context of “omics” field where Mass Spectrometry (MS) alone fails to resolve isomeric compounds. The remarkable structural resolution of gas phase IR spectroscopy is now widely accepted as a powerful strategy for enabling “Post-isomeric barrier Mass Spectrometry”, with promising applications in glycomics and metabolomics. Early demonstrations of the resolving power of IRMPD spectroscopy (Infrared Multiple Photon Dissociation) for isobaric[2] and isomeric monosaccharides (stereoisomers and anomers)[3–7] and metabolites[8,9] have extended to more complex systems (oligosaccharides)[10–13] and more sophisticated spectroscopic schemes.[14–16]

Once restricted to synthetic standards or pure compounds, IR ion spectroscopy is now taking full advantage of the performance of MS in terms of mass separation, and is moving towards to the analysis of biological samples.[10,17] Overlapping m/z compounds often coexist in biological samples however. If they are not separated prior to spectroscopic analysis (for instance by Ion Mobility or chromatographic techniques), this results in heterogeneous IR fingerprints, which are difficult to interpret.

In this context, several strategies have been proposed to reduce the structural heterogeneity of overlapping compounds using upfront isomer-sensitive separative methods. Liquid Chromatography (LC) allows temporal separation of isomers prior to ionization and was first used in combination with IRMPD spectroscopy as a semi-separative method: fractions corresponding to isomers were separated offline and spectroscopic analysis was performed on the individual compounds.[10,17] Recent improvements in the acquisition speed to IRMPD spectra have allowed online LC-MS-IRMPD coupling.[18] Alternatively, Ion Mobility technologies offers temporal or spatial separation of ions. And has been coupled to IRMPD[19] and cryogenic IR spectroscopic schemes[20–23] for the analysis of mixtures of conformers or isomers.

Should some level of heterogeneity remain, it is essential to devise further strategies for the deconvolution and interpretation of the resulting heterogeneous spectroscopic signatures in order to identify - and ideally to quantify - the coexisting species. When the coexisting ions feature at least one resolved diagnostic band, the deconvolution is efficiently performed by experimental means. One population can then be eliminated from the mixture for quantification purpose.[24,25] Using a two-IR wavelengths scheme, Johnson and coworkers demonstrated that individual spectroscopic fingerprints can be fully reconstructed.[26] Alternatively, linear combinations of band patterns obtained by ab initio simulations can be constructed and compared to the experimental spectra. Arbitrary ratios maybe be chosen to achieve a visual match with the experimental data. Kinetic[27] and energetic[28] criteria have also been proposed for a less empirical assessment of the relative populations, but remain a matter of debate: the calculated relative energies of the coexisting structures could be used to derive relative populations at the temperature of the experiment in case of a Boltzmann distribution, but this approach is strongly biased by the existence of kinetic trapping. Besides, energetic criteria do not readily apply to the assessment of populations of isomers, which is governed by multiple factors such as steric constraints, chemical reactivity or even enzymatic expression in the case of biological samples.

Here we propose an original approach for the deconvolution and quantification of heterogeneous IR spectra. Our strategy relies on a traditional DFT simulation of the vibrational frequencies of candidate structures followed by a manual selection of the most pertinent forms based on the comparison with the experimental bands. Then, a linear combination of the DFT spectra is produced and a putative heterogeneous IR spectrum is constructed using a gaussian convolution of the linear combination. Finally, the experimental and simulated spectra are compared and the RMSD is calculated as a function of the ratio of populations. The method is applied to the analysis of two-compounds mixtures (alpha and beta anomer's of monosaccharides) and validated using NMR analysis. We further assess the relative populations of three-compounds mixtures (anomers and conformers of monosaccharides) for which NMR resolution is inadequate.

Material and methods

Samples

Glucosamine (GlcN), galactosamine (GalN) and mannosamine (ManN) were purchased from Sigma Aldrich. Samples were prepared for IRMPD spectroscopy in water:methanol solutions at a concentration of 50 µM and let to rest to reach mutarotation equilibrium. Samples were prepared for NMR in deuterated water:methanol solutions and let to rest to reach mutarotation equilibrium. The following α/β ratio were obtained by NMR analysis for validation of our IRMPD population analysis: GlcN 69/31; GalN 65/35.

Additionally, in the case of galactosamine, fresh samples were prepared and immediately submitted to IRMPD and NMR analysis before reaching mutarotation equilibrium. The α/β ratio obtained by NMR was 7/93.

IRMPD spectroscopy

IRMPD spectroscopy is performed in a modified commercial quadrupolar ion trap (Thermofinnigan LCQ classic) equipped with an electrospray source. The trapping electrode was drilled and an IR-transparent window was installed to allow irradiation of the ion cloud by

the IR laser beam produced by a YAG-pumped OPO/OPA system (LaserVision) tunable in the 3 µm region. Mass-selected ions were isolated and irradiated for 700 ms, resulting in wavelength-dependent photofragmentation. The photofragmentation yield is calculated from the resulting photofragmentation spectrum using the following variation of the Beer-Lambert formula: $-log\frac{I_p}{I_p+\sum I_f}$

were I_p is the intensity of the precursor ion and I_f is the intensity of the photofragments. The IRMPD spectrum of the precursor ion is retrieved by monitoring its photofragmentation yield as a function of the wavelength in the 2700-3700 cm^{-1} spectral range. The data points shown thorough the manuscript are averaged 3 times and a trend line (5-points Fourier Transform rolling averaging) is added to guide the eye. Note that the photofragmentation yield is not normalized by the laser power. This choice is justified by the fact that the laser power is relatively constant around 12 mJ/pulse is the spectral range of interest (OH and NH vibrations) and only decreases to 8-10 mJ/pulse in the CH range.

Simulations

Molecular dynamic was used to explore the potential energy surface of the protonated ions with the PM7[29] potential in OpenMopac[30] as described in a previous work.[2] One trajectory was calculated for each anomer. 1000 geometries were generated during each trajectory and optimized with PM7, yielding typically a few hundreds stable conformations after discarding the identical ones (based on atom positions). The geometries were further optimized with two steps of Density Functional Theory (DFT): firstly B3LYP[31–33]/6-31G*, then CAM-B3LYP[34]/6-311++G(2df,2pd).[35–37] This results in the reduction of the number of stable geometries. The energies are further refined by single-point calculations with MP2[38]/6- 311++G(3df,3pd). At this stage, the 20 lowest energy conformers were selected for computing harmonic frequencies with CAM-B3LYP/6-311++G(2df,2pd). An empirical scaling factor of 0.947 was used for comparison with the experimental IRMPD spectra. The mode analysis was performed using Gabedit,[39] Gaussian09[40] was used for DFT calculations and Orca[41] for MP2 single-point energies.

The heterogeneous IR spectra are constructed as a linear combination of individual theoretical spectra convoluted with a Gaussian:

$$I_{IR}(\nu) = \Lambda \sum_{j=1}^{j=M} \gamma_j I_j(\nu) \qquad (1.1)$$

with $I_j(\nu) = \sum_{i=1}^{Nm} I_i e^{-log(2)(\frac{f_i-\nu}{\sigma})^2}$ (1.2)

and $\sum_j \gamma_j = 1$ (1.3)

where Λ is a conversion factor between DFT intensities and IRMPD intensities; M is the number of overlapping forms contributing to the simulated IR spectrum (here anomers and conformers); γ_j is the ratio of the j^{th} form; $I_j(\nu)$ is the simulated spectrum of the j^{th} form; Nm is the number of vibrational modes present in the frequency range of interest for comparison with the IRMPD spectrum; I_i is the DFT intensity of the i^{th} mode; f_i is the DFT frequency of the i^{th} mode (scaled down with a scaling factor of 0.947); σ is the full width at half maximum (FWHM), here the value is fixed to 12 cm^{-1} to best reproduce the experimental bandwidth.

The root mean square deviation (RMSD) between the simulated IR spectrum and the IRMPD spectrum is calculated as follow:

$$RMSD = \sqrt{\frac{\sum_{\nu\in\{\nu_{exp}\}} \Delta I(\nu)^2}{N}} \qquad (2.1)$$

with $\Delta I(\nu) = I_{IR}(\nu) - I_{IRMPD}(\nu)$ (2.2)

where ν_{exp} is the set of IRMPD frequencies; N is the number of datapoints collected between 2700 and 3700 cm^{-1}); I_{IRMPD} is the IRMPD intensity.

Finally, the γ_j ratios are obtained by minimising RMSD as a function of M-1 independent γ variables in a space of M-1 dimensions.

Results and discussion

Conformational preferences

Molecular Dynamics exploration of the conformational landscape of αGlcN, βGlcN, αGalN, βGalN, αManN and βManN yielded 79, 69, 87, 40, 108 and 114 structures, respectively. After geometry re-optimisation with DFT B3LYP/6-31G*, the 20 lowest energy conformers

were selected for further geometry optimisation and computation of the IR spectra at the CAM-B3LYP/6-311++G(2df,2pd) level of theory. The 20 low energy structures of αGlcN, βGlcN, αGalN, βGalN, αManN and βManN and their IR spectra are shown in the figures of the supporting information. For both αGalN and βGalN, the lowest energy conformer is a $^4C_1$ chair with hydrogen bonds between O3H and O4H and between O4H and O6H, in agreement with the structures reported in Fraschetti *et al.*[25] For αGlcN, the lowest energy conformation is not a $^4C_1$ chair but a $^1C_4$ chair with a strong hydrogen bond between O3H and O6. The second lowest energy structure (ΔE = 7.1 kJ/mol) is a $^4C_1$ chair with a network of hydrogen bonds between O3H, O4H and O6H. For βGlcN, the most stable conformer is a $^4C_1$ chair featuring the same same network of hydrogen bonds. These three low energy structures, previously proposed in[18] were further confirmed by Scutelnic and Rizzo.[42] The case of ManN is more complex: a $^2H_1$ half chair (α anomer); a $^3S_1$ skew (β anomer) and $^4C_1$ chairs (both anomers) are present within less than 3 kJ/mol at the MP2 level of calculation.

The IRMPD spectra of protonated GlcN, GalN and ManN were previously published in Schindler *et al.*[7] and are shown in Fig.1 (top panels) in the 3000-3700 cm^{-1} spectral range, which features the OH and NH modes. For each monosaccharide, both α and β anomers coexist in the electrosprayed solution, and hence in the ion trap. Therefore, their IRMPD signatures must be interpreted as a two-component mixture at minima (two anomers), or as more complex mixtures in case of further conformational heterogeneity (as suggested by the conformational analysis). A selection of forms contributing to the IRMPD spectra is shown in Figure 1 (middle and lower panel).

The case of GalN is straightforward: the lowest energy conformation ($^4C_1$) features a distinctive pattern of OH modes between 3400 and 3700 cm^{-1} in excellent agreement with the experimental bands (Fig. 1, center panel). The NH pattern between 3100 and 3400 cm^{-1} also matches the experimental spectra and provides additional anomeric diagnostic: the low frequency asymmetric NH_3 stretching mode of αGalN reproduces the band at 3261 cm^{-1} while the one for β GalN reproduces the low intensity band at 3292 cm^{-1}. The case of ManN

is particularly difficult because several close-lying low energy conformations feature similar vibrational patterns and remain ambiguous at our spectroscopic resolution. The lowest energy structures (shown in Fig. S21) adopt a $^{2}H_{1}$ conformation characterized by four free OH bands around 3650 cm^{-1} and two NH bands at 3240 and 3330 cm^{-1}, which accounts for the majority of the experimental features. Additionally, the weak experimental feature at 3548 cm^{-1} matches an intense, diagnostic mode of a $^{4}C_{1}$ conformer (shown in Fig. 1, right panel) which suggests its presence in small amount. Note that the frequency of the symmetric NH_3^+ mode is in principle a good anomer diagnostic for this conformer (α: 3140 cm^{-1}; β: 3210 cm^{-1}). Unfortunately, this mode is relatively weak and is not observed in the experimental spectrum. For GlcN, the $^{1}C_{4}$ and $^{4}C_{1}$ conformations are easily identified: the $^{4}C_{1}$ form features a mildly H-bonded OH mode at 3550 cm^{-1} and a compact NH pattern 3180 and 3340 cm^{-1}, whereas the $^{1}C_{4}$ form features a strongly H-bonded OH mode at 3420 cm^{-1} and a spread out NH pattern between 3050 and 3350 cm^{-1}. The matching OH bands observed at 3545 and 3425 cm^{-1} in the experimental spectrum are a clear evidence of the presence of both $^{4}C_{1}$ and $^{1}C_{4}$ conformers. Further inspection of the complex experimental NH pattern reveals the presence of both anomers of the $^{4}C_{1}$ conformer. In particular, the low frequency asymmetric NH_3 stretching modes predicted at 3258 cm^{-1} (α anomer) and 3294 cm^{-1} (β anomer) match experimental features.

Finally, the IRMPD spectrum of GalN can be interpreted as a simple contribution of α and β anomers in the $^{4}C_{1}$ conformation. Unsurprisingly, ManN - which has revealed the most complex potential energy surface - shows the presence of the $^{4}C_{1}$ conformer but remains difficult to resolve in further detail. GlcN is particularly interesting: its IRMPD signature unambiguously shows the presence of three distinctive forms, namely α $^{1}C_{4}$, α $^{4}C_{1}$ and β $^{4}C_{1}$.

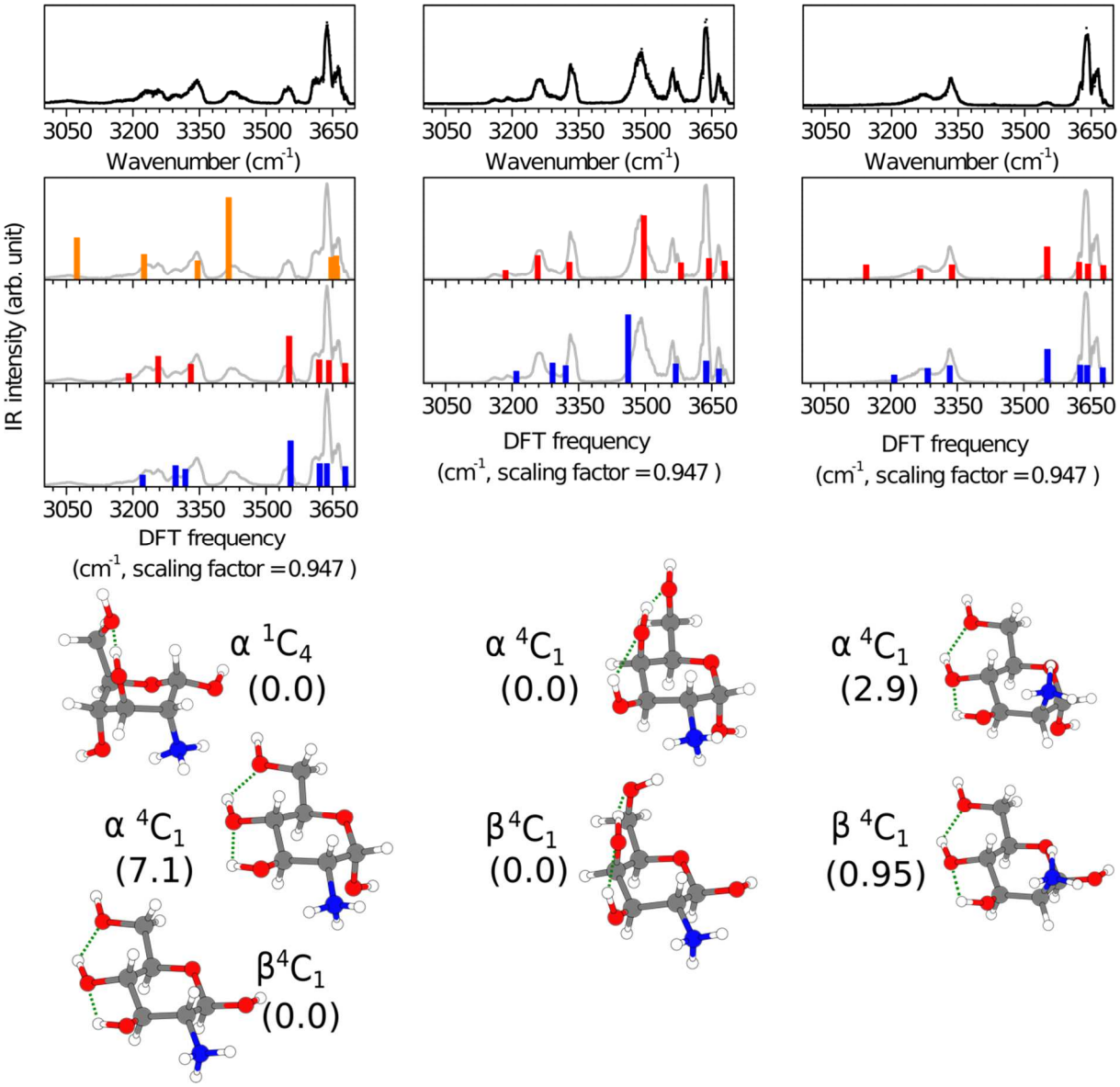


Fig. 1: IRMPD spectroscopy of GlcN (left panels), GalN (center panels) and ManN (right panels). Top: IRMPD spectra from ref 7. Center and bottom: vibrational frequencies and 3D representation of the low energy forms contributing to the experimental spectra; blue bars: β anomer; red and orange bars: α anomers; grey line: experimental trace. The energy of each form, relative to the lowest lying conformer of the corresponding molecule, is given in bracket (kJ/mol).

Population analysis

As seen above, GalN only shows anomeric heterogeneity, and its IRMPD spectrum can be interpreted as a simple two-component mixture. Moreover, the NH pattern (3100 and 3400 cm^{-1} spectral range) is distinctive of both anomers. Therefore, GalN was used to

validate our population analysis approach. The commercial dry GalN powder is conveniently sold in the β form, which allowed us to validate our analysis for two different anomeric ratios. Firstly, samples were prepared in solution and immediately submitted to NMR analysis and IRMPD spectroscopy. The α/β ratio measured by NMR was 7/93. The IRMPD spectrum obtained in these conditions is reported in Fig. 2a (green trace, T_0). Then, the solution was let to rest for 3 days to reach mutarotation equilibrium. In these conditions, the α/β ratio measured by NMR was 65/35. The IRMPD spectrum is shown in Fig. 2a (black trace, T_{eq}). As a reminder, the theoretical spectra of both anomers are reported below the experimental trace, and the frequencies previously identified as anomer-diagnostic are highlighted in red (α anomer: 3261 cm^{-1}) and blue (β anomer: 3292 cm^{-1}). The α band is weak in the IRMPD spectrum recorded at T_0 and becomes very intense at T_{eq}, while the β band is dominant at T_0 and is weak at T_{eq}. This observation is consistent with an increase of the population of α anomer in solution as mutarotation occurs. Heterogeneous, two-component theoretical spectra are constructed using the DFT frequencies of the α and β anomers, as described in the methods section. The deviation (RMSD) between the IRMPD spectra and the two-component IR spectra is plotted in Fig. 2b as a function of the relative population of β anomer. The minimal value of RMSD (best match between experiment and theory) is found for an α/β ratio of 22/78 and 58/42 at T_0 and T_{eq}, respectively. Finally, the constructed two-component spectra corresponding to these optimal ratios are shown in Fig. 2c in the diagnostic spectral range (NH modes) and compared to the IRMPD spectra for visual inspection. This representation confirms the excellent match between the experimental data and the constructed two-component spectra. At T_0, the α population (22%) is higher than the NMR value (7%). This can be explained by the different acquisition time for both methods: the "T_0" NMR value corresponds to an acquisition time of 5 minutes, while the "T_0" IRMPD value corresponds to an acquisition time of 34 minutes. The higher α population measured by IRMPD is thus consistent with the dynamics of mutarotation occurring during the experiment. At equilibrium, the populations derived from the NMR and IRMPD measurements are in good qualitative agreement.

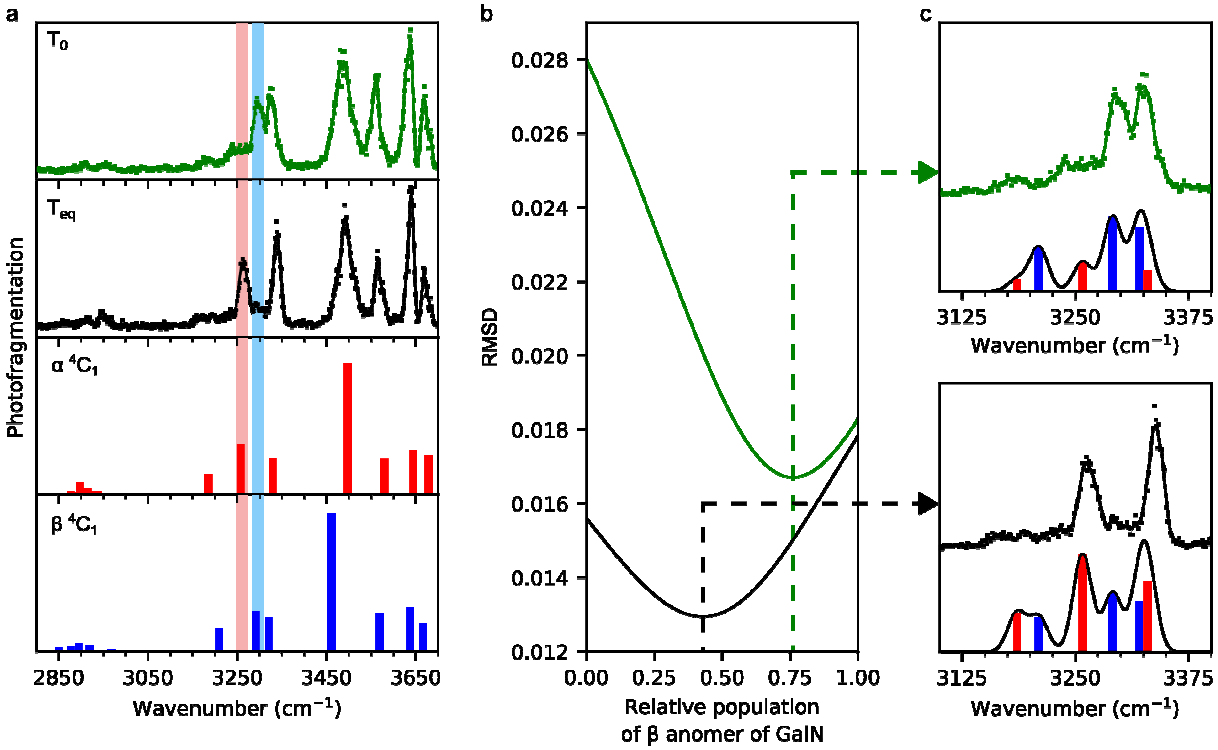


Fig. 2: analysis of the anomers population of GalN. (a) IRMPD spectra of GalN before reaching mutarotation equilibrium (green trace) and after reaching mutarotation equilibrium (black trace), compared with the simulated spectra of the α anomer (red bars) and the β anomer (blue bars). The main diagnostic modes for α and β are highlighted in red and blue, respectively. (b) graphic representation of the match between the IRMPD spectra and the simulated two-component spectrum as a function of the β/α ratio. (c) comparison of the IRMPD spectra and the simulated heterogeneous spectra at minimal RMSD (black lines). α frequencies: red bars, β frequencies: blue bars.

GlcN shows both anomeric and conformational heterogeneity and three forms were found to contribute to the IRMPD spectrum, namely α $^1C_4$, α $^4C_1$ and β $^4C_1$. The α/β ratio measured by NMR is 69/31, the ratio of conformers is not readily accessible by this technique however. As for GalN, the NH pattern (3100-3400 cm^{-1} region) offers the best spectral range for the population analysis because it is specific of both anomers and conformers. In this case, the heterogeneous spectrum is a three-component construct and a 2D RMSD surface is plotted in Fig. 3a as a function of two independent variables: the ratio of β anomer, and the ratio of

$^1C_4$ conformer. The minimum of the 2D RMSD surface is obtained for α/β=72/28 and $^1C_4/^4C_1$=31/69, which correspond to the following populations: α $^1C_4$ 31%; α $^4C_1$ 41%; β $^4C_1$ 28%. The three-component theoretical spectrum is plotted in Fig. 3b for comparison with the IRMPD spectrum. The features of the complex experimental NH pattern are remarkably well reproduced by the theoretical construct. Besides, the α/β ratio at equilibrium is in excellent agreement with the NMR value.

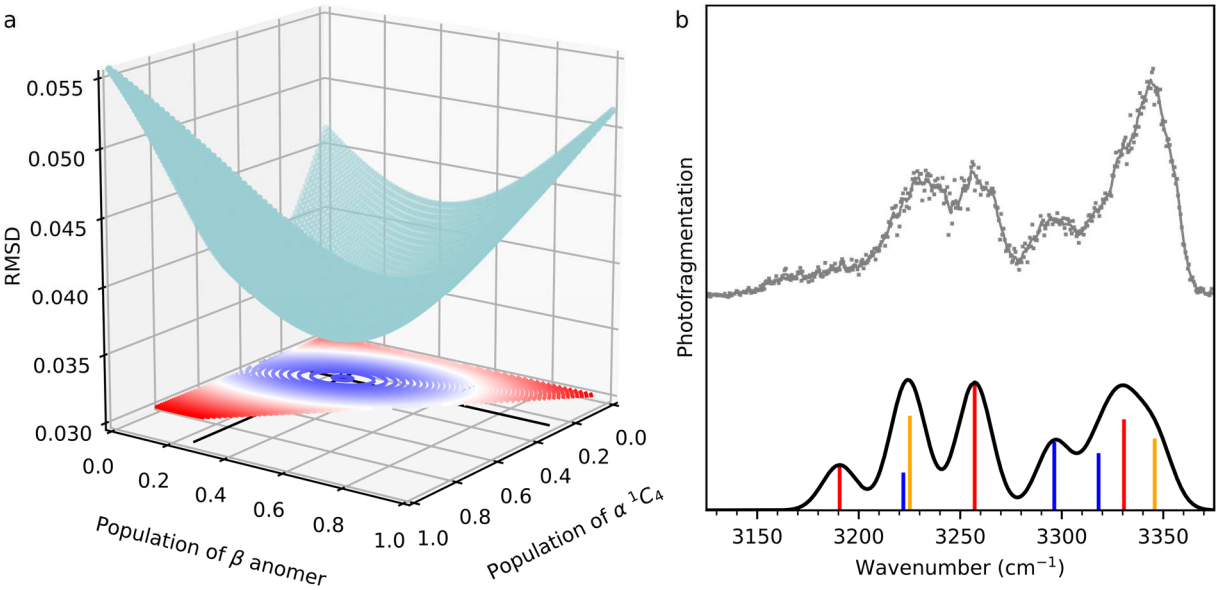


Fig. 3: analysis of the anomers and conformers population of GlcN. (a) graphic representation of the match between the IRMPD spectrum and the simulated three-component heterogeneous spectrum as a function of two independent variables. (c) comparison of the IRMPD spectra (grey trace) and the simulated three-component spectrum at minimal RMSD (black line). $^1C_4$ α frequencies: orange bars, $^4C_1$ α frequencies: red bars, $^4C_1$ β frequencies: blue bars.

This simple approach is purely empirical: it relies on the optimization of the match between the linear combination of computed IR spectra and the experimental IRMPD spectrum. When experimental features clearly originate from more than one form (as an example, we've seen here both $^1C_4$ and $^4C_1$ diagnostic bands in the spectrum of GlcN), it is tempting to combine the simulated spectra in arbitrary ratios, such as "50/50", to better account for the experimental data. Here we propose to add a quantitative dimension to this intuitive

approach: instead of introducing arbitrary ratios, the RMSD plots are used to identify optimal, empirical ratios.

In contrast with the evaluation of the populations based on their relative energy, our strategy does not rely on first principles, which might turn to an advantage if the experimental conditions are not properly described by a Boltzmann distribution. Yet, both approaches rely on the highly time-consuming exploration of the conformational space and production of equilibrium geometries in the first place. Therefore, no extra computational cost is associated to the empirical population analysis.

In the present examples, the population analysis was performed after two stages of pre-analysis. Firstly, the candidate structures contributing to the IRMPD spectra were manually selected based on the presence of diagnostic frequencies, as in traditional conformational analysis. Secondly, the fits were performed on a restricted spectral range, corresponding to the diagnostic modes (NH modes in the present examples). Modes that were common to several structures were not included in the fit (typically the free OH modes). However, we expect that these prerequisites should not limit the applicability of the approach. The pre-selection of candidate structures doesn't seem necessary, as the algorithm will readily discard any structure that does not contribute to the IRMPD spectrum. The choice of the frequency range appears more critical, especially in the case of H-bonded modes which are very intense in harmonic DFT simulations but weaker and broader, experimentally. The inclusion of H-bonded modes in the fitting procedure would thus result in a dramatic underestimation of the structures featuring such modes. This effect is due to the use of harmonic DFT frequencies however, rather than to the fitting approach itself. One can anticipate that this issue would be solved by using other theoretical approaches, such as the simulation of the IR profiles by molecular dynamics.[2]

Conclusion

The structure of gas phase protonated hexosamines was investigated by a combination of IRMPD spectroscopy and DFT simulations, and an original approach for population analysis was proposed. Both anomers are observed in the experimental spectra of hexosamines and

several ring conformations are observed at low energy: $^4C_1$ and $^1C_4$ chairs but also skew or half-chair conformations for the complex case of mannosamine.

The interpretation of multi-component IRMPD spectra has been discussed in the past in cases of conformational heterogeneity. In the case of glycans, this point can never be disregarded, as they are present in the form of anomeric mixtures. The anomeric ratio calculated using a linear combination of theoretical IR spectra are in good agreement to the ratio obtained by NMR in the case of galactosamine and glucosamine. In the case of glucosamine, where both anomeric and conformational heterogeneity are present a $^4C_1/^1C_4$ ratio for glucosamine has been proposed. This strategy does not rely on the use of a second IR laser or separative methods, making it an interesting complementary approach for population analysis.

Acknowledgements

This work was granted access to the HPC resources of the FLMSN, "Fédération Lyonnaise de Modélisation et Sciences Numériques", partner of EQUIPEX EQUIP@MESO and to the "Centre de calcul CC-IN2P3" at Villeurbanne, France.

References

1 A. M. Rijs and J. Oomens, in *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, eds. A. M. Rijs and J. Oomens, Springer International Publishing, Cham, 2014, vol. 364, pp. 1–42.

2 B. Schindler, J. Joshi, A.-R. Allouche, D. Simon, S. Chambert, V. Brites, M.-P. Gaigeot and I. Compagnon, Distinguishing isobaric phosphated and sulfated carbohydrates by coupling of mass spectrometry with gas phase vibrational spectroscopy, *Physical Chemistry Chemical Physics*, 2014, **16**, 22131–22138.

3 E. B. Cagmat, J. Szczepanski, W. L. Pearson, D. H. Powell, J. R. Eyler and N. C. Polfer, Vibrational signatures of metal-chelated monosaccharide epimers: gas-phase infrared spectroscopy of Rb+-tagged glucuronic and iduronic acid, *Physical Chemistry Chemical Physics*, 2010, **12**, 3474–3479.

4 Wright L. Pearson, C. Contreras, D. Powell, G. Berden, J. Oomens, B. Bendiak and J. R. Eyler, Differentiation of Rubidiated Methyl-D-Glycoside Stereoisomers by Infrared Multiple-

Photon Dissociation Spectroscopy in the O–H and C–H Stretching Regions, *The Journal of Physical Chemistry B*, 2015, **119**, 12970–12981.

5 L. Barnes, B. Schindler, S. Chambert, A.-R. Allouche and I. Compagnon, Conformational preferences of protonated N-acetylated hexosamines probed by InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy and ab initio calculations, *International Journal of Mass Spectrometry*, 2017, **421**, 116 – 123.

6 B. Schindler, G. Renois-Predelus, N. Bagdadi, S. Melizi, L. Barnes, S. Chambert, A.-R. Allouche and I. Compagnon, MS/IR, a new MS-based hyphenated method for analysis of hexuronic acid epimers in glycosaminoglycans, *Glycoconjugate Journal*, 2017, **34**, 421–425.

7 B. Schindler, L. Barnes, G. Renois, C. Gray, S. Chambert, S. Fort, S. Flitsch, C. Loison, A.-R. Allouche and I. Compagnon, Anomeric memory of the glycosidic bond upon fragmentation and its consequences for carbohydrate sequencing, *Nature Communications*, 2017, **8**, 973.

8 J. Martens, G. Berden, H. Bentlage, K. L. M. Coene, U. F. Engelke, D. Wishart, M. van Scherpenzeel, L. A. J. Kluijtmans, R. A. Wevers and J. Oomens, Unraveling the unknown areas of the human metabolome: the role of infrared ion spectroscopy, *J Inherit Metab Dis*, 2018, **41**, 367–377.

9 A. P. Cismesia, M. R. Bell, L. F. Tesler, M. Alves and N. C. Polfer, Infrared ion spectroscopy: an analytical tool for the study of metabolites, *Analyst*, 2018, **143**, 1615–1623.

10 J. Wattjes, B. Schindler, S. Trombotto, L. David, B. M. Moerschbacher and I. Compagnon, Discrimination of patterns of N-acetylation in chitooligosaccharides by gas phase IR spectroscopy integrated to mass spectrometry, *Pure and Applied Chemistry*, 2017, **89**, 1349–1357.

11 A. Depraz Depland, G. Renois-Predelus, B. Schindler and I. Compagnon, Identification of sialic acid linkage isomers in glycans using coupled InfraRed Multiple Photon Dissociation

(IRMPD) spectroscopy and mass spectrometry, *International Journal of Mass Spectrometry*, 2018, **434**, 65–69.

12 G. Renois-Predelus, B. Schindler and I. Compagnon, Analysis of Sulfate Patterns in Glycosaminoglycan Oligosaccharides by MSn Coupled to Infrared Ion Spectroscopy: the Case of GalNAc4S and GalNAc6S, *Journal of The American Society for Mass Spectrometry*, 2018, **29**, 1242–1249.

13 E. Mucha, M. Lettow, M. Marianski, D. A. Thomas, W. B. Struwe, D. J. Harvey, G. Meijer, P. H. Seeberger, G. von Helden and K. Pagel, Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angewandte Chemie International Edition*, 2018, **57**, 7440–7443.

14 E. Mucha, A. I. González Flórez, M. Marianski, D. A. Thomas, W. Hoffmann, W. B. Struwe, H. S. Hahm, S. Gewinner, W. Schöllkopf, P. H. Seeberger, G. von Helden and K. Pagel, Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy, *Angewandte Chemie International Edition*, 2017, **56**, 11248–11251.

15 N. Khanal, C. Masellis, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS, *Analytical Chemistry*, 2017, **89**, 7601–7606.

16 C. Masellis, N. Khanal, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 2217–2222.

17 J. Martens, V. Koppen, G. Berden, F. Cuyckens and J. Oomens, Combined Liquid Chromatography-Infrared Ion Spectroscopy for Identification of Regioisomeric Drug Metabolites, *Analytical Chemistry*, 2017, **89**, 4359–4362.

18 B. Schindler, G. Laloy-Borgna, L. Barnes, A.-R. Allouche, E. Bouju, V. Dugas, C. Demesmay and I. Compagnon, Online Separation and Identification of Isomers Using Infrared Multiple Photon Dissociation Ion Spectroscopy Coupled to Liquid Chromatography: Application to the Analysis of Disaccharides Regio-Isomers and Monosaccharide Anomers, *Analytical Chemistry*, 2018, **90**, 11741–11745.

19 B. Schindler, A. D. Depland, G. Renois-Predelus, G. Karras, B. Concina, G. Celep, J. Maurelli, V. Loriot, E. Constant, R. Bredy, C. Bordas, F. Lépine and I. Compagnon, FAIMS-MS-IR spectroscopy workflow: a multidimensional platform for the analysis of molecular isoforms, *International Journal for Ion Mobility Spectrometry*, 2017, **20**, 119–124.

20 L. Voronina and T. R. Rizzo, Spectroscopic studies of kinetically trapped conformations in the gas phase: the case of triply protonated bradykinin, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25828–25836.

21 J. Seo, W. Hoffmann, S. Warnke, X. Huang, S. Gewinner, W. Schöllkopf, M. T. Bowers, G. von Helden and K. Pagel, An infrared spectroscopy approach to follow β-sheet formation in peptide amyloid assemblies, *Nature Chemistry*, 2016, **9**, 39.

22 M. Z. Kamrath and T. R. Rizzo, Combining Ion Mobility and Cryogenic Spectroscopy for Structural and Analytical Studies of Biomolecular Ions, *Accounts of Chemical Research*, 2018, **51**, 1487–1495.

23 A. Ben Faleh, S. Warnke and T. R. Rizzo, Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal. Chem.*, 2019, **91**, 4876–4882.

24 J. S. Prell, T. M. Chang, J. A. Biles, G. Berden, J. Oomens and E. R. Williams, Isomer Population Analysis of Gaseous Ions From Infrared Multiple Photon Dissociation Kinetics, *J. Phys. Chem. A*, 2011, **115**, 2745–2751.

25 C. Fraschetti, L. Guarcini, C. Zazza, L. Mannina, S. Circi, S. Piccirillo, B. Chiavarino and A. Filippi, Real time evolution of unprotected protonated galactosamine probed by IRMPD spectroscopy, *Physical Chemistry Chemical Physics*, 2018, **20**, 8737–8743.

26 C. M. Leavitt, A. B. Wolk, J. A. Fournier, M. Z. Kamrath, E. Garand, M. J. Van Stipdonk and M. A. Johnson, Isomer-Specific IR–IR Double Resonance Spectroscopy of D_2-Tagged Protonated Dipeptides Prepared in a Cryogenic Ion Trap, *The Journal of Physical Chemistry Letters*, 2012, **3**, 1099–1105.

27 M. U. Munshi, J. Martens, G. Berden and J. Oomens, Protoisomerization of Indigo and Isoindigo Dyes Confirmed by Gas-Phase Infrared Ion Spectroscopy, *J. Phys. Chem. A*, 2019, **123**, 8226–8233.

28 B. Lucas, G. Grégoire, J. Lemaire, P. Maître, F. Glotin, J. P. Schermann and C. Desfrançois, Infrared multiphoton dissociation spectroscopy of protonated N-acetyl-alanine and alanyl-histidine, *International Journal of Mass Spectrometry*, 2005, **243**, 105–113.

29 J. J. P. Stewart, Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters, *Journal of Molecular Modeling*, 2013, **19**, 1–32.

30 Stewart, J. J. P., *MOPAC2012*, 2012.

31 A. D. Becke, Density functional exchange energy approximation with correct asymptotic behavior, *Physical Review A*, 1988, **38**, 3098–3100.

32 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Physical review B*, 1988, **37**, 785.

33 A. D. Becke, A new mixing of hartree-fock and local density-functional theories, *Journal of Chemical Physics*, 1993, **98**, 1372–1377.

34 T. Yanai, D. P. Tew and N. C. Handy, A new hybrid exchange correlation functional using the Coulomb attenuating method (CAM-B3LYP), *Chemical Physics Letters*, 2004, **393**, 51–57.

35 W. J. Hehre, R. Ditchfield and J. A. Pople, Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules, *The Journal of Chemical Physics*, 1972, **56**, 2257–2261.

36 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, Self consistent molecular orbital methods .20. basis set for correlated wave functions, *Journal of Chemical Physics*, 1980, **72**, 650–654.

37 M. J. Frisch, J. A. Pople and J. S. Binkley, Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets, *The Journal of Chemical Physics*, 1984, **80**, 3265–3269.

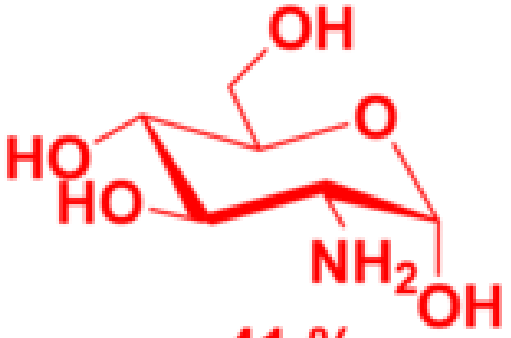
38 C. Moller and M. S. Plesset, Note on an approximation treatment for many-electron systems, *Physical Review*, 1934, **46**, 0618–0622.

39 A.-R. Allouche, Gabedit-A graphical user interface for computational chemistry softwares, *Journal of Computational Chemistry*, 2011, **32**, 174–182.

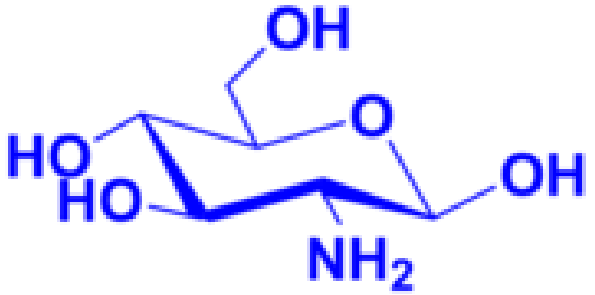
40 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09*, Gaussian, Inc., Wallingford, CT, USA, 2009.

41 F. Neese, The ORCA program system, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2012, **2**, 73–78.

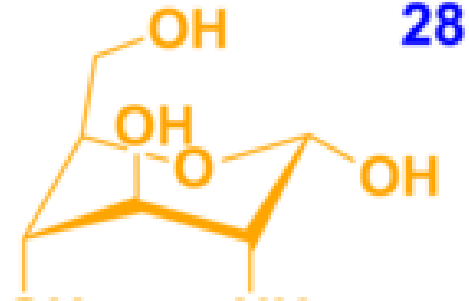
42 V. Scutelnic and T. R. Rizzo, Cryogenic Ion Spectroscopy for Identification of Monosaccharide Anomers, *J. Phys. Chem. A*, 2019, **123**, 2815–2819.


OH
O
HO
HO
NH2
OH

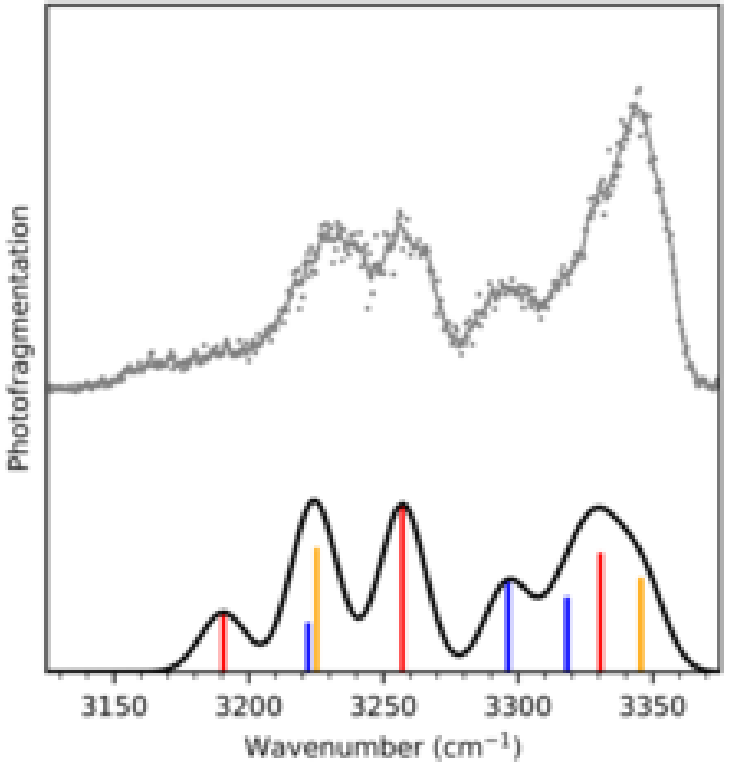
41 %


OH
O
HO
HO
OH
NH2

28 %


OH
OH
O
OH
OH
NH2

31 %


Photofragmentation
3150
3200
3250
3300
3350
Wavenumber (cm−1)