Combining Cryogenic Infrared Spectroscopy with Selective Enzymatic Cleavage for Determining Glycan Primary Structure

Irina Dyukova, Eduardo Carrascosa, Robert P. Pellegrinelli, and Thomas R. Rizzo*

Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, EPFL SB ISIC LCPM, Station 6, CH-1015 Lausanne, Switzerland.

ABSTRACT: Given the biological relevance and intrinsic structural complexity of glycans, increasing efforts are being directed towards developing a general glycan database that includes information from different analytical methods. As recently demonstrated, cryogenic infrared (IR) spectroscopy is a promising technique for glycan analysis, as it provides unique vibrational fingerprints of specific glycan isomer ions. One of the main goals of a glycan database is the identification and detailed characterization of unknown species. In this work, we combine enzymatic digestion with cryogenic IR-spectroscopy and demonstrate how it can be used for glycan identification. We measured the IR-spectra of a series of cationic glycan standards of increasing complexity and compared them with spectra of the same species after enzymatic cleavage of larger glycans. We show that the cryogenic IR spectra of the cleaved glycans are highly structured and virtually identical to those of standards after both single and multiple cleavages. Our results suggest that the combination of these methods represents a potentially powerful and specific approach for the characterization of unknown glycans.

INTRODUCTION

The surface of nearly every cell of living organisms is covered by a layer of glycans that play a central role in the cell-to-cell recognition and signaling.[1,2] Glycans are responsible for triggering cellular immune response and inflammatory processes to protect cells from viruses and bacteria.[3] Moreover, they are involved in most major diseases,[2,4,5] and their characterization is essential for assessing the quality of biotherapeutics.[6] As the biological function of glycans is directly related to their structure, many different approaches have been introduced to unravel their structural and isomeric complexity. Commonly used methods include liquid chromatography (LC)[7-11] with enzymatic digestion,[9,12-14] mass spectrometry (MS),[15-18] and ion mobility spectrometry (IMS).[19-23] Current glycan databases[24-26] that include information from these techniques represent powerful tools for fundamental research in chemistry and molecular biology as well as for the pharmaceutical industry.

While the above-mentioned experimental approaches cannot distinguish all types of glycan isomerism when used alone, combining them often provides a more complete structural characterization. For instance, coupling IMS or LC to MS allows one to determine many glycan isomers that are indistinguishable by MS alone.[20-23] Nevertheless, these so-called hyphenated techniques are blind to many of the subtle structural details that distinguish isomeric glycans.[21] In the last few years vibrational spectroscopy has been proven to provide unique structural fingerprints of isolated glycans.[27-33] We have recently combined IMS-MS with cryogenic messenger-tagging IR-spectroscopy to obtain highly resolved, isomer-specific vibrational spectra of glycan ions,[28,34,35] which is not possible with room temperature spectroscopic techniques.[29-31] Our goal is to complement currently existing databases by adding a vibrational fingerprint for each glycan species. While this spectroscopic database will be constructed initially using standard/known glycans, one needs a mechanism to associate a vibrational fingerprint with glycan structures that are presently unknown or those for which standards are not easily obtained.

For the identification of unknown glycans, we propose a bottom-up approach using cryogenic spectroscopy to identify specifically cleaved glycan fragments. Currently, the most standard technique for determining glycan structures from their fragments is to use MS^n together with collision-induced dissociation (CID).[17,29,30,36-38] Using infrared multiphoton dissociation (IRMPD) spectroscopy, Compagnon and coworkers recently demonstrated that the anomericity of glycosidic linkages can be retained in the dissociation process,[36] which is important structural information. Nevertheless, CID and IRMPD are intrinsically non-specific fragmentation techniques, and the multitude of isomeric fragments can lead to an ambiguous interpretation of the parent structure.

In contrast, the more time-consuming enzymatic digestion approach allows identification of complex glycans through a series of bond-specific cleavages.[9,12-14] Unknown glycans can be sequenced by a set of exoglycosidases, which selectively cleave monosaccharides from the non-reducing end. By degrading a parent glycan until its spectrum is one that is in our database, and knowing the isomeric specificity of the applied enzymes, one can reconstruct the primary structure of the unknown parent glycan and then add its vibrational spectrum to the database.

The aim of this work is to evaluate the feasibility of combining exoglycosidase digestion with cryogenic IR-spectroscopy for glycan identification. For this purpose, we compare vibrational spectra of four representative reference glycans with those produced after single or multiple enzymatic cleavage. To demonstrate our approach NA2 N-linked glycan was used as a model for an unknown glycan.

EXPERIMENTAL APPROACH

Exoglycosidase digestion of N-linked glycans

In the present work, we monitor the enzymatic degradation of glycans using UPLC-MS. We chose the N-linked glycans Man-1, Man-3, NGA2, and NA2 (Figure 1) as standards.

Figure 1. The schematic structures of N-linked glycans used in this study.

All glycan reference compounds were purchased from Dextra (UK) and exoglycosidases (β(1-3,4)-galactosidase (Bovine testis), α(1-2,3,6)-mannosidase (Jack bean), β-N-acetylhexosaminidase (Jack bean)) were purchased from Prozyme (Denmark). Glycan purity is certified at a minimum of 85 %. Enzymatic exoglycosylations were carried out in a total glycan concentration of 0.4 mg/mL. A solution of oligosaccharide and exoglycosidase in 50 μL of buffer solution was incubated at 37 °C for various periods of time (Table S1, Supporting Information).

After digestion, all samples were analyzed and purified by hydrophilic interaction liquid chromatography using an AQUITY UPLC H-Class Plus System (Waters) coupled to a Q-TOF mass spectrometer (Waters Premier). We used an XBridge Glycan BEH Amide Column (Waters) with a bisolvent system as a mobile phase – acetonitrile and ammonium formate buffer solution (100 mM, pH = 4-5). After digestion, glycans were collected using a Waters Fraction Collector III coupled to the UPLC-MS system, and the solvents were evaporated to concentrate the samples up to 15-20 μM final concentration, calculated assuming 100% elution of samples from the column.

Cryogenic tagging IR-spectroscopy

We performed our spectroscopy experiments on a home-built, cryogenic, tandem mass spectrometer, a detailed description of which has been reported previously.[39] Briefly, cationized glycans are produced *via* nanoelectrospray ionization (nESI) of 15-20 μM standard and cleaved glycan samples (Man-1, Man-3, NGA2, NA2) dissolved in 50:50 water/acetonitrile. The ions of a particular *m/z* ratio are selected in a quadrupole mass filter and guided into a cryogenic octupole ion trap maintained at 40-50 K, where they are cooled in collisions with a mixture of nitrogen and helium buffer gases (1:9) and tagged with N_2. The nitrogen-tagged ions are then irradiated with an infrared pulse from a Nd:YAG pumped tunable optical parametric oscillator (OPO). When the OPO wavenumber is tuned in resonance with a vibrational transition, IR radiation is absorbed and then internally redistributed, causing the weakly bound N_2 molecule to evaporate. Monitoring the depletion of the tagged ions as a function of the IR wavenumber allows us to obtain an infrared absorption spectrum.[40] The acquisition time for each scan was 27.30 min for the range 3200—3750 cm^{-1}. We analyzed our samples at 10—15 μM concentration using nano-ESI. With a flow rate for nano-ESI in the range of 1—100 nL/min, the total amount of sample used for this experiment is approximately 0.5—40 pmol.

RESULTS AND DISCUSSION

Exoglycosidase digestion of N-linked glycans

For the N-glycan core, Man-3, we performed a single exoenzyme cleavage with α(1-2,3,6)-mannosidase. During the incubation period, two mannose residues were removed from the non-reducing end of Man-3, leaving Man-1. The chromatogram (Figure 2B) of the mixture after cleavage confirms the full conversion of Man-3 to Man-1, showing a single peak for Man-1 with retention time 6.07 min.

For NGA2 a single enzymatic cleavage with β-N-acetylhexosaminidase was carried out. During the incubation period, two N-acetylglucosamine monosaccharides were removed from the non-reducing end of NGA2, and the resulting smaller glycan was Man-3. The chromatogram (Figure 2C) of the mixture after cleavage confirms the full conversion of NGA2 to Man-3, showing a single peak with retention time 10.82 min.

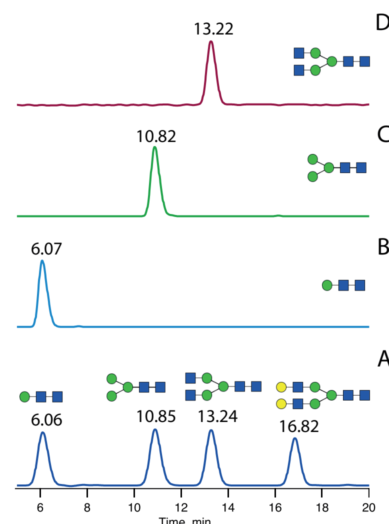

Figure 2. Chromatograms of a mixture of glycan standards (A) and smaller glycans after enzymatic cleavage (B–D). B – conversion of Man-3 to Man-1; C – conversion of NGA2 and NA2 to Man-3; D – conversion of NA2 to NGA2.

For NA2 we performed single cleavage with β(1-3,4)-galactosidase and multiple exoenzyme cleavage using a combination of β(1-3,4)-galactosidase and β-N-acetylhexosaminidase. The chromatogram (Figure 2D) of the mixture after cleavage confirms the full conversion of NA2 to NGA2, showing a single peak for NGA2 with retention time 13.22 min. After multiple digestion we observed the same chromatogram (Figure 2C) as was obtained after a single exoenzyme cleavage of NGA2. The observation of a single peak with retention time 10.82 min confirms the full conversion of NA2 to Man-3 in which two galactose units and two N-acetylglucosamine monosaccharides were removed from the non-reducing end.

Infrared spectroscopy of glycan standards and smaller glycans after enzymatic cleavage

Since N-linked glycans, such as those investigated in this work contain at least two N-acetylglucosamine monosaccharides, they can be easily protonated in acidic medium. Since we use low pH (4-5) during the clean-up procedure after enzymatic cleavage, the corresponding protonated glycans were the predominant species observed in all mass spectra of our collected samples. The IR spectra of protonated glycan stand-

ards and glycans after enzymatic cleavage were measured using the messenger tagging technique, monitoring the depletion of the weakly bound N_2-tagged molecules.

Man-1 IR-spectra

We performed a single enzymatic digestion of Man-3 with α(1-2,3,6)-mannosidase, which converts all of the starting material to Man-1, as shown in Figure 2B. Figure 3 shows the measured IR-spectra of this cleaved Man-1 along with that from the Man-1 reference compound. Both consist of a number of overlapping transitions in the range 3450–3550 cm^{-1} and four distinct lines in the "free" OH-stretch region (3550–3700 cm^{-1}). The good agreement in both the position and intensities of the vibrational bands of the reference and cleaved Man-1 confirms the identity of the degradation product.

Figure 4 shows the corresponding spectra of the sodiated Man-1. As in the case of the protonated species, the spectrum of the enzymatic digestion product matches that of the reference compound unambiguously. One expects the spectra of the sodiated and protonated species to differ, since the proton and sodium cation will bind differently, leading to a different geometry and hence different vibrational frequencies. The spectral matching of both the protonated and sodiated species with their respective standards provide independent information with which to identify the enzymatic digestion product.

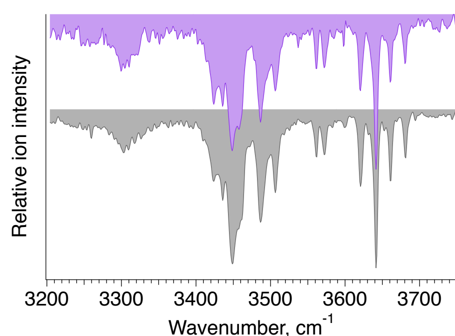

Figure 3. IR-spectra of protonated Man-1 reference (purple) and Man-1 after enzymatic cleavage of Man-3 (grey).

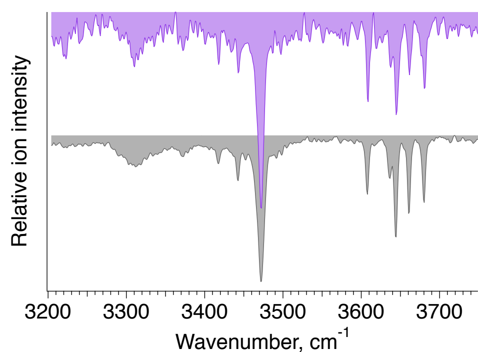

Figure 4. IR-spectra of sodiated Man-1 reference (purple) and Man-1 after enzymatic cleavage of Man-3 (grey).

To date, thorough high-level theoretical calculations of glycan structure have been restricted to disaccharides, mainly due to the challenges associated with long computational times and with carefully sampling the full conformational space of such complex molecular ions. Recently, first molecular dynamics investigations have been performed to understand the structure and isomeric complexity of larger glycans.[33,41] Because we are interested only in the primary (i.e., covalent) structure of glycans including all isomeric forms, rather than comparing our measurements with calculated spectra, we take a completely different approach. We aim to identify an unknown glycan molecule by measuring its mass and vibrational fingerprint spectrum and comparing them with entries in a database that we will construct. This means that our determination of primary structure does not rely upon being able to perform accurate, high-level, quantum-chemical calculations.

Man-3 IR-spectra

Man-3 was obtained by both the single digestion of NGA2 with β-N-acetylhexosaminidase and multiple digestion of NA2 with a mixture of β(1-3,4)-galactosidase and β-N-acetylhexosaminidase. We measured the IR-spectra of these cleaved Man-3 oligosaccharides and compared them with the reference Man-3 spectra for protonated (Figure 5), ammoniated and sodiated adducts (Figure S3-S4).

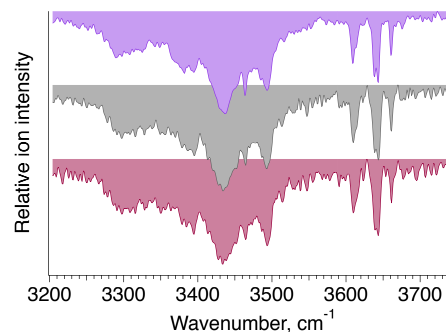

Figure 5. IR-spectra of protonated Man-3 reference (purple), Man-3 after enzymatic cleavage of NGA2 (grey) and Man-3 after multiple enzymatic cleavage of NA2 (red).

The spectra of Man-3 after single enzymatic digestion of NGA2 and multiple digestion of NA2 are in excellent agreement with the reference spectra of Man-3. These results demonstrate that the spectra of smaller glycans do not depend on the means by which they are produced and reinforce the practicality of using a spectroscopic database to identify them.

NGA2 IR-spectra

For NA2 we performed the single digestion with β(1-3,4)-galactosidase and obtained the smaller glycan NGA2. We show the IR-spectra of the cleaved species together with that of the standard for protonated (Figure 6) and sodiated (Figure S5) NGA2.

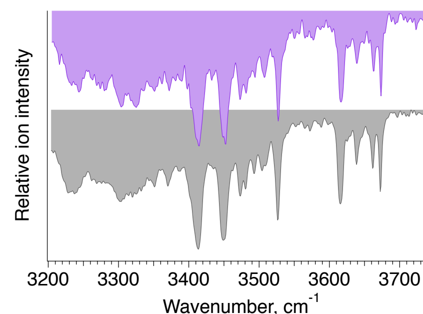

Figure 6. IR-spectra of protonated NGA2 reference (purple) and NGA2 after enzymatic cleavage of NA2 (grey).

NA2 IR-spectrum

The goal of our approach is to identify either unknown glycans or known glycans for which we cannot obtain standards. In the present case, we have used NA2 as a model for an unknown glycan to demonstrate our procedure. Once we have identified the target parent glycan, we can then add its cryogenic IR spectrum to our database and identify it subsequently

by its spectroscopic fingerprint. The spectrum of the NA2 parent, shown in Figure 7, is sufficiently structured that it would serve well as such a fingerprint. Moreover, it provides the basis for identifying larger glycans that we can enzymatically degrade to NA2.

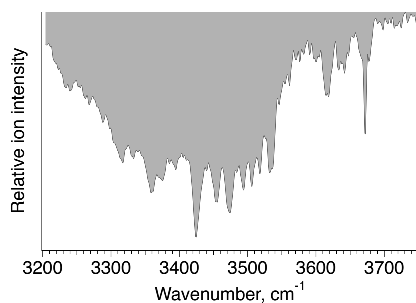

Figure 7: IR-spectrum of protonated NA2 reference.

CONCLUSIONS

The use of cryogenic vibrational spectroscopy for glycan identification requires the development of a spectroscopic database. While we will initially measure the spectra of as many known glycan structures as possible, one major goal is to analyze glycans from biological samples, which will certainly include species that are not (at least initially) contained in our database. By combining the existing protocols for exoglycosidase cleavage with cryogenic infrared spectroscopy, we have developed a mechanism by which to determine unknown glycan primary structures and add them to the database. The fact that glycans as large as those investigated in this work still show discrete spectral lines by which we can identify them bodes well for the principle of using a spectroscopic database for glycan analysis.

It is important to note that the methodology described in this work represents a procedure for constructing rather than using the database. We only have to go through this time-consuming process once for each unknown species. Each time we add a new species to the database, we gain the ability to identify it from a mixture on the basis of its spectral fingerprint and obtain a new core structure that can be used in identifying still larger glycans that are built upon it. By initially focusing on N-linked glycans, we should be able to quickly construct a functional database that will progressively expand. We also plan to include cryogenic IR spectra from glycan fragments generated by CID, as these will provide complementary information for determining glycan structure.

In conclusion, the present work demonstrates the feasibility of combining enzymatic digestion with cryogenic IR-spectroscopy for glycan analysis. The IR-spectra of enzymatically cleaved glycans do not depend on enzymatic cleavage procedure and are in very good agreement with those of reference standards. Even though the glycans investigated in this work have as many as nine monosaccharide units, all the spectra exhibit sufficiently distinct spectral fingerprints that can serve as unique identifiers. These results serve as a starting point for constructing a glycan database and can be useful to benchmark future electronic structure and vibrational frequency calculations of increasingly complex glycans.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI:

Conditions for exoglycosidase cleavage of glycans, mass spectrum of the glycan after enzymatic cleavage, IR-spectra of ammoniated and sodiated glycan adducts. (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: thomas.rizzo@epfl.ch.

ORCID

Eduardo Carrascosa 0000-0003-4338-8669

Thomas R. Rizzo: 0000-0003-2796-905X

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank the European Research Council (Grant 788697-GLYCANAL) and the EPFL for their generous support of this work.

REFERENCES

(1) Varki, A. *Glycobiology* **1993**, *3*, 97-130.
(2) Varki, A. *Glycobiology* **2017**, *27*, 3-49.
(3) Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370-2376.
(4) *Transforming Glycoscience: A Roadmap for the Future*; The National Academies Press: Washington, D.C., 2012.
(5) Dwek, R. A. *Chemical Reviews* **1996**, *96*, 683-720.
(6) Dalziel, M.; Crispin, M.; Scanlan, C. N.; Zitzmann, N.; Dwek, R. A. *Science* **2014**, *343*, 37-+.
(7) Stockmann, H.; Duke, R. M.; Martin, S. M.; Rudd, P. M. *Anal. Chem.* **2015**, *87*, 8316-8322.
(8) Buszewski, B.; Noga, S. *Anal. Bioanal. Chem.* **2012**, *402*, 231-247.
(9) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. *Nat Chem Biol* **2010**, *6*, 713-723.
(10) Liu, S.; Gao, W. J.; Wang, Y.; He, Z. Y.; Feng, X. J.; Liu, B. F.; Liu, X. *Plos One* **2017**, *12*.
(11) Shah, B.; Jiang, X. G.; Chen, L.; Zhang, Z. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 999-1011.
(12) Gotz, L.; Abrahams, J. L.; Mariethoz, J.; Rudd, P. M.; Karlsson, N. G.; Packer, N. H.; Campbell, M. P.; Lisacek, F. *Bioinformatics* **2014**, *30*, 3131-3133.
(13) Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. In *Glycobiology Protocols*, Brockhausen, I., Ed.; Humana Press: Totowa, NJ, 2007, pp 125-143.
(14) Walsh, I.; Nguyen-Khuong, T.; Wongtrakul-Kish, K.; Tay, S. J.; Chew, D.; Josee, T.; Taron, C. H.; Rudd, P. M. *Bioinformatics* **2019**, *35*, 688-690.
(15) Han, L.; Costello, C. E. *Biochemistry-Moscow* **2013**, *78*, 710-720.
(16) Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J. *Anal. Chem.* **2014**, *86*, 196-212.
(17) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.
(18) Yu, X.; Jiang, Y.; Chen, Y. J.; Huang, Y. Q.; Costello, C. E.; Lin, C. *Anal. Chem.* **2013**, *85*, 10017-10021.
(19) Gray, C. J.; Thomas, B.; Upton, R.; Migas, L. G.; Eyers, C. E.; Barran, P. E.; Flitsch, S. L. *Biochim Biophys Acta* **2016**, *1860*, 1688-1709.
(20) Pagel, K.; Harvey, D. J. *Anal. Chem.* **2013**, *85*, 5138-5145.
(21) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. *Nature* **2015**, *526*, 214-244.
(22) Liu, Y. S.; Clemmer, D. E. *Anal. Chem.* **1997**, *69*, 2504-2509.
(23) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. *Anal. Chem.* **2019**, *91*, 12030-12037.
(24) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Bioinformatics* **2008**, *24*, 1214-1216.

(25) Hayes, C. A.; Karlsson, N. G.; Struwe, W. B.; Lisacek, F.; Rudd, P. M.; Packer, N. H.; Campbell, M. P. *Bioinformatics* **2011**, *27*, 1343-1344.
(26) Struwe, W. B.; Pagel, K.; Benesch, J. L. P.; Harvey, D. J.; Campbell, M. P. *Glycoconjugate Journal* **2016**, *33*, 399-404.
(27) Masellis, C.; Khanal, N.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2217-2222.
(28) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. *Anal. Chem.* **2019**, *91*, 4876-4882.
(29) Hernandez, O.; Isenberg, S.; Steinmetz, V.; Glish, G. L.; Maitre, P. *J. Phys. Chem. A* **2015**, *119*, 6057-6064.
(30) Schindler, B.; Barnes, L.; Gray, C. J.; Chambert, S.; Flitsch, S. L.; Oomens, J.; Daniel, R.; Allouche, A. R.; Compagnon, I. *J. Phys. Chem. A* **2017**, *121*, 2114-2120.
(31) Tan, Y.; Zhao, N.; Liu, J.; Li, P.; Stedwell, C. N.; Yu, L.; Polfer, N. C. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 539-550.
(32) Mucha, E.; González Flórez, A. I.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schöllkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K. *Angew. Chem. Int. Ed.* **2017**, *56*, 11248-11251.
(33) Mucha, E.; Stuckmann, A.; Marianski, M.; Struwe, W. B.; Meijer, G.; Pagel, K. *Chemical Science* **2019**.
(34) Warnke, S.; Ben Faleh, A.; Scutelnic, V.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2204-2211.
(35) Warnke, S.; Ben Faleh, A.; Pellegrinelli, R. P.; Yalovenko, N.; Rizzo, T. R. *Faraday Discuss. Chem. Soc.* **2019**, *217*, 114-125.
(36) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. *Anal. Chem.* **2017**, *89*, 4540-4549.
(37) Schindler, B.; Renois-Predelus, G.; Bagdadi, N.; Melizi, S.; Barnes, L.; Chambert, S.; Allouche, A.-R.; Compagnon, I. *Glycoconjugate Journal* **2017**, *34*, 421-425.
(38) Tsai, S. T.; Liew, C. Y.; Hsu, C.; Huang, S. P.; Weng, W. C.; Kuo, Y. H.; Ni, C. K. *Chembiochem* **2019**, *20*, 2351-2359.
(39) Svendsen, A.; Lorenz, U. J.; Boyarkin, O. V.; Rizzo, T. R. *Rev. Sci. Instrum.* **2010**, *81*, 073107.
(40) Kamrath, M. Z.; Garand, E.; Jordan, P. A.; Leavitt, C. M.; Wolk, A. B.; Van Stipdonk, M. J.; Miller, S. J.; Johnson, M. A. *J. Am. Chem. Soc.* **2011**, *133*, 6440-6448.
(41) Struwe, W. B.; Baldauf, C.; Hofmann, J.; Rudd, P. M.; Pagel, K. *Chem. Commun.* **2016**, *52*, 12353-12356.

For table of contents only: