Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 141–145
Carbohydrate RESEARCH

Oligosaccharide identification and mixture quantification using Raman spectroscopy and chemometric analysis

Melissa F. Mrozek, Dongmao Zhang and Dor Ben-Amotz*

Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA

Received 21 May 2003; accepted 11 September 2003

Abstract—This work demonstrates the feasibility of using Raman spectroscopy for the analysis of small quantities of chemically similar oligosaccharides and their mixtures. Raman spectra were obtained from 10-μL aliquots of 1 mM solutions of maltotetraose and/or stachyose after deposition onto an electrochemically roughened silver substrate (and the resulting spectral features are attributed to a combination of normal and surface-enhanced Raman scattering). These compounds were selected because they are representative of glycans derived from post-translationally modified proteins which, like these compounds, often consist of isomers of equal mass and similar shape. Replicate spectral measurements were recorded and processed using a partial-least-squares (PLS) classification and quantification algorithms with a leave-one-batch-out (LOBO) training and testing procedure. Spectra derived from solutions of individual sugars were identified with 100% accuracy, and mixtures of the two sugars were quantified with an average error of 2.7% in the relative maltotetraose/stachyose composition for mixtures with a total oligosaccharide concentration of 1 mM.

Keywords: Raman; SERS; PLS; Oligosaccharide; Chemometrics

1. Introduction

Carbohydrates comprise one of the most widespread classes of organic molecules and play important roles in a variety of biological processes. One of the challenges in glycobiology is finding analytical tools for the structural elucidation of carbohydrate isomers with similar chemical and physical properties, such as glycans derived from post-translationally modified proteins. This work is focused on demonstrating the utility of surface-enhanced Raman scattering (SERS) for such difficult analysis applications and as a useful alternative to mass spectroscopy (MS), nuclear magnetic resonance (NMR) spectroscopy, and/or circular dichroism (CD).[1] The advantages of SERS include sensitivity to the structural differences between equal-mass isomers, lower detection limits than NMR spectroscopy, and higher chemical information content than CD (i.e., Raman spectra contain a larger number of narrower spectral features than CD spectra). Furthermore, SERS is more sensitive than normal Raman spectroscopy and is better suited to the analysis of aqueous solutions than mid-infrared (IR) vibrational spectroscopy (since IR spectra are obscured by strong water absorption, while Raman spectroscopy is relatively insensitive to water).

Previous studies have demonstrated that vibrational spectroscopy can be an effective tool for studying the molecular structures and interactions of carbohydrates.[1] Due to the strong sensitivity of vibrational spectra to molecular composition and structure, spectral 'fingerprints' for very structurally similar saccharides can be obtained. Zhbankov et al. have described the structural specificity of the vibrational spectra of carbohydrates with the same chemical composition but differing only in the spatial arrangement of individual fragments. They have attributed this specificity to the redistribution of vibrational energy between the bands of the skeletal base and monomer normal mode vibrations of similar

* Corresponding author. Tel.: +1-765-494-5256; fax: +1-765-494-0239; e-mail: dor@purdue.edu

0008-6215/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.09.019

frequency, thus leading to drastic changes in the band intensities in infrared and Raman spectra.[2] The unique spectral fingerprints reported for mono- and disaccharides using far infrared,[3] Raman,[4] Raman optical activity (ROA),[5] and SERS[6] suggest that such methods may be useful for the identification and characterization of biologically relevant oligosaccharides.

The present report demonstrates the use of SERS, coupled with multivariate analysis, for the classification and quantification of two isomeric oligosaccharides, maltotetraose and stachyose, and their mixtures. These compounds were chosen to represent glycans derived from post-translationally modified proteins.[7] Since they have identical molecular weights and very similar shapes, they would be difficult or impossible to identify using mass spectroscopic and/or chromatographic techniques, which are the current methods of choice in glycoprotein analysis. Under the conditions used in this work the observed spectra are attributed to a combination of SERS and normal Raman scattering with a nominal enhancement of 100–1000.

Partial-least-squares (PLS) classification is performed using normalized second-derivative spectra derived from solutions of each oligosaccharide. The percent composition of mixtures of the two oligosaccharides is determined using PLS calibration. These results indicate that such methods may indeed be valuable both for identification of individual oligosaccharides as well as for the quantification of oligosaccharide mixtures.

These applications are relevant to glycoprotein analysis, both because of the nature of the compounds studied and because of the improved detection limits that we have demonstrated. The concentration (1 mM) and total quantities (<10 μg) analyzed are each more than 10 times lower than those used in previous Raman studies of saccharide solutions.[1b] There is also significant room for additional improvement since less than 1/1000 of the deposited saccharide sample is actually probed by the Raman excitation laser. Thus improved microdeposition methods should facilitate detection of submicrogram quantities of saccharides. Furthermore, the present results were obtained using 100 times lower laser power and >10 times shorter signal collection times, as compared to the conditions used in previous normal Raman measurement of saccharide solutions.[1b]

2. Experimental

Details of the micro-Raman system employed in this work have been previously reported.[6b] Briefly, the Raman excitation, from a Spectra Physics Stabilite model 2017 Kr^+ laser operated at 647.1 nm, was fiber-optically coupled to a custom-built Raman Microprobe System (SpectraCode, West Lafayette, IN). Scattered light was collected by the microprobe and directed through a fiber bundle coupled to an Acton SP 300i monochromator, equipped with a Princeton Instruments Spec-100 400BR digital CCD camera.

Two oligosaccharides, maltotetraose (α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc) and stachyose hydrate (α-D-Gal-(1→6)-α-D-Gal-(1→6)-α-D-Glc-(1→2)-β-D-Fru), were provided by Dr. Milos Novotny of Indiana University. Aqueous oligosaccharide solutions were prepared using ultrapure water from a MilliQ Plus system (Millipore).

The silver electrode substrates were of rotating-disk construction, consisting of 4-mm diameter disks sheathed in Teflon. After polishing with 1.0- and 0.3-μm alumina, they were electrochemically roughened to engender stable SERS activity by means of five potential steps from −0.4 to 0.4 V versus a saturated calomel electrode (SCE) and returned in 0.1 M KCl as outlined previously.[6,8] The surface was then rinsed thoroughly with Millipore water and dried in a stream of N_2 gas. A 10-μL aliquot of the oligosaccharide solution (1 mM) was deposited onto the roughened silver substrate with a micropipette, and the solvent was allowed to evaporate. Raman spectra were acquired with ca. 15 mW power focused to a 120 μm diameter spot size on the sample, and spectral acquisition times were typically 120 s, unless otherwise noted. The ~4-mm diameter region over which the solution is deposited is much larger than the area nondestructively probed by the Raman excitation laser. Thus, less than 10 pmol of saccharide produced a high-quality spectrum, although a total of 10 nmol was deposited.

Under these experimental conditions, the observed spectra are believed to result from a combination of SERS and normal Raman scattering. Comparisons with Raman measurements performed using similar saccharide deposits on nominally flat silver and gold foils, as well as glass substrates, imply an enhancement of the order of 100–1000 for the oligosaccharide Raman cross-section when deposited on roughened silver electrode substrates. The estimated surface density of the saccharides deposited on roughened silver electrode substrates is ~800 pmol mm^{-2} (which represents an upper bound to the true surface density, since this estimate was obtained without correcting for the increased surface area induced by surface roughness).

3. Results and discussion

Figure 1 compares the normal Raman spectra of maltotetraose and stachyose obtained from 1 M aqueous solutions (a and c, respectively) and the corresponding spectra obtained following deposition of 10 μL of 1 mM solution onto a roughened silver substrate (b and d, respectively). The wavenumber range shown between 1600 and 300 cm^{-1} contains features repre-

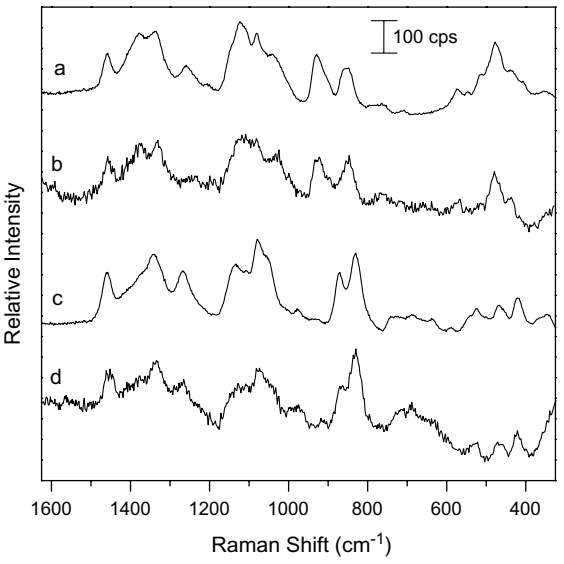

Figure 1. (a) Raman spectrum of 1 M maltotetraose solution; 75 mW laser power; 60 s integration time. (b) Raman spectrum of 10 μL of 1 mM maltotetraose solution deposited onto a roughened Ag electrode substrate; 15 mW power; 120 s integration time. (c) Same as (a) but for stachyose. (d) Same as (b) but for stachyose.

senting most of the vibrational normal modes, including HCH and CH_2OH vibrations (1500–1200 cm^{-1}), C–O stretching vibrations (1200–950 cm^{-1}), side-group deformations (950–700 cm^{-1}), and skeletal deformations (below 700 cm^{-1}).[1] The aqueous Raman spectrum of maltotetraose corresponds well to data in a previous report;[9] however, no prior publication of the Raman spectrum of stachyose has been found. The lower signal-to-noise ratio of the Raman spectra of the deposited samples as compared to that of the aqueous solutions is undoubtedly due to the lower concentration and laser power used, as well as the increased luminescence background typically observed with SERS substrates.

Comparison of the spectra in Figure 1 reveals that the Raman spectra of the deposited oligosaccharide samples correspond very closely with the normal Raman spectra of the solutions. The similarity of normal Raman and SERS spectra for monosaccharides has been previously noted.[10] Although the solvent is allowed to evaporate from the sample before spectral measurement, this correlation suggests that the deposited oligosaccharide molecules are still significantly solvated. Most noteworthy, however, is that the spectra of maltotetraose and stachyose differ substantially, especially in the regions 950–800 and 600–400 cm^{-1}, even though the two oligosaccharides are of identical mass. The highly coupled nature of saccharide vibrations results in a marked sensitivity of the Raman spectra to small structural differences, thereby enabling the spectra of maltotetraose and stachyose to be readily distinguished.

The reproducibility of the Raman spectra was investigated with repeated measurements performed on independently prepared roughened silver substrates. For each oligosaccharide, 30 spectral measurements were made at different positions on five independently prepared Ag substrates. Figure 2A and B shows representative spectra (out of over 300 measurements, total) for each oligosaccharide collected using a different substrate. The normal Raman spectra of each oligosaccharide obtained from 1 M solutions are also shown (bottom traces). The good reproducibility of the Raman spectral signatures for the two oligosaccharides suggests that such spectra can be utilized for automated spectral classification purposes using multivariate analysis to correlate observed spectral changes with properties such as analyte identity or concentration.

In order to minimize any spectral variance due to background or noise, the spectral data were first pre-processed using a 15-pixel Savitsky–Golay second derivative (SGSD) algorithm and then normalized (so $\sum_i I_i^2 = 1$) to produce normalized second derivative (NSD) spectra.[10b] Such pre-processing is advantageous as it serves to enhance structurally significant Raman features and suppress spurious background and noise interference. Representative NSD spectra for maltotetraose and stachyose are shown in Figure 2C and D, respectively.

The NSD spectra of the two oligosaccharides (with 150 spectra of each oligosaccharide, collected from five different substrates) were then processed using a PLS[10]

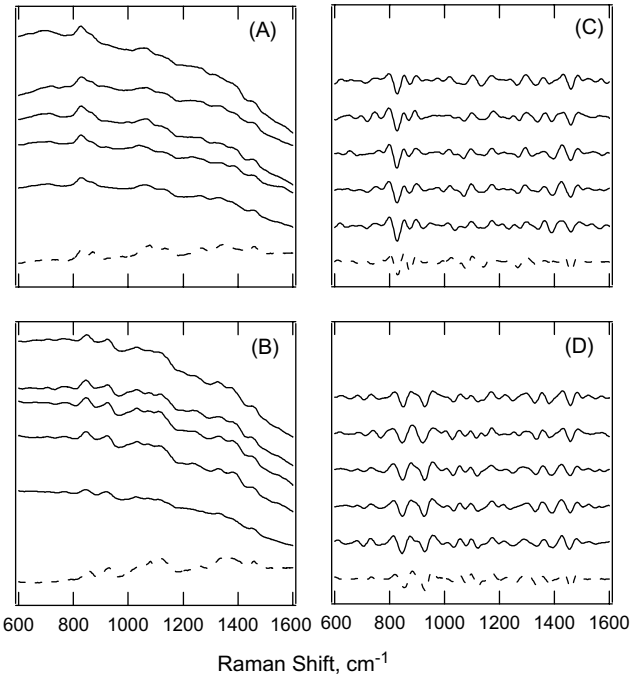

Figure 2. Reproducibility of Raman spectra of 1 mM maltotetraose (A) and stachyose (B) solutions deposited onto roughened Ag electrode substrates, and the corresponding NSD pre-processed spectra for maltotetraose (C) and stachyose (D) (see text). The lower spectra (dashed curves) are those obtained from the corresponding 1 M aqueous solution. Spectra are shifted vertically for clarity.

classification algorithm to produce an optimized procedure for identifying these oligosaccharides from the spectra. In order to insure that the procedure is reliable and generally applicable, the leave-one-batch-out (LOBO) method[11] was used for training and testing. This involves training using spectra from all but one of the substrate batches and then testing using spectra from the excluded-substrate batch. This cross-validation procedure is repeated for all the substrate batch tests (in each case training using all other batches).

The resulting PLS classification score plot is shown in Figure 3. Each point in this plot represents an entire Raman spectrum, and the three clouds of points represent the background (circles), maltotetraose (squares), and stachyose (triangles). The good separation between the clouds of points indicates that the spectra are easily distinguishable and thus that the two oligosaccharides may be readily identified from their corresponding spectra with 100% accuracy. No mis-classifications were found in testing a total of 300 spectra.

An even more challenging task is that of determining whether such methods can be used to quantify the percent composition in a mixture of the two oligosaccharides. In order to demonstrate this, a total of nearly 360 spectra were collected from mixtures of the above two oligosaccharides, each with a total concentration of 1 mM and relative maltotetraose/stachyose compositions of 0/100%, 25/75%, 50/50%, 75/25%, and 100/0%. Again, only 10 μL of each solution was used for spectral measurement, and spectra from independent substrates were used for training and testing. The corresponding NSD spectra were utilized in the PLS calibration, and a LOBO cross-validation procedure was again performed to evaluate the quality of the method.

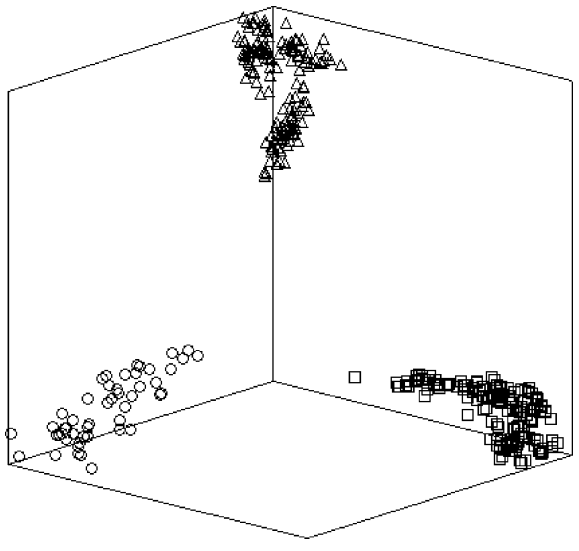

Figure 3. PLS classification score plot showing good differentiation of oligosaccharide Raman spectra obtained by depositing 1 mM solutions onto roughened Ag electrode surfaces: background, maltotetraose, and stachyose spectra are represented as circles, squares, and triangles, respectively.

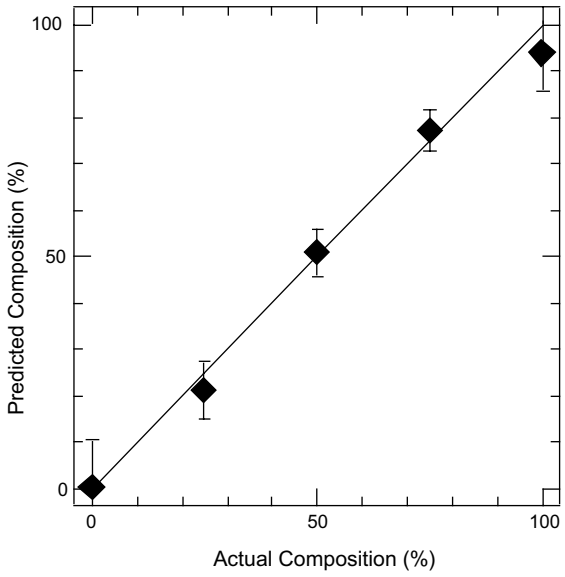

Figure 4. Plot of PLS predicted oligosaccharide mixture percent composition versus actual percent composition using LOBO cross-validation. Each point represents the average of between 18 and 120 independent Raman measurements. Error bars represent 90% confidence intervals.

Figure 4 shows the resulting cross-validated oligosaccharide mixture percent composition predictions. Each point in Figure 4 represents the average of between 18 and 120 measurements. The error bars represent 90% confidence intervals, and the average error of 2.7% is determined from the difference between the average predicted composition and the actual composition. Note that although proper training of the PLS algorithm requires extensive statistical sampling of Raman spectra, once this training is complete, only a single Raman spectrum is required in order to determine the composition of the oligosaccharide mixture. Thus such methods may be used to readily determine the compositions of oligosaccharide mixtures.

4. Concluding remarks

The foregoing demonstrates the utility of Raman spectroscopy and chemometric analysis for the differentiation of two isomeric oligosaccharides, as well as the determination of percent composition of oligosaccharide mixtures. We have also used similar methods to quantify the percent composition of tertiary monosaccharide mixtures, to produce preliminary results of only moderately lower accuracy than shown in Figure 4.[12] Since carbohydrates are difficult to detect using SERS due to their modest Raman cross-sections and weak (or

no) surface adsorption, alternative methods such as preconcentrating the oligosaccharide within the zone of the electromagnetic field enhancement using alkanethiol monolayers may be useful for enhanced detection. Such tactics have been successfully employed for the quantitative detection of glucose using SERS.[13]

Following this work, we have found that high-quality Raman spectra of millimolar concentrations of aqueous mono- and oligosaccharides can also be obtained by drop-coating deposition onto nominally flat gold or silver foil substrates, when the sample is concentrated to a small ($\sim$0.3 mm^2) deposition area following solvent evaporation.[14] Although little to no surface-enhancement is obtained with nominally flat gold or silver foil substrates, the low substrate background facilitates the detection of comparably small analyte quantities without requiring a roughening pre-treatment.[14] The substrates and deposition conditions used in the present study produce spectra that likely derive from a combination of SERS and normal Raman scattering (as discussed in Section 2). The resulting spectra are moderately enhanced and highly reproducible, with little or no variation in intensity or spectral shape, either within the deposition area or from one deposition to another.

Acknowledgements

This work was supported by the Indiana Proteomics Consortium. The authors thank Prof. Milos Novotny and Dr. Yehia Mechref of Indiana University for valuable discussions and for providing the glycan samples used in this analysis.

References

1. (a) Matlouthi, M.; Koenig, J. L. *Adv. Carbohydr. Chem. Biochem.* **1986**, *44*, 7–89; (b) Arboleda, P. H.; Loppnow, G. R. *Anal. Chem.* **2000**, *72*, 2093–2098, and references cited therein.
2. (a) Zhbankov, R. G.; Firsov, S. P.; Grinshpan, D. D.; Baran, J.; Marchewka, M. K.; Ratajczak, H. *J. Mol. Struct.* **2003**, *645*, 9–16; (b) Zhbankov, R. G.; Andrianov, V. M.; Marchewka, M. K. *J. Mol. Struct.* **1997**, *436/437*, 637–654.
3. Yang, L.; Weng, S.; Ferraro, J. R.; Wu, J. *Vib. Spectrosc.* **2001**, *25*, 57–62.
4. Arboleda, P. H.; Loppnow, G. R. *Anal. Chem.* **2000**, *72*, 2093–2098.
5. (a) Bell, A. F.; Hecht, L.; Barron, L. D. *J. Mol. Struct.* **1995**, *349*, 401–404; (b) Wen, Z. Q.; Barron, L. D.; Hecht, L. *J. Am. Chem. Soc.* **1993**, *115*, 285–292; (c) Barron, L. D.; Gargaro, A. R.; Wen, Z. Q. *Carbohydr. Res.* **1991**, *210*, 39–49.
6. (a) Mrozek, M. F.; Weaver, M. J. *Anal. Chem.* **2002**, *74*, 4069–4075; (b) Mrozek, M. F.; Weaver, M. J. *J. Phys. Chem. B* **2001**, *105*, 8931–8937.
7. (a) Mechref, Y.; Novotny, M. V. *Anal. Chem.* **1998**, *70*, 455–463; (b) Wuang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem.* **2001**, *73*, 6063–6069.
8. Roth, E.; Hope, G. A.; Schweinsberg, D. P.; Keifer, W.; Fredericks, P. M. *Appl. Spectrosc.* **1993**, *47*, 1794–1800.
9. Sekkal, M.; Dincq, V.; Legrand, P.; Huvenne, J. P. *J. Mol. Struct.* **1995**, *349*, 349–352.
10. (a) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193–1202; (b) Zhang, D.; Ben-Amotz, D. *Appl. Spectrosc.* **2000**, *54*, 1379–1383.
11. Martens, J.; Naes, T. *Multivariate Calibration*; Wiley: Chichester, UK, 1989.
12. Mrozek, M. F.; Zhang, D.; Ben-Amotz, D. Unpublished results.
13. Shafer-Peltier, K. E.; Haynes, C. L.; Glucksberg, M. R.; Van Duyne, R. P. *J. Am. Chem. Soc.* **2003**, *125*, 588–593.
14. Zhang, D.; Xie, Y.; Mrozek, M. F.; Ortiz, C.; Ben-Amotz, D. *Anal. Chem.*, in press.