Automatically Full Glycan Structural Determination with Logically Derived Sequence Tandem Mass Spectrometry

Shang-Ting Tsai,[a] Chia Yen Liew,[a,b] Hsu Chen Hsu,[a] Shih-Pei Huang,[a] Wei-Chien Weng,[a] Yu-Hsiang Kuo[a] and Chi-Kung Ni,*[ac]

a Institute of Atomic and Molecular Sciences, Academia Sinica, P.O. Box: 23-166, taipei, 10617, Taiwan.

b Department of Chemistry, National Taiwan Normal University, Taipei, 11677, Taiwan

c Department of Chemistry, National Tsing Hua University, Hsinchu, 30013, Taiwan.

Glycans have diverse functions and play vital roles in many biological systems such as influenza, vaccines, and cancer biomarkers. However, fully structural identification of glycans remains challenging. Glycan structure was conventionally determined by chemical methods or nuclear magnetic resonance spectroscopy which require a large amount of sample and is not readily applicable for glycans extracted from biological samples. Although it has a high sensitivity and widely used for structural determination of molecules, current mass spectrometry can only reveal parts of the glycan structure. Here we show the full structures of glycans, including diastereomers and anomericity of each monosaccharide and linkage position of each glycosidic bond, can be determined using tandem mass spectrometry guided by a logically derived sequence (LODES). This new method provides de novo oligosaccharide structural identifications with high sensitivity and was applied to automatically in situ structural determination of the oligosaccharides eluted by high performance liquid chromatography. We showed that the structure of a given trisaccharide from trisaccharide mixture and bovine milk were determined from near three thousand isomers by using six to seven logically selected CID spectra. The entire procedure of mass spectrum measurement guided by LODES can be programmed in computer for automatically full glycan structural identification, a goal that remains a great challenge in glycan analysis.

Introduction

The four basic categories of molecules involved in creating life are nucleic acids, proteins, glycans (carbohydrates), and lipids. Among these four categories, lipids have relatively simple structures that can be determined easily. Nucleic acids, proteins, and carbohydrates are macromolecules comprising of nucleotides, amino acids, and monosaccharides, respectively. These three categories of molecules have complex structures due to variations in the sequences and the subunits. Although methods used for elucidating the sequences and residue constituents of nucleic acids and proteins are available, there is yet a strategy for determining the structure of carbohydrates.[1]

Glycans have been demonstrated to play important roles in many biological processes such as protein folding, cell-cell interactions, cell targeting and cell differentiation.[2-4] They act as binding and recognition sites for a variety of viruses and bacteria,[5] and are used in medicine such as carbohydrate-based vaccines[6] and anticoagulant.[7] Hence obtaining their complete structural information is necessary to

enable our thorough understanding glycan chemical and biological properties. However, there is no simple method with high sensitivity that can fully determine the structures of glycans.

Traditional analysis of oligosaccharides involves reduction and permethylation, followed by acid hydrolysis, borodeuteride reduction, acetylation, and then the identification of linkage positions and monosaccharide composition by a combination of GC and MS.[8] This procedure only provides linkage positions of the glycosidic bonds. Sequence of the linkages and anomeric configurations of the glycosidic bonds have to be determined by the other methods, e.g., enzyme digestion.[9] Many techniques, such as nuclear magnetic resonance (NMR) spectroscopy,[10-12] mass spectrometry (MS),[13-16] infrared spectroscopy,[17-19] microwave spectroscopy,[20,21] have been developed and used to reveal carbohydrate structures recently. High sensitivity is crucial when the analytical technique is applied to biological samples, from which only minuscule amounts of glycans can be extracted. MS provides structural information at a sensitivity three to four orders of magnitude higher than the other methods, and this method is particularly useful when applied to samples of limited supply. In structural analysis by MS, accurate mass measurement from single-stage MS only elucidates oligosaccharide composition. Combination of MS with various methods, such as derivatization of carbohydrates,[22,23] fragmentation (e.g., collision-induced dissociation (CID),[24-27] higher-energy collision dissociation,[28,29] infrared multiphoton dissociation,[30,31] electronic excitation dissociation,[32] electron capture disso-ciation,[33] and electron transfer dissociation,[34] post source decay of matrix-assisted laser desorption ionization),[35] ion mobility,[36] enzyme digestion,[37] or possible structural constrains of biological system (e.g., oligosaccharides in bovine milk usually have lactose at the reducing end) is required for structural determination. However, complete characterization of the glycan structure by MS remains difficult because of the numerous possible isomers of a given chemical formula[38] and the subtle differences between these isomers. The aforementioned methods either reveal only parts of the structures, or are limited to the few well-characterized oligosaccharides in MS libraries, or are time and sample consuming for completely structural determination.

In this study, a new MS method for the structural determination of oligosaccharides was demonstrated. Full structure, including linkage positions, anomeric configurations, and constituent of each unit monosaccharide can be determined. This method was applied to the in situ structural determination of the oligosaccharides eluted from high performance liquid chromatography. We showed that the structure of a given trisaccharide from trisaccharide mixture and bovine milk were determined from near three thousand isomers by using six to seven appropriately selected CID spectra. The structural determination procedure involves the dissociation of oligosaccharides into monosaccharides or disaccharides in an ion trap of mass spectrometer and the in situ structural identification of these monosaccharides and disaccharides by subsequent MS in the same ion trap. The complete structures of oligosaccharides are reconstructed from the structures of disaccharides and monosaccharides. This concept is simple, but the success of this method requires three key technologies. One is a method by

which the diastereomers and anomericity of each monosaccharide can be identified. We found the CID spectra of monosaccharide lithium adducts exhibit significant differences between α-glucose, β-glucose, α-galactose, β-galactose, α-mannose, β-mannose, N-acetylgalactosamine (GalNAc), and N-acetylglucosamine (GlcNAc). These monosaccharides can be differentiated by the CID spectra. The second crucial technology is the capability to differentiate the linkage position and anomeric configurations of glycosidic bonds connecting these monosaccharides. Studies have reported that the fragments produced from cross ring dissociation can be used to determine the linkage positions of glucose disaccharide lithium adducts.[39,40] In this study, we extended the use of cross ring fragments to identify the linkage positions and anomeric configurations of disaccharides containing glucose, galactose, mannose, GalNAc, and GlcNAc based on the understanding of dissociation mechanism.[41-43] The third one is the capability to dissociate selective chemical bond and generate the chosen monosaccharide or disaccharide for subsequent structural determination. We have developed a procedure, namely logically derived sequence (LODES) for tandem MS, to generate the desired monosaccharide or disaccharide by collision induced dissociation of oligosaccharide sodium adducts.[44,45] We extended the method of LODES to oligosaccharide lithium adducts in this study. These three key technologies are described in the subsequent paragraphs, followed by the applications of the new MS method to structural determination of oligosaccharides.

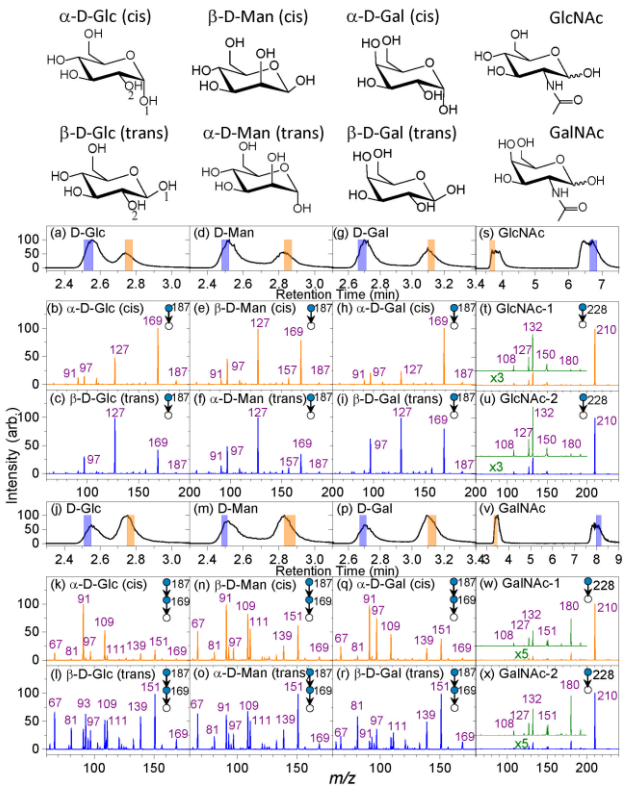


Fig. 1. HPLC chromatograms (selected-ion monitoring (SIM) of *m/z* 187 for (a), (d), and (g); *m/z* 169 for (j), (m), and (p); *m/z* 228 for (s) and (v)) and CID spectra of monosaccharide hexose lithium adducts (*m/z* 187) and HexNAc lithium adducts (*m/z* 228). Since the α- and β-anomers of each monosaccharide typically coexist in solution, we separated these two anomers by HPLC and measured the CID spectra of the separated anomers immediately after separation. The blue and orange bars surrounding each peak in the chromatograms represent the period in which the CID spectra were measured. The corresponding spectra are shown in blue and orange, respectively.

Results and discussion

Identifications of monosaccharides

Fig. 1 shows the CID spectra of various monosaccharide lithium adducts. A major difference between the CID spectra of two anomers of a given hexose (e.g., Fig. 1(b) and 1(c) for α-glucose and β-glucose) is the relative intensity of ion *m/z* 169, which represents H_2O elimination. High-level quantum chemistry calculations have revealed that the water elimination barriers of the O1 and O2 atoms in the cis configuration of the sodium adduct was substantially smaller than that in the trans configuration.[41,42] We found lithium adducts have the similar properties. Therefore, the CID spectrum with high ion *m/z* 169 intensity, representing a high branching ratio in the dehydration reaction, can be assigned to the anomer that has the O1 and O2 atoms in the cis configuration (α-glucose, α-galactose, and β-mannose). Fig. 1 also shows that the CID spectra of the hexose dehydration products, 187→169→ fragments, are very different between these hexoses. For example, ion *m/z* 91 and 151 have the highest intensity for cis and trans configuration, respectively, and the relative intensities of ions *m/z* 81, 93, 97, 109, 111, 139, and 151 are very different between these hexoses. These CID spectra can be used to identify the constitution and anomeric configuration of hexoses. By contrast, CID spectra of GlcNAc shows significant difference from that of GalNAc, but no difference between α- and β-GlcNAc or α- and β-GalNAc was discovered. We could only distinguish GlcNAc and GalNAc but not the anomeric configuration of these HexNAc. The anomeric configurations of these HexNAc were determined using the disaccharide described in the next section. Details on spectrum identification and differentiation achieved through spectral similarity calculations are described in Supplementary Information. Additional CID spectra of ^{18}O labelled monosaccharides were measured and showed similar spectra.

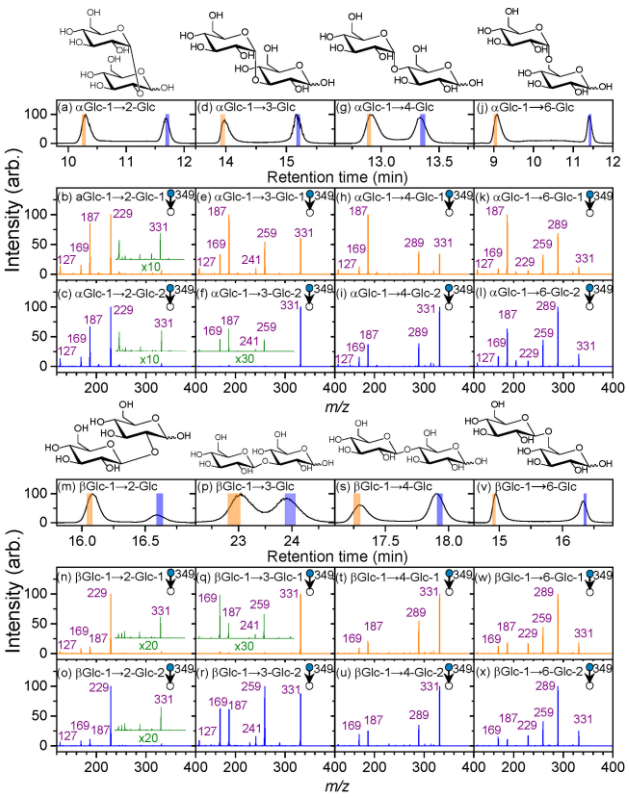


Fig. 2. HPLC chromatograms (SIM of *m/z* 349) and CID spectra of glucose disaccharide lithium adducts (*m/z* 349). The CID spectrum of each anomer was measured immediately after the separation of two anomers by HPLC. The orange and blue bars surrounding each peak in the chromatograms represent the period in which the CID spectra were measured. The corresponding spectra are displayed in orange and blue, respectively.

Identifications of linkages

Linkage identifications include the determination of linkage positions and anomeric configuration of glycosidic bond. The fragmentation patterns of different linkage positions for glucose disaccharide sodium or lithium adducts have been reported in previous studies.[39,40] The major dissociation patterns include the loss of neutral fragments with m = 120 u (and m = 18 u for mannose or and some galactose disaccharides) for the 1→2 linkage, m = 18 and 90 u for the 1→3 linkage, m = 18 and 60 u for the 1→4 linkage, and m = 18, 60, 90, and 120 u for the 1→6 linkage. These fragmentation patterns were explained by the retro-aldol reaction.[39] Recent high-level quantum chemistry calculations[41,42] confirmed that retro-aldol reaction starting from the O1 atom of the sugar at the reducing end is the dominant mechanism of cross-ring dissociation. This mechanism does not depend on the orientation of hydroxyl groups or the sugar at the non-reducing end. Consequently, we can extend this mechanism to explain the similar fragmentation patterns of hexose disaccharides (Hex-Hex) and HexNAc-Hex disaccharides

with hexoses at the reducing end, (e.g., GalNAc-Gal). These fragmentation patterns were used to determine the linkage positions.

The α and β anomeric configurations of the sugar at the reducing end of disaccharides were not separated prior to CID spectrum measurement in the aforementioned studies.[39,40] In the present study, the CID spectrum of each disaccharide anomer was measured immediately after the separation of anomers by HPLC. Our previous study[44,45] on hexose disaccharide sodium adducts demonstrated that both the anomeric configurations of the glycosidic bond and the sugar at the reducing end can be determined using the relative intensities of glycosidic bond cleavage, dehydration, and cross-ring dissociation. We ascertained that a similar concept can be applied to hexose disaccharide lithium adducts. However, the differences of the relative ion intensities between anomers for hexose disaccharide lithium adducts are smaller than those of sodium adducts. Alternatively, we used the fragmentation patterns of hexose disaccharides to determine linkage positions in this study, but anomeric configurations were determined by the identification of monosaccharide anomericity. For HexNAc-Hex disaccharide, the spectrum difference for disaccharide lithium adducts with different anomeric configurations of the glycosidic bond was clear, and the disaccharide database was used to determine the anomeric configuration. Spectrum identification and differentiation achieved through spectral similarity calculations.

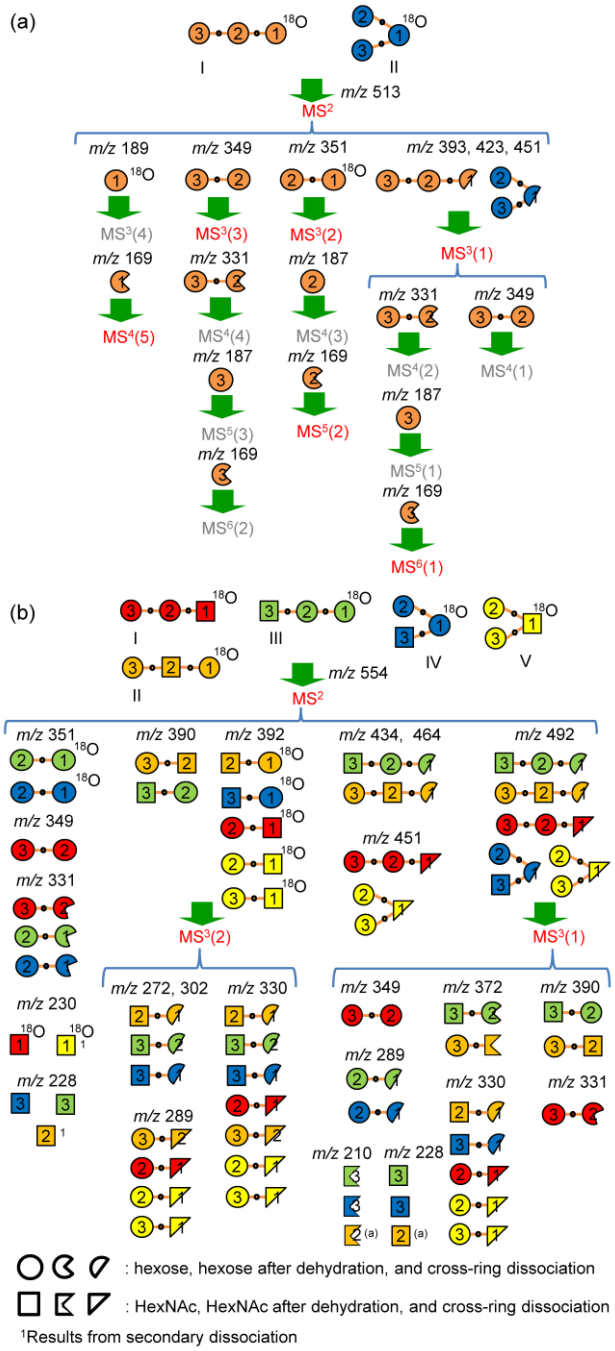


Fig. 3. LODES for structural determination of trisaccharides. (a), trisaccharides containing three hexoses (only the fragments used to differentiate types I and II, and the fragments for type I structural determination are shown); (b), trisaccharides containing two hexoses and one HexNAc (only the fragments to differentiate type I-V are shown). Reducing end is located at the monosaccharide unit 1. The colors used in this figure are simply to distinguish different types of trisaccharides, they are not related to the color codes used in the symbol nomenclature of Consortium for Functional Glycomics (CFG).

Logically derived sequences (LODES) for MS^n

A logically derived sequence (LODES) for CID tandem mass spectrometry was developed based on the understanding of the dissociation mechanisms of carbohydrate sodium adducts[44,45]. This sequence can be used for generating the chosen monosaccharides and disaccharides for subsequent structural determination. It can also determine the branch locations and the linkage positions at the branch locations. The same LODES can be applied to lithium adducts because of the similar dissociation mechanisms. However, in the process of LODES, the first step is to generate the fragment by dehydration or cross ring dissociation at the reducing end such that B (or C) ions can be differentiated from Z (or Y) ions in subsequent dissociation. Our previous study[46] showed that dehydration and cross-ring dissociation from the non-reducing end of oligosaccharide lithium adducts were not negligible compared to sodium adducts; therefore, labelling of O1 atom of the sugar at the reducing end by ^{18}O is necessary for some oligosaccharides when lithium adducts are used.

The LODES employed for the structural determination is illustrated in Fig. 3; lithium adducts of trisaccharides containing three hexoses or two hexoses and one HexNAc are used as examples. A ^{18}O labelled trisaccharide containing three hexoses (*m/z* 513) can be linear or branched, denoted as type I and II, respectively, in Fig. 3(a). Linear trisaccharides undergo cross-ring dissociation at the sugar of reducing end in MS^2, producing fragment ions *m/z* 451, 423, or 393. In contrast, branched trisaccharides do not produce these fragments except a branched trisaccharide with (1→6, 1→4) linkages. Linear trisaccharides and the branched trisaccharide with (1→6, 1→4) linkages can be distinguished in $MS^3(1)$: only linear trisaccharides produce fragment ions *m/z* 331 and 349 from ions *m/z* 451, 423, or 393.

When the carbohydrate is identified as a linear trisaccharide, the following procedures are used to determine the detailed structure of the linear trisaccharide. The lithiated disaccharides produced in MS^2 (Fig. 3(a)) are the lithiated disaccharide ion *m/z* 351 at the reducing end (②-①, representing a disaccharide comprising monosaccharide units 2 and 1) and the lithiated disaccharide ion *m/z* 349 at the non-reducing end (③-②) of trisaccharide. The CID spectra of these disaccharides, $MS^3(2)$ and $MS^3(3)$, provide the information necessary to determine the linkage positions of the glycosidic bonds between units 2 and 1 and between units 3 and 2, respectively. Diastereomers and the anomeric configurations of the monosaccharides are identified using the CID spectra of monosaccharide product ion *m/z* 169, $MS^5(2)$, $MS^6(1)$, and $MS^4(5)$ (Fig. 3(a)). Other CID sequences can be used to crosscheck the structure. For example, the linkage between monosaccharide units 2 and 1 can be verified using MS^2, and $MS^6(2)$ can be employed to confirm the structure of monosaccharide unit 3. A similar approach can be applied to trisaccharides containing HexNAc, as illustrated in Fig. 3(b).

For practical applications, the CID spectra of these monosaccharide and disaccharide lithium adducts were obtained in advance and recorded in a database. The CID spectra of the monosaccharide and disaccharide lithium adducts produced from the dissociation of oligosaccharides were compared with those in the database for structural identification. For the structural determination of the sugar at the reducing end, which is labelled by ^{18}O, a separate database containing ^{18}O was used.

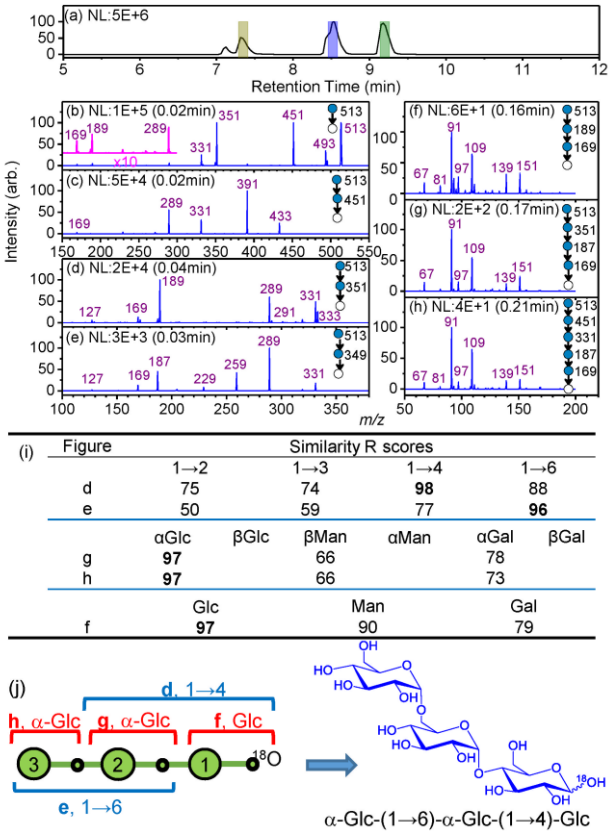


Figure 4. In situ structural determination of trisaccharides (^{18}O labeled) appears in chromatogram. These CID spectra were measured when the peak in the chromatogram (a) (SIM of m/z 513) highlighted in blue appeared. The first, second, third and fourth injections of sample were used to obtain CID spectra of (b)-(e), (f), (g) and (h), respectively. The structure was assigned correctly according to the highest values of R scores shown in (i), and the details are shown in (j). The partial splitting of the peak highlighted in blue in (a) is due to different anomeric configuration of the sugar at the reducing end.

Oligosaccharide structural determination

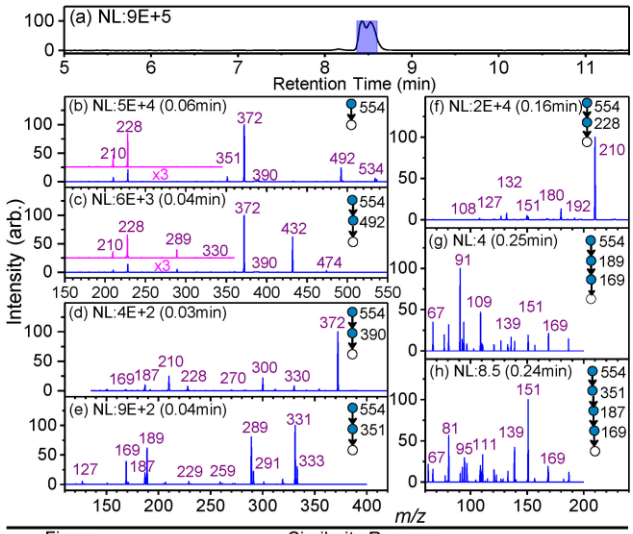
The LODES for tandem MS of hexose trisaccharides, as illustrated in Fig. 3(a), was programmed into the mass spectrometer for in situ structural determination of the trisaccharides eluted from HPLC. The mixture of three trisaccharides, ^{18}O-labelled α-Glc-(1→6)-α-Glc-(1→4)-Glc, ^{18}O-labelled α-Gal-(1→

3)-β-Gal-(1→4)-Glc, and ^{18}O-labelled α-Man-(1→2)-α-Man-(1→6)-Man, was injected into a HPLC coupled with ESI-MS. Fig. 4 shows the chromatograms and CID mass spectra of these saccharides. Here we use the second oligosaccharide eluted from HPLC, highlighted in blue in Fig. 4(a), as an example to demonstrate how a structure can be determined. Precursor ion *m/z* 513 suggests that it is a lithium adduct of ^{18}O labelled trisaccharide containing three hexoses. Ion *m/z* 331 observed in Fig. 4(c) suggests it is a linear trisaccharide, according to the MS^3(1) of LODES in Fig. 3(a). Fig. 4(f), (g), and (h) show the CID spectra of monosaccharides at the reducing end, center, and non-reducing end of the trisaccharide, respectively. Comparison of these spectra with the monosaccharide database suggests that they are glucose, α-glucose, and α-galactose, respectively. Large ion intensity of I(*m/z* 451) in MS^2 (Fig. 4(b))—suggest that it is the fragmentation pattern of the glycosidic bond at the reducing end with (1→4) linkage. Ions *m/z* 351 and 349 represent the disaccharide at the reducing end (a disaccharide comprising monosaccharide units 2 and 1 of type I in Fig. 3(a)) and the disaccharide at the non-reducing end (a disaccharide comprising monosaccharide units 3 and 2 of type I in Fig. 3(a)), respectively. The ion intensities I(*m/z* 289) > I(*m/z* 259) >I (*m/z* 229) in CID of *m/z* 349 (Fig. 4(e)) confirm that the glycosidic bond at the non-reducing end is a (1→6) linkage. Large intensity of ion *m/z* 289 in CID of *m/z* 351 (Fig. 4(d)) indicates that the glycosidic bond at the reducing end is a (1→4) linkage. Combining these spectra reveals that the structure of this trisaccharide is α-Glc-(1→6)-α-Glc-(1→4)-Glc. Spectral similarity calculations between the CID spectra in Fig. 4 and CID spectra in database used in structural assignments are shown in Fig. 4(i), and the structure determined from highest values of *R* scores in Fig. 4(i) is shown in Fig. 4(j). Similar procedures were used to determine the structures of oligosaccharides in the other peaks of chromatogram.

This method is rapid and sensitive, which is suitable for limited amount of oligosaccharides extracted from biological sample. Fig. 5 shows the chromatograms and CID mass spectra of free oligosaccharides extracted from bovine milk. The CID spectra were measured when the peak in the chromatograms appeared. Precursor ion *m/z* 554 suggests that it is a lithium adduct of ^{18}O labelled trisaccharide containing two hexoses and one HexNAc.

According to the LODES of trisaccharides containing two hexoses and one HexNAc (Fig. 3(b)), fragment ion *m/z* 351 in CID spectrum of Fig. 5(b) suggests it is type III or IV trisaccharides, and fragment ion *m/z* 289 and 372 in CID spectrum 554→492→fragments (Fig. 5(c)) indicates the trisaccharide is type III. Fig. 5(f), (g), and (h) show the CID spectra of monosaccharides at the non-reducing end, reducing end, and center of the trisaccharide, respectively. Comparison of these spectra with the monosaccharide database suggests that they are GalNAc, Glc, and β-Gal, respectively. Ions *m/z* 390 and 351 represent the disaccharide lithium adducts at the non-reducing end and reducing end, respectively. They are identified as α-GalNAc-(1→3)-Gal, and ^{18}O labelled Gal-(1→4)-Glc by comparing Fig. 5(d) and (e) to the database of disaccharides. Consequently, we can determine the structure of the trisaccharide as α-GalNAc-(1→3)-β-Gal-(1→4)-Glc. Information Fig. S7). The same retention time and very similar CID spectra to that of

the trisaccharide in Fig. 5 were obtained. The similar method was used to determine the hexose trisaccharides extracted from bovine mile.



(i)

Figure	Similarity R scores			
d	α-GalNAc-(1→3)-β-Gal	α-GalNAc-(1→3)-α-Gal	β-GalNAc-(1→3)-α-Gal	β-GalNAc-(1→3)-β-Gal
	95	87	60	30
e	1→2	1→3	1→4	1→6
	60	62	**98**	87
f	GlcNAc1	GlcNAc2	GalNAc1	GalNAc2
	87	84	**97**[1]	**98**[1]
h	βGlc	αMan	βGal	
	61	59	**97**	
g	Glc	Man	Gal	
	94	87	81	

[1]Both of them were assigned to correct structure.

(j)

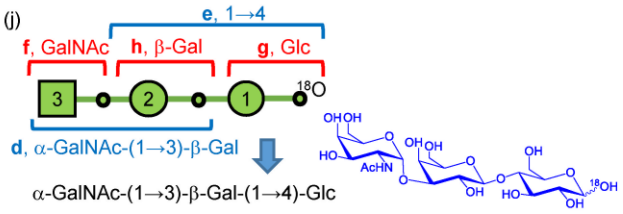


Fig. 5. In situ structural determination of free oligosaccharides (^{18}O labeled) from bovine milk. (a), HPLC chromatograms (SIM of *m/z* 554); (b)-(h), CID spectra were measured when the peak in the chromatograms highlighted in blue appeared. Spectra (b)-(e), (f), (g), and (h) were obtained in the first, second, third, and fourth injections of sample, respectively. (j), the structure concluded from CID spectra. The structures were assigned according to the highest values of *R* scores shown in (i), and the details are shown in (j). The partial splitting of the peak highlighted in blue in (a) is due to different anomeric configuration of the sugar at the reducing end.

Conclusions

The total number of isomers for hexose trisaccharides is 2376 by considering the possible monosaccharides (glucose, galactose, and mannose), linkages (1→2, 1→3, 1→4, and 1→6), branching location, and anomeric configurations (α and β) of all glycosidic bonds. The number of isomers increases to 3960 when one hexose is replaced by HexNAc (GlcNAc or GalNAc). In this study, we have

demonstrated that the complete structure of a given trisaccharide can be identified using CID tandem mass spectra guided by LODES. The necessary spectra for structural determination were all obtained during the time in which the trisaccharide was present in chromatograms of HPLC, although multiple sample injections were needed. In addition to the aforementioned oligosaccharides, we verified our method by another 13 trisaccharides. In the entire structural determination process, neither the assumption of possible oligosaccharide structure was made, nor structural constrains based on biological knowledge was used. It is a de novo method and the structural determination procedure is equivalent to blind test. This is the first time that the complete structures of oligosaccharides are determined using a simple MS approach.

Although more sample injections are expected when the oligosaccharides are larger than trisaccharides, the sensitivity improvement using nanospray ionization to increase the ionization efficiency and using a mass spectrometer with multiple ion traps (e.g., Orbitrap Fusion mass spectrometer, Thermo Fisher Scientific Inc.) to increase the sample usage duty cycle are two feasible methods for reducing the consumption of sample. Similar CID tandem mass spectra guided by LODES can be used to determine the structures of oligosaccharides containing other monosaccharides. This method is rapid and has high sensitivity. The entire procedure of mass spectrum measurement guided by LODES and spectrum comparison to database for structural assignment can be programmed in computer for automatically in situ structural identification, a goal that remains a great challenge in carbohydrate analysis.

References

1 National Research Council (US), *Transforming Glycoscience: A Roadmap for the Future*, The National Academies Press, Washington, DC, 2012.

2 I. Bucior and M. M. Burger, *Curr. Opin. Struct. Biol.,* 2004, **14**, 631.

3 R. Kleene and M. Schachner, *Nat. Rev. Neurosci.,* 2004, **5**, 195.

4 A. Varki, *Nature*, 2007, **446**, 1023.

5 A. Imberty and A. Varrot, A. *Curr. Opin. Struct. Biol*., 2008, **18**, 567.

6 C. J. Wei et al. *Sci. Transl. Med*., 2010, **2**, 24ra21.

7 C. H. Chang et al. *Angew. Chem Int. Ed. Engl*., 2014, **53**, 9876.

8 B. Mulloy, A. Dell, P. Stanley and J. H. Prestegard, Ch. 50 in *Essentials of Glycobiology. 3rd edition*. Edited by A. Varki et al. Cold Spring Harbor Laboratory Press, New York, USA, 2015-2017.

9 D. L. Aldredge et al. *Glycobiology*, 2013, **23**, 664.

10 J. O. Duus, C. H. Gotfredsen and K. Bock, *Chem. Rev*., 2000, **100**, 4589.

11 J. Jiménez-Barbero and T. Peters, *NMR Spectroscopy of Glycoconjugates*. Wiley-VCH, Weinheim, Germany, 2003.

12 M. D. Battistel, H. F. Azurmendi, B. Yu and D. I. Freedberg, *Prog. Nucl. Magn. Reson. Spectrosc*., 2014, **79**, 48.

13 A. Dell and H. R. Morris, *Science*, 2001, **291**, 2351.

14 J. Zaia, *Mass Spectrom. Rev*., 2004, **23**, 161.

15 Y. Park and C. B. Lebrilla, *Mass Spectrom. Rev*., 2005, **24**, 232.

16 M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla and I. J. Amster, *Anal. Chem*., 2014, **86**, 196.

17 E. J. Cocinero, P. Carcabal, T. D. Vaden, J. P. Simons and B. G. Davis, *Nature*, 2011, **469**, 76.

18 E. J. Cocinero, D. P. Gamblin, B. G. Davis and J. P. Simons, *J. Am. Chem. Soc*., 2009, **131**, 11117.

19 E. Mucha et al. *Angew. Chem. Int. Ed. Engl*. 2017, **56**, 11248.

20 E. J. Cocinero et al. *J. Am. Chem. Soc*., 2013, **135**, 2845.

21 J. L. Alonso et al. *Chem. Sci*., 2014, **5**, 515.

22 J. A. Saba, X. Shen, J. C. Jamieson and H. Perreault, *J. Mass Spectrom*., 2001, **36**, 563.

23 D. J. Harvey, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*., 2011, **879**, 1196.

24 D. J. Harvey, *J. Am. Soc. Mass. Spectrom*., 2005, **16**, 622.

25 D. Ashline, S. Singh, A. Hanneman and V. Reinhold, *Anal. Chem*., 2005, **77**, 6250.

26 C. E. Costello, J. M. Contado-Miller and J. F. Cipollo, *J. Am. Soc. Mass Spectrom*., 2007, **18**, 1799.

27 A. S. Palma et al. *Mol. Cell Proteom*., 2015, **14**, 974.

28 Z. M. Segu and Y. Mechref, *Rapid Commun. Mass Spectrom.,* 2010, **24**, 1217.

29 C. Singh, C. G. Zampronio, A. J. Creese and H. J. Cooper, *J. Proteome Res*., 2012, **11**, 4517.

30 B. Schindler et al. *Nat. Commun*., 2017, **8**, 973.

31 Y. Tan and N. C. Polfer, *J. Am. Soc. Mass Spectrom*., 2015, **26**, 359.

32 X. Yu, Y. Jiang, Y. Chen, Y. Huang, C. E. Costello and C. Lin, *Anal. Chem*., 2013, **85**, 10017.

33 H. J. Cooper, K. Hakansson and A. G. Marshall, *Mass Spectrom. Rev*., 2005, **24**, 201.

34 E. S. Witze, W. M. Old, K. A. Resing and N. G. Ahn, *Nat. Methods*, 2007, **4**, 798.

35 B. Guan and R. B. Cole, *J. Am. Soc. Mass Spectrom*., 2008, **19**, 1119.

36 P. Both et al. *Nat. Chem*., 2014, **6**, 65.

37 T. Song, D. Aldredge and C. B. Lebrilla, *Anal. Chem*., 2015, **87**, 7754.

38 R. A. Laine, *Glycobiology*, 1994, **4**, 759.

39 G. E. Hofmeister, Z. Zhou and J. A. Leary, *J. Am. Chem. Soc*., 1991, **113**, 5964.

40 C. S. Azenha, M. A. Coimbra, A. S. Moreira, P. Domingues and M. R. M. Domingues, *J. Mass Spectrom*., 2013, **48**, 548.

41 J. L. Chen et al. *Phys. Chem. Chem. Phys*., 2017, **19**, 15454.

42 H. T. Huynh et al. *Phys. Chem. Chem. Phys*., 2018, **20**, 19614.

43 C. C. Chiu et al. Phys. Chem. Chem. Phys. (submitted).

44 C. H. Hsu, C. Y. Liew, S. P. Huang, S. T. Tsai and C. K. Ni, *J. Am. Soc. Mass Spectrom*., 2018, **29**, 470.

45 C. H. Hsu, C. Y. Liew, S. P. Huang, S. T. Tsai and C. K. Ni, *Sci. Rep*., 2018, **8**, 5562.

S. T. Tsai, J. L. Chen and C. K. Ni, *Rapid Commun. Mass Spectrom*., 2017, **31**, 1835.