De novo determination of fucose linkages in *N*-glycans and the unusual *N*-glycans in bee venom

Jien-Lian Chen[a], Hsien-Wei Tseng[b], Chun-Cheng Lin[b], and Chi-Kung Ni*[ab]

[a] Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 10617, Taiwan

[b] Department of Chemistry, National Tsing Hua University, Hsinchu, 30013, Taiwan

*Corresponding author, e-mail addresses: ckni@po.iams.sinica.edu.tw

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Abstract

N-linked glycosylation are one of the most important post-translational modification of proteins. Fucosylated *N*-glycans are related to many biological processes and they are commonly used as biomarkers. However, determining fucose linkages in *N*-glycans remains challenging, and most of the fucose linkages, particularly the linkages in the antenna of *N*-glycans remain unidentified. In this work, a simple multi-stage tandem mass spectrometry method based on dissociation mechanisms was developed for the de novo determination of fucose linkages in *N*-glycans. Using collision induced dissociation (CID), fucose linkages in intact *N*-glycans (i.e., without permethylation, reduction, or other derivatization) extracted from pine nuts and human milk were identified. We demonstrate that fucose linkages in *N*-glycans, including the 1-3 linkage in the core of *N*-glycans, the 1-6 linkage in the core of *N*-glycans, the linkage in the α(1→3) antenna of biantennary *N*-glycans, and the linkage in the α(1→6) antenna of biantennary *N*-glycans, were identified through CID spectra. The procedure does not require the *N*-glycan standards, and the entire structural determination process is summarized in a flowchart for automation. Applications of this method to *N*-glycans extracted from bee venom led to the discovery of two unusual *N*-glycans, which are not expected to exist according to the current biosynthetic pathways.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Introduction

Glycosylation is one of the most important post-translational modification of proteins. It plays important role in many biological processes, such as protein folding, protein stability, and cell-cell interaction. [1-2] *N*-glycan is one major class of glycans in protein glycosylation. Among various structures of *N*-glycans, fucosylated *N*-glycans are related to many processes, including inflammation, immune responses, cancer metastasis, and they are commonly used as biomarkers for various diseases. [3-13]

Traditional methods to determine the structures of *N*-glycan involve repeating enzyme digestion [14-16] or mass spectrometry. [17-19] However, it is challenging to determine the fucose linkages in *N*-glycans, and most of the fucose linkage in *N*-glycans, particularly the fucose linkages in the antenna of *N*-glycans remain unidentified. Part of the difficulty results from the limited available enzymes in enzyme digestion, and another part of the difficulty results from the migration of fucose in protonated native *N*-glycans during tandem mass spectrometry analysis. [20-22] Permethylation prevents fucose migration [15, 23, 24], but it still introduces ambiguity and is laborious. [25] Turiak and co-workers showed that low energy collision energy dissociation (CID) reduced the fucose migration, and the core and antenna fucosylations can be distinguished using the ratios of fragment ions. [26] Alternatively, various labels at the reducing end of *N*-glycans were developed to differentiate the core and antenna fucosylations. [25, 27, 28] Harvey et al. shows that negative ion CID coupled with ion mobility is able to provide significant structural information on fucosylated *N*-glycans. [29] Hwang et al. reported that the core and antenna fucosylations can be differentiated using more than 800 tandem mass spectra of *N*-glycan standards as the training and test sets of machine learning. [30] These

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

studies demonstrated the differentiation of the fucose at the core from the fucose at the antenna of *N*-glycans. However, the differentiation of different linkages at the core (e.g., 1→3 linkage vs 1→6 linkage), the identification of the fucose at antenna (e.g., α(1→3) antenna vs α(1→6) antenna of biantennary *N*-glycans), and the determination of the fucose linkages in antenna (e.g., 1→2, 1→3, 1→4, or 1-6 in the α(1→3) or α(1→6) antenna) remain challenging.

Recently, we have developed a new mass spectrometry method for carbohydrate structure determination. [31-32] The new method is based on the dissociation mechanisms of carbohydrate in gas phase, and has been applied to determine the structures of carbohydrates in glycosphingolipid and milk [33, 34], and the *N*-glycans without fucose [35, 36]. In this work, we applied this method to determine the fucose linkages in *N*-glycans. Using collision induced dissociation of *N*-glycan sodium ion adducts, the intact *N*-glycans (i.e., no permethylation, reduction, or other derivatization) extracted from pine nuts and human milk was used for demonstration. We show that the fucose linkages, including the 1-3 linkage in the core of *N*-glycans, the 1-6 linkage in the core of *N*-glycans, the linkage in the 1-3 antenna of biantennary *N*-glycans, the linkage in the 1-6 antenna of biantennary *N*-glycans can be distinguished easily through CID spectra. The procedure does not require the *N*-glycan standards, and the entire structural determination procedures are summarized in flowcharts for automation. Applying this method to the *N*-glycans extracted from bee venom discovered two unusual *N*-glycans, Manα-(1→6)-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc and Manα-(1→3)-[Manα-(1-6)]-Man-β-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc. These structures are not expected to exist according to the current biosynthetic pathways.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Experimental methods

The pine nuts and bee venom were purchased from local market. Human milk was given by donors. PNGase A, PNGase F, and α1-2,3,4,6 fucosidase were purchased from New England Biolabs (Ipswich, MA, USA).

N-linked glycans were released from proteins through reactions with enzymes PNGase A or PNGase F. The released glycans were purified through ethanol precipitation, solid-phase extraction (C18 cartridge and nonporous graphitized carbon cartridge), followed by the separation from each other using two-dimensional high performance liquid chromatography (HPLC). A HPLC instrument equipped with a TSKgel amide-80 column (150 mm × 2.0 mm, particle size of 5 μm, Tosoh Bioscience GmbH, Griesheim, Germany) was used in the first dimension for the separation of *N*-glycans into different sizes, which avoids the interference in the structural determination of small *N*-glycans by the large *N*-glycans due to the in-source cracking of electrospray ionization [37]. The fractions collected from the first HPLC eluents were individually injected into the second HPLC instrument with a Hypercarb column (2.1 mm × 150 mm or 2.1 mm × 100 mm, particle size of 3 μm, Thermo Fisher Scientific, Waltham, MA, USA) for the second-dimension separation, which separated *N*-glycan isomers.

For nano-electrospray ionization mass spectrometry, the MS^n spectra were obtained using a linear ion trap mass spectrometer (LTQ XL, Thermo Fisher Scientific). Samples were prepared in a water/methanol mixture and 2 μL of each sample was loaded into a borosilicate glass nano-ESI emitter. For HPLC-electrospray mass spectrometry, the chromatograms and MS^n spectra were obtained using the same linear ion trap mass spectrometer with a heated electrospray ionization coupled to a HPLC system.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Results and Discussion

Instead of protonated *N*-glycans, sodium ion adducts of *N*-glycans were used in this work to eliminate fucose migration. The determination of fucose linkages in *N*-glycan is based on the collision induced dissociation mechanisms of *N*-acetylhexosamine and hexose sodium ion adducts reported in previous studies. [38-41] The dissociation mechanism of *N*-acetylhexosamine is illustrated in Figure 1(a). The major dissociation channels are cross-ring dissociation and dehydration. Cross-ring dissociation starts with a ring-opening reaction by transferring the H atom from O1 to O5, followed by retro-aldol reaction, resulting in the loss of neutral m = 101 Da. The major dehydration also starts with ring opening reaction, followed by a two-step H atom transfer, first from C2 to acetyl group and then to O3, and finally the cleavage of C3-O3 bond. Since both the cross-ring dissociation and the major dehydration start from ring-opening, these two reactions only occur for the *N*-acetylhexosamine located at the reducing end of oligosaccharides.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

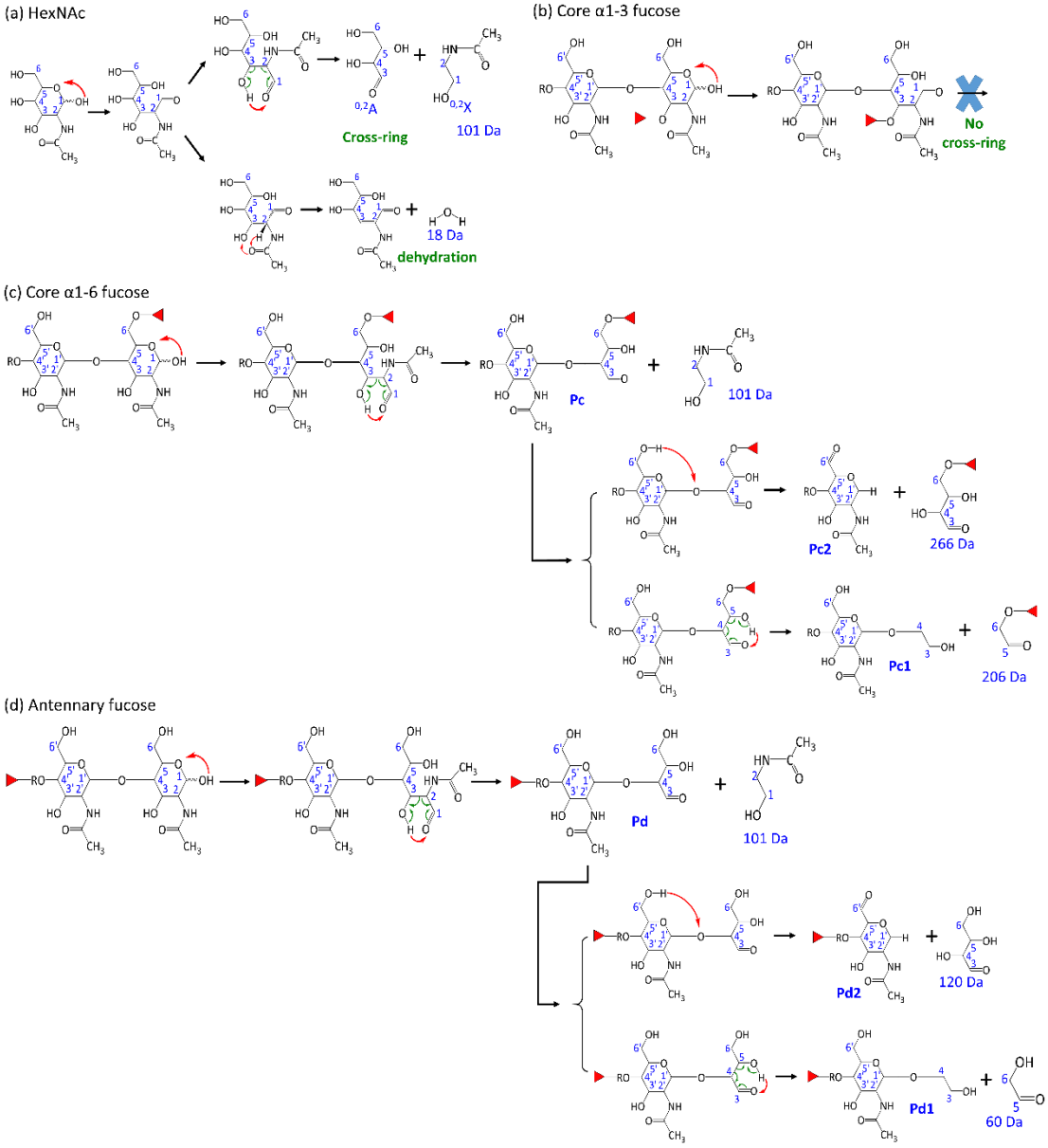


Figure 1. Dissociation mechanisms of *N*-acetylhexosamine sodium ion adducts. (a) *N*-acetylhexosamine, (b) *N*-glycans with a fucose located at the core of *N*-glycans through 1→3 linkage, (c) *N*-glycans with a fucose located at the core of *N*-glycans through 1→6 linkage, (d) *N*-glycans with a fucose located at antenna of *N*-glycans.

The location of fucose in *N*-glycans can be in the core of *N*-glycans with 1→3 linkage or 1→6 linkage, or in the α(1→3) or α(1→6) antenna of biantennary *N*-glycans. The aforementioned dissociation mechanisms [Figure 1(a)] suggest that (1)

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

cross-ring dissociation (loss of neutral m = 101 Da) and major dehydration do not occur for the *N*-glycans with the fucose located at the core of *N*-glycans through 1→3 linkage [Figure 1(b)], and (2) cross-ring dissociation and the major dehydration occur for the *N*-glycans with the fucose located at the core of *N*-glycans through 1→6 linkage [Figure 1(c)] and the *N*-glycans with the fucose located at the antenna (α(1→3) or α(1→6) antenna) of *N*-glycans [Figure 1 (d)].

To further distinguish between *N*-glycans with the fucose located at the core through the 1→6 linkage and at the antenna, the fragments generated from cross-ring dissociation [Pc and Pd in Figure 1(c) and (d)] are dissociated in the subsequent CID. Both Pc and Pd undergo additional retro-aldol reaction; however, resulting products differ. Pc results in the losses of neutrals m = 206 Da [Pc1 in Figure 1(c)] and 266 Da [Pc2 in Figure 1(c)], while Pd results in the losses of neutrals m = 60 Da [Pd1 in Figure 1(d)] and 120 Da [Pd2 in Figure 1(d)].

In multi-stage CID process, one of the fragment ions produced in CID is chosen as the precursor ion for the next stage CID. During this process, the fragments of the selected *m/z* value are isolated in the ion trap by rejecting ions of other *m/z* values from the ion trap. Then, the selected fragments are resonance excited by the electric field, resulting in the collisions with buffer gas and dissociation into fragments (product ions). The electric field is designed such that the product ions are confined in the same ion trap but they are not resonance excited by the electric field. Ideally, these product ions are quickly stabilized by the collisions with buffer gas and no subsequent dissociation, i.e., secondary dissociation, occurs. However, the stabilization is through vibration-to-translation energy transfer which is not very efficient. [42] If there is sufficient energy remaining in the product ions and if there is a dissociation channel with a low barrier of the product ions, secondary dissociation may take place. The

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

facile cleavage of the fucose glycosidic bond makes this secondary dissociation, i.e., elimination of fucose, easily. As a result: after Pd [Figure 2(d)] dissociates into Pd1 or Pd2 by CID (resulting in the loss of neutral m =60 or 120), the secondary dissociation can occur through the elimination of fucose. Consequently, the loss of neutral m = 206 and 266 Da also occurs for the *N*-glycans with fucose located at antenna. However, the intensities of ions resulting from the losses of neutrals m = 206 and 266 Da are lower compared to those resulting from losses of neutrals m =60 and 120 Da, because not all Pd1 or Pd2 undergo secondary dissociation. In contrast, ion Pc only undergo the losses of neutrals m = 206 and 266 Da, with almost no losses of neutrals m=60 and 120 observed. Therefore, the relative intensity of ions from the losses of neutrals m = 60, 120, 206, 266 Da can be used to distinguish the *N*-glycans with the fucose located at the core of *N*-glycans through α(1→6) linkage and at the antenna.

The aforementioned procedure is used to differentiate the *N*-glycans with fucose located at the core with 1→3 linkage, the *N*-glycans with fucose located at the core with 1→6 linkage, and the *N*-glycans with fucose located at the antenna. If the fucose is determined to be located at the antenna, the cross-ring dissociation mechanisms of hexose are used to determine which antenna of *N*-glycans the fucose is located. Fucose can be located at the α-(1→3) linked antenna or at the α-(1→6) linked antenna. The cross-ring dissociation mechanisms of the hexose at the reducing end of oligosaccharides with 1→3 and 1→6 linkages are illustrated in Figure 2(a) and (b), respectively. These cross-ring dissociation mechanisms of hexose start from ring-opening reaction, followed by the retro-aldol reaction. Figure 2(a) shows that hexose with 1-3 linkage undergoes loss of neutrals m = 90 Da, while hexose with 1-6 linkage undergoes losses of neutrals m = 60, 90, and 120 Da [Figure 2(b)]. The relative intensities of ions produced from the hexose with 1→6 linkage are I(loss of m

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

= 60) > I(loss of m = 90) > I(loss of m = 120). This is because loss of m= 60 occurs directly after ring-opening; loss of m = 90 occurs after ring-opening and one step of isomerization; loss of m = 120 occurs after ring-opening and two steps of isomerization. The more steps of isomerization required, the less competitive the dissociation channels are, and the empirical values of the relative intensities are 5: 3±1: 1.5±1. The application of the hexose dissociation mechanism to determine the antenna of *N*-glycans where the fucose is located is illustrated in Figure 2(c) using biantennary *N*-glycans as an example.

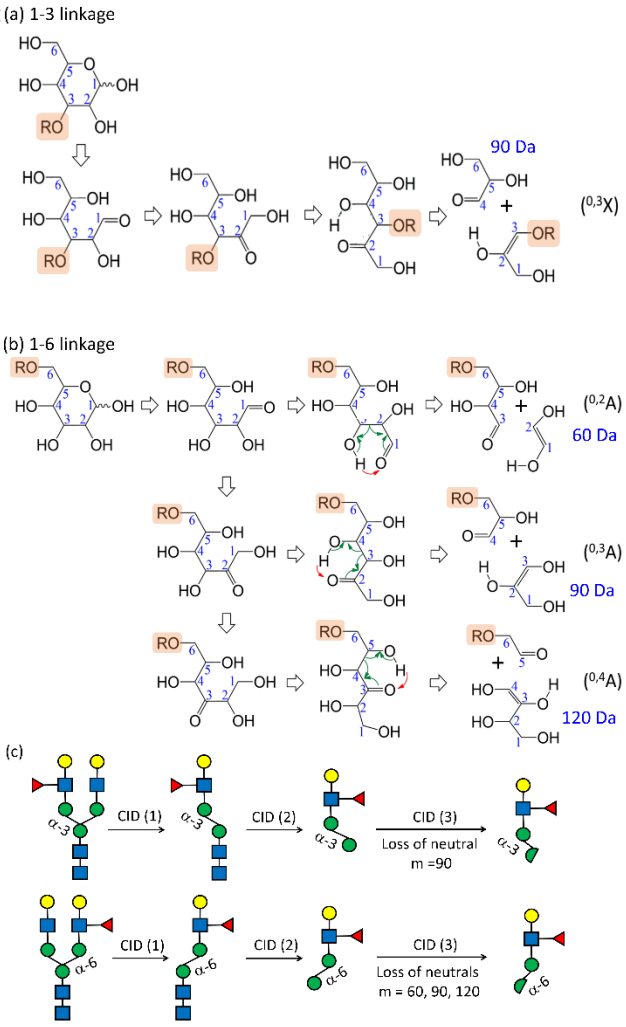


Figure 2. Dissociation mechanisms of hexose sodium ion adducts and the determination of the antenna that fucose is located. (a) hexose at the reducing end of oligosaccharide through a 1→3 linkage, (b)

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

hexose at the reducing end of oligosaccharide through a 1→6 linkage. (c) CID sequence to generate the precursor ion that mannose with 1→3 or 1-6 linkage is located at the reducing end (first two steps of CID), and subsequent CID to determine which antenna the fucose is located.

After the determination which antenna the fucose is located, the linkage of fucose in the antenna can be determined using the dissociation mechanisms of Hex and GlcNAc and the CID spectrum of the fragments consisting of fucose. For example, the linkages of fucose and galactose in the antenna can be Gal-(1-4)-[Fuc-(1-3)]-GlcNAc or Gal-(1-3)-[Fuc-(1-4)]-GlcNAc. They can be differentiated using the fragment Gal-[Fuc]-GlcNAc and the dissociation mechanism of GlcNAc illustrated in Figure 1(a). Both of them have large intensities of the fragment from the loss of neutral m= 146 Da due to the easy cleavage of fucose linkage, but Gal-(1-4)-[Fuc-(1-3)]-GlcNAc has large intensity of the fragment from the loss of neutral m = 164 Da compared to that of Gal-(1-3)-[Fuc-(1-4)]-GlcNAc due to the dissociation mechanism analogous to the dehydration of GlcNAc, while the Gal-(1-3)-[Fuc-(1-4)]-GlcNAc has large intensity of the fragment from the loss of neutral m = 180 Da compared to that of Gal-(1-4)-[Fuc-(1-3)]-GlcNAc due to the same mechanism. Figure 3 shows the difference between the CID spectra of Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc and Galβ-(1-3)-[Fucα-(1-4)]-GlcNAc. The intensity of fragment *m/z* = 388 (from the loss of neutral m = 164 Da) in the CID spectrum of Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc is larger than that of Galβ-(1-3)-[Fucα-(1-4)]-GlcNAc, while the intensity of fragment *m/z* 372 (from the loss of neutral m = 180 Da) in the CID spectrum of Galβ-(1-3)-[Fucα-(1-4)]-GlcNAc is larger than that of Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

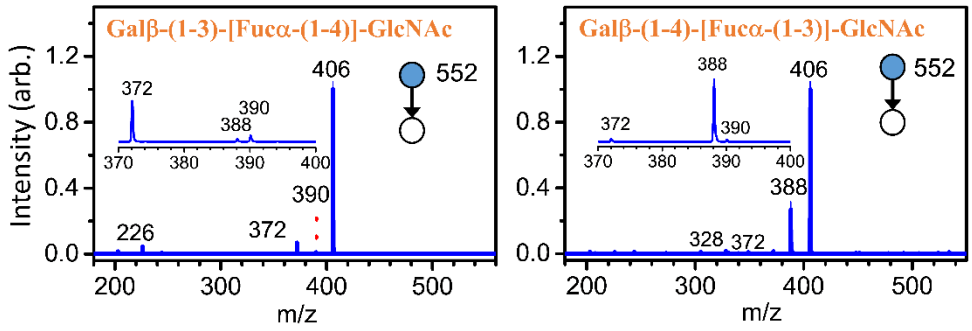
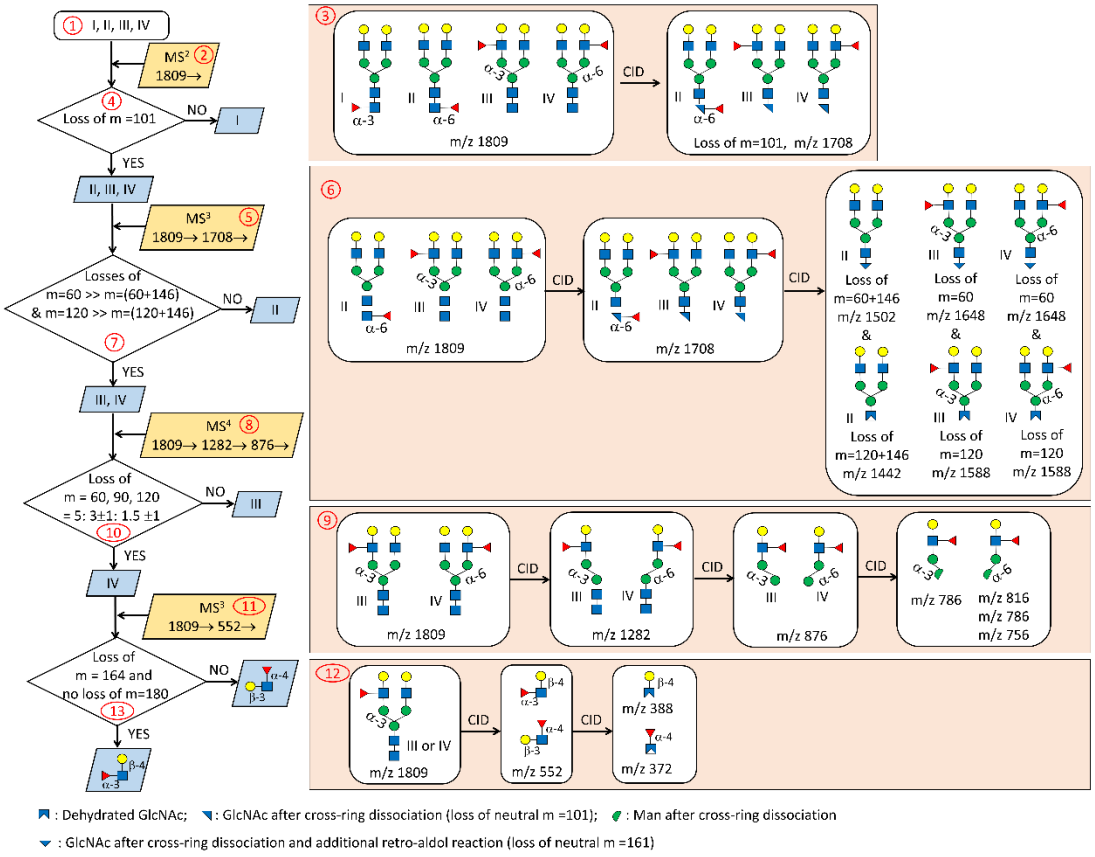


Figure 3. CID spectra of synthesized Galβ-(1-3)-[Fucα-(1-4)]-GlcNAc and Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc.

The aforementioned procedures for isomer identification are summarized in a flow chart, as illustrated in Figure 4 for a symmetric biantennary *N*-glycan consisting of one fucose. This flow chart can be used to develop programs of automatic fucose linkage determination.



https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Figure 4. Flow chart of procedures for determining fucose linkage in a symmetric biantennary *N*-glycan consisting of one fucose. The procedure is divided into four parts. In the first part (blocks 1-4), isomer I is differentiated from isomers II, III, and IV. Starting from block 1, which includes all isomers, MS^2 spectrum through the CID sequence 1809→ (block 2) is used, and the structures of isomers and fragments produced through this CID sequence are shown in block 3. Only isomers II, III, and IV generate cross-ring fragment *m/z* 1708 (loss of neutral m=101), and decision in block 4 differentiates isomer I from the other isomers. In the second part (blocks 5-7), isomer II is differentiated from isomers III and IV through the MS^3 spectrum of CID sequence 1809→ 1708→ (blocks 5). The structures of isomers and fragments produced through this CID sequence are shown in block 6. The elimination of neutrals m=60 and 120 from isomers III and IV are easier than that of isomer II, and decision in block 7 differentiates isomers II from III and IV. In the third part (blocks 8-10), isomers III and IV are differentiated. MS^4 spectrum through CID sequence 1809→ 1282→ 876→ (block 8) is used to distinguish isomers III and IV. The structures of isomers and fragments produced through this CID sequence are shown in block 9 and decision in block 10 differentiates isomers III and IV. In the fourth part (blocks 11-13), MS^3 spectrum through CID sequence 1809→ 552→ (block 11) is used to differentiate Gal-β-(1-4)-[Fuc-α-(1-3)]-GlcNAc and Gal-β-(1-3)-[Fuc-α-(1-4)]-GlcNAc.

Here, we use the fucosylated *N*-glycans extracted from pine nuts, human milk, and bee venom to demonstrate the determination of fucose linkage in *N*-glycans. Our focus is specifically on determining the fucose linkage; the identifications of the other glycosidic bond linkages in the *N*-glycans was reported in our previous study [35] and they are not repeated in this work.

(A) *N*-glycan $(Man)_3Xyl(GlcNAc)_2Fuc$ extracted from pine nuts. The *N*-glycan with molecular weight of 1188 Da extracted from pine nuts corresponds to $(Man)_3Xyl(GlcNAc)_2Fuc$ which consists of one fucose. The CID spectrum through the CID sequence, 1211 (sodium ion adduct)→, is illustrated in Figure 5(a). Fragment ion *m/z* 1110, corresponding to the loss of neutral m = 101, was not found, indicating the fucose is located at the core of the *N*-glycan through a 1→3 linkage. Additionally, the intensity of fragment from dehydration, *m/z* 1193, is low , further supporting the fucose linkage at the core of *N*-glycan is 1→3. The fucose linkage was crosschecked

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

by the CID spectrum after fucose elimination. After fucose is eliminated by CID [ion *m/z* 1065 in Figure 5(b)], fragment ion *m/z* 964 produced from cross-ring dissociation (loss of neutral m = 101) and fragment *m/z* 1047 produced from dehydration were found. These two fragments support the fucose located at the core of the *N*-glycan through 1→3 linkage.

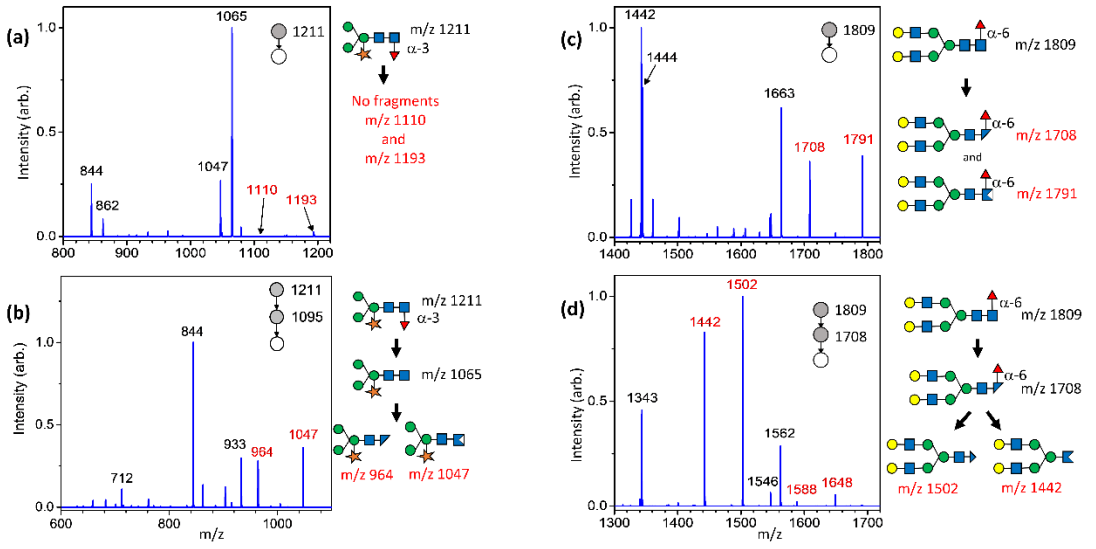


Figure 5. CID spectra, CID sequences, and structural changes in CID for determining fucose linkage in *N*-glycans. (a) and (b) *N*-glycan $(Man)_3Xyl(GlcNAc)_2Fuc$ extracted from pine nut; (c) and (d) *N*-glycan of isomer 1 of $(Hex)_5(GlcNAc)_4Fuc$ extracted from human milk. The *m/z* values of the fragments used to determine the structures are highlighted in red.

(B) *N*-glycan $(Hex)_5(GlcNAc)_4Fuc$ extracted from human milk. The chromatogram reveals two isomers of *N*-glycans with molecular weight of 1786 Da extracted from human milk. The composition is $(Hex)_5(GlcNAc)_4Fuc$, containing one fucose.

Isomer 1 of $(Hex)_5(GlcNAc)_4Fuc$. The CID spectrum through the CID sequence, 1809 (sodium ion adduct)→, is illustrated in Figure 5(c). A fragment ion *m/z* 1708, corresponding to the loss of neutral m = 101 Da from cross-ring dissociation, was found, indicating the fucose is located either at the core of the *N*-glycan through a

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

1→6 linkage or at the antenna. The large intensity of fragment ion produced from H_2O elimination, fragment *m/z* 1991, supports that fucose is not located at the core of the *N*-glycan through a 1→3 linkage. The CID spectrum through the CID sequence, 1809 (sodium ion adduct)→1708 (loss of neutral m = 101 Da)→, [Figure 5(d)] shows the intensities of ions from the losses of neutrals m = 206 Da (ion *m/z* 1502) and m =266 Da (ion *m/z* 1442) are much larger than those from the losses of neutrals m =60 Da (ion *m/z* 1648) and m = 120 Da (ion *m/z* 1588), indicating that the fucose is located at the core of the *N*-glycan through a 1→6 linkage.

Isomer 2 of $(Hex)_5(GlcNAc)_4Fuc$. It was found that the intensities of doubly charged ion (two sodium ions attached to the *N*-glycan, *m/z* 916) and singly charge ion (one sodium ion attached to the *N*-glycan, *m/z* 1809) are similar for this isomer. Both of them can be used for structural analysis. The CID spectrum obtained through the CID sequence, 916 (adduct of two sodium ions)→, is illustrated in Figure 6(a). Fragment ions *m/z* 856.5 and 907, corresponding to the loss of neutral m = 101 Da and dehydration, were found, indicating fucose is located either at the core of the *N*-glycan through a 1→6 linkage or at the antenna. The CID spectrum through the CID sequence, 916→865.5 (loss of neutral m =101)→, [Figure 6(b)], shows the intensities of ions from the losses of neutrals m = 60 Da (ion *m/z* 835.5) and m = 120 Da (ion *m/z* 805.5) are larger than those resulting from losses of neutrals m = 206 Da (ion *m/z* 762.5) and m = 266 Da (ion *m/z* 732.5), indicating that the fucose is located at the antenna of *N*-glycan. The presence of ion *m/z* 534, which consists of a hexose, a *N*-acetylhexosamine, and a fucose in Figure 6(b), supports that the fucose is located at the antenna of the *N*-glycan.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

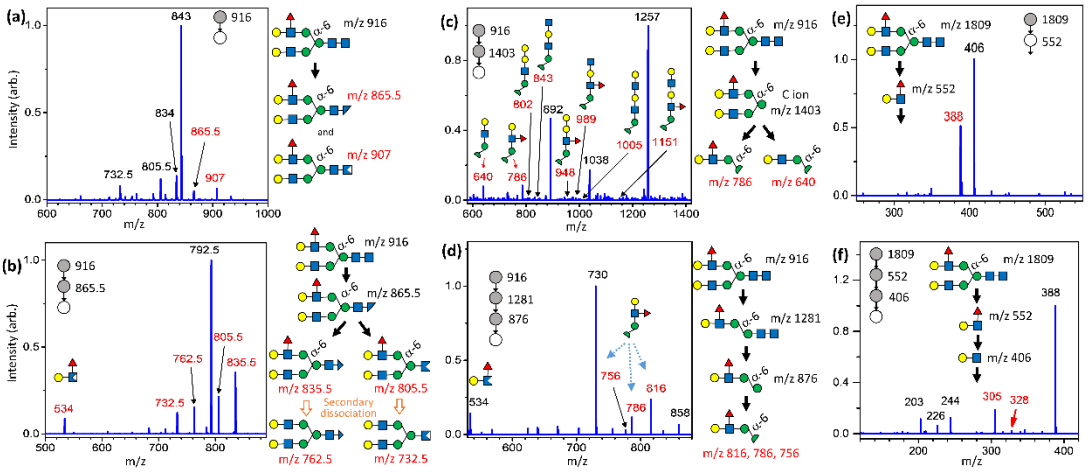


Figure 6. CID spectra, CID sequences, and structural changes in CID for determining fucose linkage in *N*-glycan of isomer 2 of $(Hex)_5(GlcNAc)_4Fuc$ extracted from human milk. The *m/z* values of the fragments used to determine the structures are highlighted in red.

To determine which antenna the fucose is located, the structure of antenna must first be determined. In brief, the C ions (*m/z* 1403) produced by the cleavage of the β-1→4 glycosidic bond between Man and GlcNAc in the core of *N*-glycan is dissociated, and the fragments produced from the cross-ring dissociation of the mannose at the reducing end of this C ion are used to determine the length of the antenna. The CID spectrum obtained through the CID sequence 916 (adduct of two sodium ions, doubly charged) →1403 (singly charged C ion)→fragments is illustrated in Figure 6(c). Fragment ions *m/z* 640 and 786 were found, but fragment ions *m/z* 802, 843, 948, 989, 1005, or 1151 were not found [structures of these ions are illustrated in Figure 6(c)], indicating the biantennary *N*-glycan without fucose is symmetric and each antenna consists of two hexoses and one *N*-acetylhexosamine.

After the structure of antenna is identified, the next step is to determine the linkage of fucose in the antenna. The CID spectrum obtained through the CID sequence, 916 (adduct of two sodium ions, doubly charged) →1281 (adduct of one

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

sodium ion, singly charged) →876 (adduct of one sodium ion, singly charged) →fragments, is presented in Figure 6(d). In this CID sequence, ion *m/z* 1281 represents the elimination of the antenna without fucose from precursor ions [structure is shown in Figure 6(d)], and ion *m/z* 876 represents the sodium ion adduct of the antenna consisting of fucose [structure is shown in Figure 6(d)]. Ion *m/z* 876 contains a mannose at the reducing end, the cross-ring dissociation mechanism of hexose is used to determine the linkage of antenna. The CID of ion *m/z* 876 generates fragment ions *m/z* 816, 786, and 756 from the loss of neutrals m = 60, 90, and 120 Da with relative intensities 5 : 2.5 : 1, indicating the linkage of the antenna consisting of fucose is 1→6. Consequently, the fucose is determined to be located at the α(1→6) antenna of the biantennary *N*-glycan.

The fucose linkage at the α(1→6) antenna was determined using two CID sequences, 1809 (adduct of one sodium ions, singly charged) →552 (adduct of one sodium ion, singly charged) →fragments and 1809 →552 →406 →fragments. Comparing the CID spectrum in Figure 6(e) to that in Figure 3 suggests that the linkage is Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc. The CID spectrum in Figure 6(f) shows the fragment ions *m/z* 328 (from the loss of neutral m = 78) and *m/z* 305 (from the cross-ring dissociation by elimination of neutral m = 101 Da), indicating the linkage is Gal-(1-4)-GlcNAc, which supports the structural assignment of Figure 6(e).

(C) *N*-glycan $(Hex)_5(GlcNAc)_4Neu5Ac(Fuc)_2$ extracted from human milk. This is another *N*-glycan extracted from human milk, with a molecular weight of 2223 Da. This *N*-glycan consists of one sialic acid and two fucoses. Both sialic acid and fucose are well known to be easily eliminated during the fragmentation process in tandem mass spectrometry, which makes the structure determination difficult. To overcome this challenge, the structural determination process is divided into two parts. In the

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

first part, the fragment produced from precursor ion after the loss of sialic acid is used to determine the fucose linkage. In the second part, the fragment generated from precursor ion after the loss of fucose is used to determine the sialic acid linkage. In our previous report, we demonstrated how to determine the linkage of sialic acid in glycosphingolipids and *N*-glycans. [43, 44] Here, our focus is on the determination of fucose linkage, and the determination of sialic acid linkage is not repeated in this work.

The singly charged ion of this *N*-glycan is exceeds the upper limit of the mass range of our mass spectrometer, necessitating the use of the doubly charged ion. The doubly charged ion *m/z* 1145.5 represents the *N*-glycan with three sodium ions, one sodium ion replaces the H^+ in the carboxylic acid of sialic acid and the other two sodium ions are attached to the *N*-glycan resulting in a doubly charged state. The CID spectrum, acquired through the CID sequence 1145.5 (doubly charged, with three sodium ions attached) →1955 (singly charged with one sodium ion attached)→fragments, is illustrated in Figure 7(a). In this CID sequence, ion *m/z* 1955 represents the *N*-glycan after the elimination of sialic acid. Ion *m/z* 1854, which corresponds to the loss of neutral m = 101 Da (found in Figure 7a)), indicating fucose is not located at the core of the *N*-glycan with a 1→3 linkage. The CID spectrum through the CID sequence, 1145.5 →1955→ 1854→fragments, is illustrated in Figure 7(b). The intensities of ions resulting from the losses of neutrals m = 206 Da (ion *m/z* 1648) and m =266 Da (ion *m/z* 1588) are much larger than those resulting from the losses of neutrals m =60 Da (ion *m/z* 1794) and m = 120 Da (ion *m/z* 1734). This indicates that one fucose is located at the core of the *N*-glycan through a 1→6 linkage.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

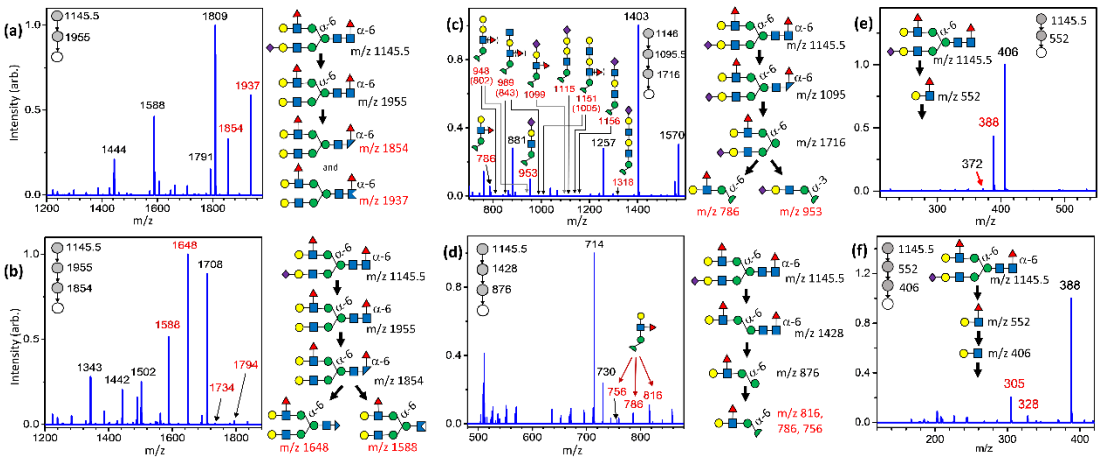


Figure 7. CID spectra and CID sequences for determining fucose linkage in the *N*-glycan of $(Hex)_5(GlcNAc)_4Neu5Ac(Fuc)_2$ extracted from human milk. The *m/z* values of the fragments used to determine the structures are highlighted in red.

The procedure for determining the linkage of the other fucose located at the antenna of the *N*-glycan follows similar steps to those in the previous example. The length of antennary is determined using the CID spectrum through the CID sequence, 1145.5→ 1095→ 1716→fragments, as presented in Figure 7(c). In this CID spectrum, ions at *m/z* 786 and 953 were observed, whereas ions *m/z* 802, 843, 948, 989, 1005, 1099, 1115, 1151, 1156, and 1318 were not found. The structures corresponding to these ions are illustrated in Figure 7(c). This CID spectrum suggests (1) the biantennary N-glycan without fucose and sialic acid exhibits symmetry, with each antenna comprising one *N*-acetylhexosamine and two hexoses, and (2) sialic acid and fucose are located on different antenna. In the final step, the CID spectrum through the sequence, 1145.5→ 1428→ 876→fragments is utilized to determine the location of the fucose on the antenna. Ions resulting from the losses of neutrals m = 60 (ion *m/z* 816), 90 (ion *m/z* 786), and 120 (ion *m/z* 756) with relative intensities 5 : 2.6 : 1.2 suggest the fucose is located at the α(1→6) antenna. The fucose linkage in the α(1→6) antenna of the biantennary *N*-glycan is determined using the CID spectra in Figure 7(e)

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

and (f). Comparing the CID spectrum in Figure 7(e) to that in Figure 3 suggests that the linkage is Galβ-(1-4)-[Fucα-(1-3)]-GlcNAc. The fragment ions *m/z* 328 (from the loss of neutral m = 78) and *m/z* 305 (from the cross-ring dissociation by elimination of neutral m = 101 Da) in Figure 7(f) indicates the linkage is Gal-(1-4)-GlcNAc, which supports the structural assignment of Figure 7(e).

(D) *N*-glycans extracted from been venom

The aforementioned *N*-glycans demonstrates that distinguishing fucose in the core of *N*-glycans with 1→3 and 1→6 linkages is straightforward. One might argute that these can be differentiated based on the sources of *N*-glycans, as fucose in the core of plant *N*-glycans typically has a 1→3 linkage, whereas in mammalian *N*-glycans it is a 1→6 linkage. However, relying on sample sources to identify fucose linkage in the core of *N*-glycans does not apply to insect *N*-glycans, where the linkages of fucose in the core can be either 1→3 or 1→6.

Figure 8(a) shows the HPLC chromatogram of ion *m/z* 1079, representing the $(Hex)_3(HexNAc)_2dHex$, where dHex represents deoxyhexose. Here, a PGC column was used for isomer separation. PGC column is also known for separating anomeric isomers, so two peaks in Figure 8(a) might be the alpha and beta anomers of HexNAc at the reducing end of the same isomer. To clarify, the eluents separated by PGC column were collected into different tubes, corresponding to each peak in Figure 8(a). These eluents were stored at room temperature for more than 6 hours and were subsequently reinjected individually into the same PGC column. If the two peaks in Figure 8(a) belong to one isomer and differ only in the anomericity at the reducing end, the reinjection of the eluents into the same PGC column would result in two peaks in chromatogram, even though only one anomer was collected in a tube. The relative intensities and the retention times of these two peaks should be the same as

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

those in Figure 8(a). This is because two anomers of the same isomer undergo mutarotation and reach equilibrium during the storage time. The PGC chromatograms of these eluents, as illustrated in Figure 8(b), show the peaks at t = 6.4 and 7.0 min belongs to one isomer, peaks at t = 11.1 and 15.5 min belong to the second isomer, and peaks at t = 27.3 and 32.8 min belong to the third isomer.

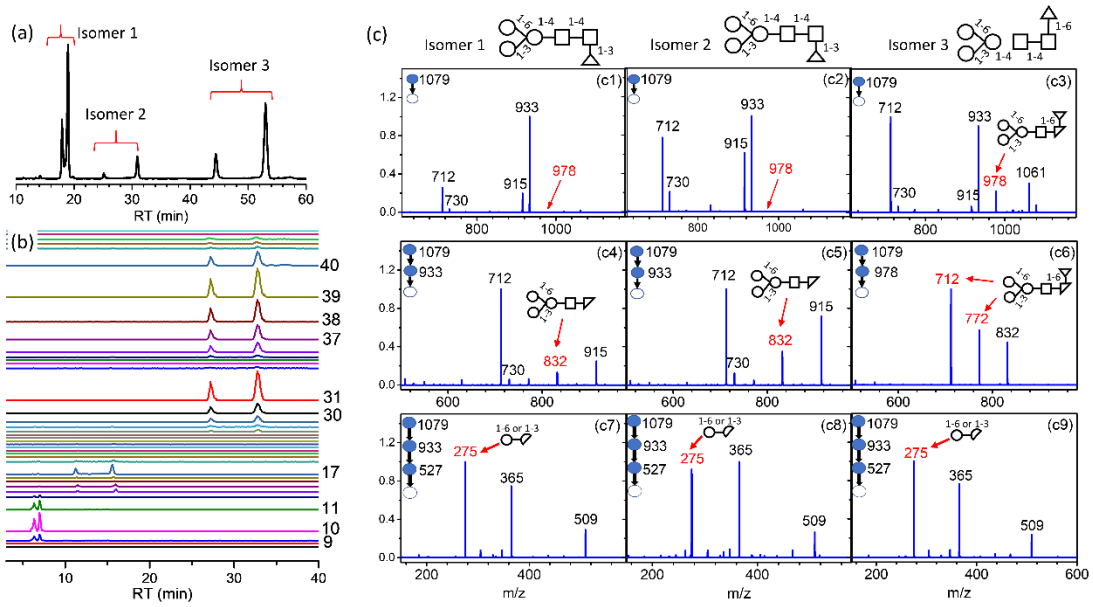


Figure 8. Chromatograms and CID spectra of the $(Hex)_3(HexNAc)_2dHex$ extracted from bee venom. (a) Chromatograms and CID spectra of the $(Hex)_3(HexNAc)_2dHex$ separately by PGC column (15 cm), y axis represents the intensity of total fragments from ion *m/z* 1079. (b) Chromatogram of the eluents fractionally collected from (a) and reinjected into PGC column (10 cm). y axis represents the intensity of total fragments from ion *m/z* 1079. Notably, the lengths of PGC column used in (a) and (b) are different. (c) CID spectra of three isomers. The fragment used for structural determination are highlighted in red.

The CID spectra of these isomers, as illustrated in Figure 8(c1)-(c3), show that fragment ion *m/z* 978, corresponding to the loss of neutral m = 101 from cross-ring dissociation, was not found in isomers 1 and 2. This suggests that the dHex is located

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

at the core of the *N*-glycan through a 1→3 linkage for isomer 1 and 2. Meanwhile, fragment ion *m/z* 978, corresponding to the loss of neutral m = 101, was found in isomer 3, suggesting dHex is located at the core with a 1-6 linkage or at the antenna. The CID spectrum of isomer 3 through the sequence 1079→ 978→ fragments [Figure 8(c6)] reveals that the intensities of fragment ions *m/z* 772 (loss of neutral m = 60+ 146) and 712 (loss of neutral m = 120 + 146) are much larger than those of *m/z* 918 (loss of neutral m = 60) and 858 (loss of neutral m = 120), suggesting the dHex is located at the core through a 1→6 linkage for isomer 3. The loss of neutral m=101 after the elimination of deoxyhexose for isomers 1 and 2 generates ion *m/z* 832 [Figure 8(c1) and (c2)], which supports that deoxyhexose is located at the core through a 1→3 linkage for these two isomers.

Because the dHex in both isomers 1 and 2 is located at the core with a 1→3 linkage, one possible difference in the structure between isomers 1 and 2 is the linkages of three mannoses. However, the surprising result is that the CID spectra through the sequence, 1079→ 933 → 527→ fragments, suggest that all three isomers have a structure with Hex-(1→3)-[Hex-(1-6)]-Hex linkage, as indicated by the ion *m/z* 275 illustrated in Figure 8(c7)-(c9).

To further determine the structure, we used the enzyme, α1-2,3,4,6 fucosidase. The reactions of these isomers with α1-2,3,4,6 fucosidase show that dHex can be removed from isomers 1 and 3 partially within 12 hours, resulting in a change from ion *m/z* 1079 to *m/z* 933; however, no product ion m/z 933 was found for isomer 2 [Figure 9]. The products *m/z* 933 generated from the removal of isomers 1 and 3 have the same retention time as the synthesized standard, Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-GlcNAc, suggesting isomer 1 is

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fucα-(1→3)]-GlcNAc and isomer 3 is Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fucα-(1→6)]-GlcNA. This supports the structural assignment by mass spectra in Figure 8.

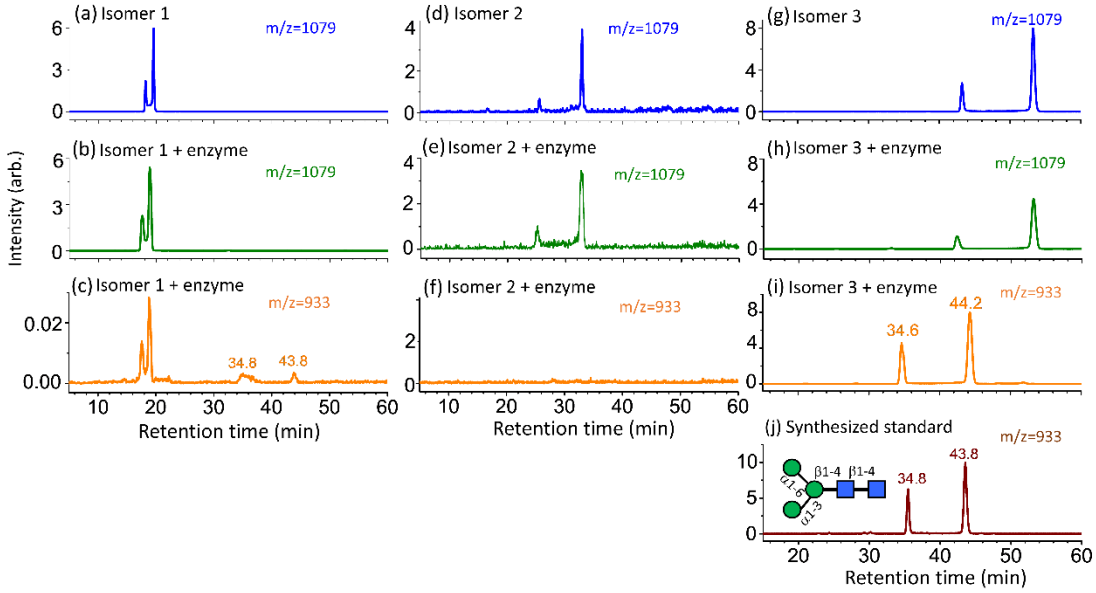


Figure 9 Chromatograms of isomers 1-3 before and after α1-2,3,4,6 fucosidase digestion (a)-(i) and the chromatogram of synthesized *N*-glycan (j).

The partial hydrolysis by acid shows that the retention time of isomer 1 after removal of fucose, is the same as that of Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-GlcNAc, while the retention time of isomer 2 after removal of dHex, is the same as that of Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-ManNAc [Figure 10(a)-(c)]. The aforementioned evidences suggest isomer 2 is Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[dHex-(1→3)]-ManNAc.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

To find out the monosaccharide composition of dHex in isomer 2, we utilized acid hydrolysis followed by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAE-PAD). The monosaccharides found by HPAE-PAD after acid hydrolysis of isomer 2 in 2M trifluoroacetic acid for 4 hours included Fuc, GlcNAc, ManNAc, Man, GlcN, and ManN [Figure 10(d)], where the GlcN and ManN not found in mass spectra results from the conversion of GlcNAc and ManNAc to GlcN and ManN during acid hydrolysis. Fucose was found in the acid hydrolysis of isomer 2, suggesting isomer 2 is Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc, where the anomericity of the fucose linkage remains unidentified. The resistance of the removal of fucose from isomer 2 by α1-2,3,4,6 fucosidase suggests two possibilities: (1) the fucose linkage is not in α anomeric configuration, or (2) the fucose linkage is in α anomeric configuration, but the activity of α1-2,3,4,6 fucosidase is affected by the ManNAc.

It has been shown that the GlcNAc at the reducing end of *N*-glycans is epimerized to ManNAc in alkaline solution. [45] However, the solution we used in the entire process of sample preparation, including enzyme release of *N*-glycans from proteins and purification of *N*-glycans, was acidic (pH=6.0, buffer solution of PNGase A) or pH=7.0. Therefore, it is not likely that GlcNAc at the reducing end of *N*-glycans is epimerized to ManNAc during sample preparation process. The potential artifact was crosschecked using the following method. Isomers 1 and 3 were added separately to the buffer solution containing PNGase A, followed by the processes of *N*-glycan release, extraction, and purification. No isomer 2 was found, indicating that isomers 1 or 3 would not convert to isomer 2 due to the artifact during the processes of *N*-glycan release, extraction, and purification. Analogous three isomers of $Man_2FucHexNAc_2$

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

were also found.

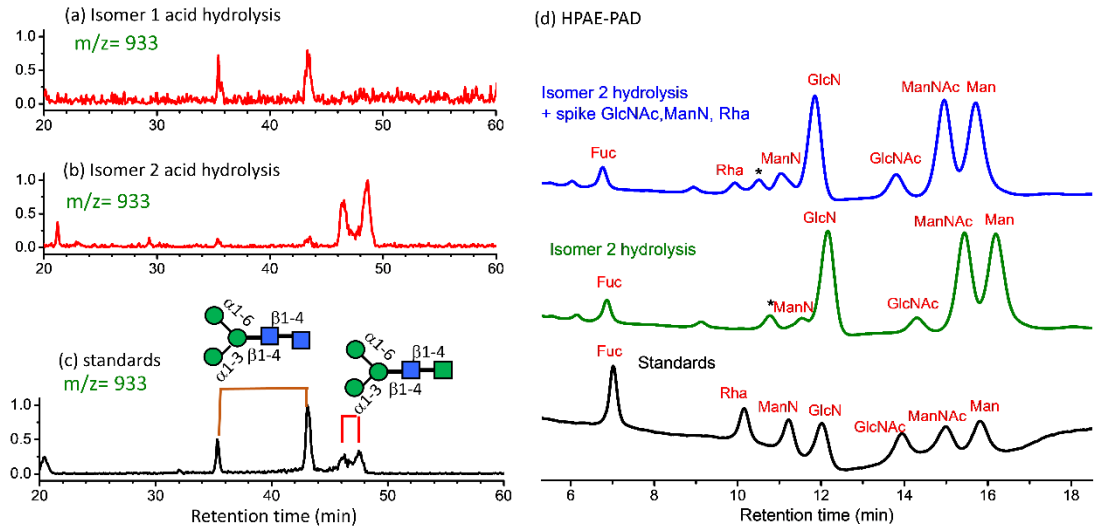


Figure 10. (a) and (b): Chromatograms of the product ion, *m/z* 933, generated from acid hydrolysis of isomer 1 and 2, respectively. (c) Chromatogram of Man-α-(1-3)-[Man-α-(1-6)]-Man-β-(1-4)-GlcNAc-β-(1-4)-GlcNAc and Man-α-(1-3)-[Man-α-(1-6)]-Man-β-(1-4)-GlcNAc-β-(1-4)-ManNAc, where Man-α-(1-3)-[Man-α-(1-6)]-Man-β-(1-4)-GlcNAc-β-(1-4)-GlcNAc is synthesized *N*-glycans, and Man-α-(1-3)-[Man-α-(1-6)]-Man-β-(1-4)-GlcNAc-β-(1-4)-ManNAc was generated from Man-α-(1-3)-[Man-α-(1-6)]-Man-β-(1-4)-GlcNAc-β-(1-4)-GlcNAc in NH_3 solution (12 hrs, 60^{o}C). (d) HPAE-PAD chromatogram of the products generated from isomer 2 acid hydrolysis, indicating dHex in isomer 2 is fucose.

These unusual *N*-glycans, Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc and Manα-(1→6)-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc, were also found in the solution of bee venom without using enzyme PNGase A to release *N*-glycans from proteins. Figure 11 shows the intensities of three isomers extracted

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

from the same batch and same amount of bee venom using following two conditions: (1) only bee venom and buffer solution, and (2) bee venom and buffer solution with enzyme PNGase A. The same amount of a disaccharide, Fuc-β-(1-3)-Gal, was added to both solutions as the internal standard for intensity calibration. When only bee venom and buffer solution were used (condition 1), isomers 1 and 2 were found with small intensities [Figure 11 (a) and (c)]. When bee venom and buffer solution with PNGase A were used (condition 2), isomers 1, 2, and 3 were found [Figure 11 (b) and (d)]. While the intensity of isomer 1 increases by a factor of 20-30 when PNGase A was used, isomer 2 only increases by less than a factor of two. The existence of isomers 1 and 2 in condition 1 indicates that part of them exists as free *N*-glycans. The small increase of isomer 2 from condition 1 to condition 2 indicates that 50% of isomer 2 found in chromatogram originally exists as a free *N*-glycan. However, it is not clear whether isomer 2 mainly exists as free *N*-glycan, or if it is also linked to a protein but cannot be efficiently released by enzyme PNGase A. Although the amount of isomer 2 is small, these unusual *N*-glycans have not been reported before and they are not expected to exist according to the current biosynthetic pathways. [46]

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

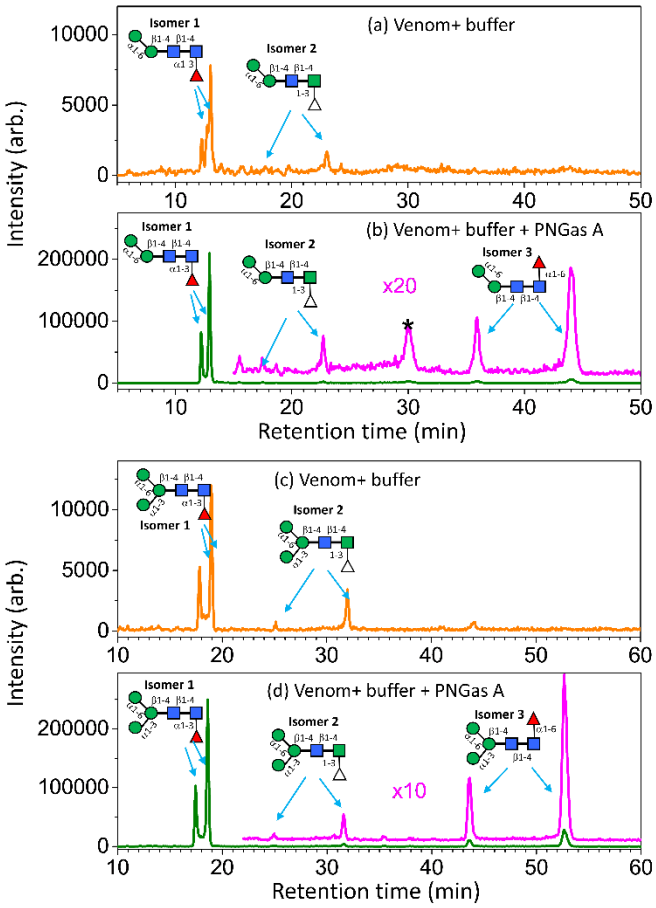


Figure 11. Relative intensities of *N*-glycan extracted from bee venom in two conditions. Condition1: bee venom and buffer solution, (a) and (c). Condition 2: bee venom and buffer solution with PNGase A, (b) and (d). * represents impurity.

Conclusions

We demonstrate a simple tandem mass spectrometry method to determine the linkages of fucose in *N*-glycans. This method is based on the collision induced dissociation mechanisms of *N*-acetylhexosamine and hexose sodium ion adducts. No derivatization, such as permethylation, reduction, or labelling at the reducing end, is required. This greatly simplifies the structural determination procedure and eliminates any possible side reaction during derivatization. Although the locations of fucose in

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

the core of *N*-glycans can be identified by the source of samples, 1→3 linkage for the *N*-glycans from plants and 1→6 linkage for the *N*-glycans from mammal, our approach provides a de novo method to determine the linkages of the fucose in *N*-glycans extracted from any species. This method does not required *N*-glycan standards and is particularly useful for identifying the antenna where fucose is located, determining the linkage of fucose in antenna, and distinguishing the linkage (1→3 or/and 1→6) of fucose in the core of *N*-glycans, such as those extracted from insects. The entire procedure can be coded in the computer that controls the mass spectrometer for automated structural determination, an ultimate goal in carbohydrate structural analysis. Applying this method to the *N*-glycans extracted from bee venom has led to the discovery of *N*-glycans, Manα-(1→3)-[Manα-(1→6)]-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc and Manα-(1→6)-Manβ-(1→4)-GlcNAcβ-(1→4)-[Fuc-(1→3)]-ManNAc, which are not expected to exist according to the current biosynthetic pathways. The discovery of these *N*-glycans suggests the existence of previously undiscovered biosynthetic pathways.

Acknowledgements

This work was supported by the National Science and Technology Council, Taiwan (Chemical bonding, 112-2113-M-007-011).

Conflict of interest

C.K.N. is one of inventors of a US patent (US 10,796,788 B2); a part of the method described in the patent for determining the carbohydrate structure was used in this work. All other authors declare no competing interests.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

References

1. Shental-Bechor, D.; Levy, Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105**,** 8256–8261. DOI: 10.1073/pnas.0801340105
2. Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. Conformational influences of glycosylation of a peptide: a possible model for the effect of glycosylation on the rate of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93**,** 12759–12761. DOI: https://doi.org/10.1073/pnas.93.23.1275
3. Huhn, C.; Selman, M. H.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. IgG glycosylation analysis. *Proteomics.* **2009,** 9(4), 882–913. DOI: https://doi.org/10.1002/pmic.200800715
4. Pucić, M.; Knezević, A.; Vidic, J.; Adamczyk, B.; Novokmet, M.; Polasek, O.; Gornik, O.; Supraha-Goreta, S.; Wormald, M. R; Redzić, I.; Campbell, H.; Wright, A.; Hastie, N. D; Wilson, J. F; Rudan, I.; Wuhrer, M.; Rudd, P. M; Josić, D.; Lauc, G. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol. Cell Proteomics* **2011,** 10(10), M111 010090. DOI: 10.1074/mcp.M111.010090.
5. Nakagawa, T.; Matsumoto, H.; Shinzaki, S.; Kamada, Y. *Biomolecules* **2012**, 2, 34−45. DOI: https://doi.org/10.3390/biom2010034
6. Chen, C. Y.; Jan, Y. H.; Juan, Y. H.; Yang, C. J.; Huang, M. S.; Yu, C. J.; Yang, P. C.; Hsiao, M.; Hsu, T. L.; Wong, C. H. Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. *Proc. Natl. Acad. Sci.* **2013**, 110, 630–635. DOI: 10.1073/pnas.1220425110
7. Yang, X.; Wang, J.; Liu, S.; Yan, Q. HSF1 and Sp1 regulate FUT4 gene expression and cell proliferation in breast cancer cells. *J. Cell Biochem.* **2014**, 115,

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

168–178. DOI: 10.1002/jcb.24645

8. Singh, S.; Pal, K.; Yadav, J.; Tang, H.; Partyka, K.; Kletter, D.; Hsueh, P.; Ensink, E.; KC, B.; Hostetter, G.; Xu, H. E.; Bern, M.; Smith, D. F.; Mehta, A. S.; Brand, R.; Melcher, K.; Haab, B. B. Upregulation of glycans containing 3' fucose in a subset of pancreatic cancers uncovered using fusion-tagged lectins. *J. Proteome Res.* **2015**, 14 (6), 2594−2605. DOI: https://doi.org/10.1021/acs.jproteome.5b00142
9. Ito, E.; Oka, R.; Ishii, T.; Korekane, H.; Kurimoto, A.; Kizuka, Y.; Kitazume, S.; Ariki, S.; Takahashi, M.; Kuroki, Y.; Kida, K.; Taniguchi, N. Fucosylated surfactant protein-D is a biomarker candidate for the development of chronic obstructive pulmonary disease. *J. Proteomics* **2015**, 127, 386−394. DOI: https://doi.org/10.1016/j.jprot.2015.07.011
10. Holst, S.; Deuss, A. J. M.; van Pelt, G. W.; van Vliet, S. J.; Garcia-Vallejo, J. J.; Koeleman, C. A. M.; Deelder, A. M.; Mesker, W. E.; Tollenaar, R. A.; Rombouts, Y.; Wuhrer, M. N-glycosylation profiling of colorectal cancer cell lines reveals association of fucosylation with differentiation and caudal type homebox 1 (CDX1)/villin mRNA expression. *Mol. Cell. Proteomics* **2016**, 15 (1), 124−140. DOI: https://doi.org/10.1074/mcp.M115.051235
11. Sakae, Y.; Satoh, T.; Yagi, H.; Yanaka, S.; Yamaguchi, T.; Isoda, Y.; Iida, S.; Okamoto, Y.; Kato, K. Conformational effects of N-glycan core fucosylation of immunoglobulin G Fc region on its interaction with Fcγ receptor IIIa. *Sci. Rep.* **2017**, 7 (1), 13780. DOI:10.1038/s41598-017-13845-8
12. Awan, B.; Turkov, D.; Schumacher, C.; Jacobo, A.; McEnerney, A.; Ramsey, A.; Xu, G.; Park, D.; Kalomoiris, S.; Yao, W.; Jao, L. E.; Allende, M. L.; Lebrilla, C. B.; Fierro, F. A. FGF2 induces migration of human bone marrow stromal cells by increasing core fucosylations on N-Glycans of integrins. *Stem Cell Rep.* **2018**, 11

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

(2), 325−333. DOI:https://doi.org/10.1016/j.stemcr.2018.06.00

13. Agrawal, P.; Fontanals-Cirera, B.; Sokolova, E.; Jacob, S.; Vaiana, C. A; Argibay, D.; Davalos, V.; McDermott, M.; Nayak, S.; Darvishian, F.; Castillo, M.; Ueberheide, B.; Osman, I.; Fenyö, D.; Mahal, L. K; Hernando, E. A systems biology approach identifies FUT8 as a driver of melanoma metastasis. *Cancer Cell* **2017**, 31, 804–819. DOI: 10.1016/j.ccell.2017.05.007
14. Ailor, E.; Takahashi, N.; Tsukamoto, Y.; Masuda, K.; Rahman, B. A.; Jarvis, D. L.; Lee, Y. C.; Betenbaugh, M. J. N-glycan patterns of human transferrin produced in Trichoplusia ni insect cells: effects of mammalian galactosyltransferase. *Glycobiology* **2000**, 10 (8), 837−847. DOI: https://doi.org/10.1093/glycob/10.8.837
15. Haslam, S. M.; Coles, G. C.; Morris, H. R.; Dell, A. Structural characterization of the N-glycans of Dictyocaulus viviparus: discovery of the Lewis x structure in a nematode. *Glycobiology* **2000**, 10 (2), 223−229. DOI: https://doi.org/10.1093/glycob/10.2.223
16. Suzuki, N.; Abe, T.; Hanzawa,K.; Natsuka, S. Toward robust N-glycomics of various tissue samples that may contain glycans with unknown or unexpected structures. *Sci. Rep.* **2021,** 11, 6334. DOI: 10.1038/s41598-021-84668-x.
17. Dell, A.; Morris, H. R. Glycoprotein structure determination by mass spectrometry. *Science* **2001**, 291, 2351–2356. DOI: 10.1126/science.1058890
18. Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V. Congruent strategies for carbohydrate sequencing. 1. Mining structural details by MSn. *Anal. Chem.* **2005**, 77, 6250–6262. DOI: https://doi.org/10.1021/ac050724z
19. Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J. Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Anal. Chem.* **2014**, 86, 196–212. DOI: https://doi.org/10.1021/ac403969n

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

20. Wuhrer, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M. Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antenna gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* **2006**, 20 (11), 1747−1754. DOI: https://doi.org/10.1002/rcm.2509
21. Ernst, B.; Muller, D. R.; Richter, W. J. False sugar sequence ions in electrospray tandem mass spectrometry of underivatized sialyl-Lewis-type oligosaccharides. *Int. J. Mass Spectrom. Ion Processes* **1997**, 160 (1−3), 283−290. DOI: https://doi.org/10.1016/S0168-1176(96)04487-4
22. Wuhrer, M.; Deelder, A. M.; Yuri, E. M. B. Mass Spectrometric glycan rearrangements. *Mass. Spectrom. Rev.* **2011**, 4, 664−680. DOI: 10.1002/mas.20337.
23. Pang, P.-C.; Chiu, P. C. N.; Lee, C.-L.; Chang, L.-Y.; Panico, M.; Morris, H. R.; Haslam, S. M.; Khoo, K.-H.; Clark, G. F.; Yeung, W. S. B.; Dell, A. Human sperm binding is mediated by the sialyl-Lewis x oligosaccharide on the Zona Pellucida. *Science* **2011**, 333 (6050), 1761−1764. DOI: 10.1126/science.120743
24. Klapoetke, S.; Zhang, J.; Becht, S.; Gu, X.; Ding, X. J. The evaluation of a novel approach for the profiling and identification of N-linked glycan with a procainamide tag by HPLC with fluorescent and mass spectrometric detection. *Phantenna. Biomed. Anal.* **2010**, 53 (3), 315−324. DOI: https://doi.org/10.1016/j.jpba.2010.03.045
25. Nwosu, C.; Yau, H. K.; Becht, S. Assignment of core versus antenna fucosylation types in protein N-glycosylation via procainamide labeling and tandem mass spectrometry. *Anal. Chem.* **2015**, 87 (12), 5905−5913. DOI: https://doi.org/10.1021/ac5040743
26. Ács, A.; Ozohanics, O.; Vékey, K.; Drahos, L.; Turiák, L. Distinguishing core and antenna fucosylated glycopeptides based on low-energy tandem mass spectra.

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

Anal. Chem. **2018**, 90, 12776−12782. DOI: 10.1021/acs.analchem.8b03140

27. Lattová, E.; Skřičková, J.; Zdráhal, Z. Applicability of phenylhydrazine labeling for structural studies of fucosylated N-glycans. *Anal. Chem.* **2019**, 91(13), 7985−7990. DOI: https://doi.org/10.1021/acs.analchem.9b01321
28. Mancera-Arteu, M.; Giménez, E.; Barbosa, J.; Peracaula, R.; Sanz-Nebot, V. Zwitterionic-hydrophilic interaction capillary liquid chromatography coupled to tandem mass spectrometry for the characterization of human alpha-acid-glycoprotein N-glycan isomers. *Anal. Chim. Acta* **2017**, 991, 76−88. DOI: https://doi.org/10.1016/j.aca.2017.07.068
29. Harvey, D. J.; Struwe, W. B. Structural studies of fucosylated *N*-glycans by ion mobility mass spectrometry and collision-induced fragmentation of negative ions. *J. Am. Soc. Mass. Spectrom.* **2018**, 29, 1179-1193. DOI: 0.1007/s13361-018-1950-x.
30. H. Hwang *et al.* Machine learning classifies core and outer fucosylation of *N*-glycoproteins using mass spectrometry. *Sci. Rep.* **2020**, 10: 318. DOI: 10.1038/s41598-019-57274-1.
31. Hsu, C. H.; Liew, C. Y.; Huang, S. P.; Tsai, S. T.; Ni, C. K. Simple method for de novo structural determination of underivatized glucose oligosaccharides. *Sci. Rep.*, **2018**, 8, 5562. DOI: 10.1038/s41598-018-23903-4.
32. Tsai, S. T.; Liew C. Y.; Hsu C.; Huang S. P.; Weng W. C.; Kuo, Y. H.; Ni, C. K. Automatic full glycan structural determination through logically derived sequence tandem mass spectrometry. *ChemBioChem*, **2019**, 20, 2351–2359. DOI: 10.1002/cbic.201900228.
33. Liew CY, Chan CK, Huang SP, Cheng YT, Tsai ST, Hsu HC, Wang CC, Ni CK. De novo structural determination of oligosaccharide isomers in glycosphingolipids using logically derived sequence tandem mass spectrometry. *Analyst* **2021**; 146

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

(23), 7345-7357. DOI: 10.1039/D1AN01448J

34. Liew, C. Y.; Luo, H. S.; Yang, T. Y.; Hung, A. T.; Magoling, B. J. A.; Lai, C. P. K., Ni, C. K. Identification of the high mannose *N*-glycan isomers undescribed by the conventional multicellular biosynthetic pathways. *Anal. Chem.,* **2023,** 95(23), 8789-8797. DOI: 10.1021/acs.analchem.2c05599.
35. Liew, C. Y.; Yen, C. C.; Chen, J. L.; Tsai, S. T.; Pawar, S.; Wu, C. Y.; Ni, C. K. Structural identification of N-glycan isomers using logically derived sequence tandem mass spectrometry. *Commun. Chem.* **2021**, 4(92), 1–11. DOI: 10.1038/s42004-021-00532-z.
36. Liew, C. Y.; Luo, H. S.; Yang, T. Y.; Hung, A. T.; Magoling, B. J. A.; Lai, C. P. K., Ni, C. K. Identification of the high mannose *N*-glycan isomers undescribed by the conventional multicellular biosynthetic pathways. *Anal. Chem.,* **2023,** 95(23), 8789-8797. DOI: 10.1021/acs.analchem.2c05599.
37. Liew, C. Y.; Chen, J. L.; Ni, C. K. Electrospray ionization in-source decay of *N*-glycans and the effects on *N*-glycan structural identification. *Rapid Commun. Mass Spectrom.* **2022**, 36, e9352. DOI: 10.1002/rcm.9352.
38. Chen, J. L.; Nguan, H. S.; Hsu,P. J.; Tsai, S. T.; Liew, C. Y.; Kuo, J. L.; Hu, W. P.; Ni, C. K. Collision-induced dissociation of sodiated glucose and identification of anomeric configuration. *Phys. Chem. Chem. Phys.* **2017**, 19, 15454–15462. DOI: 10.1039/c7cp02393f.
39. Huynh, H. T.; Phan, H. T.; Hsu, P. J.; Chen, J. L.; Nguan, H. S.; Tsai, S. T.; Roongcharoen, T; Liew, C. Y.; Ni, C. K.; Kuo, J. L. Collision-induced dissociation of sodiated glucose, galactose, and mannose, and the identification of anomeric configuration. *Phys. Chem. Chem. Phys.* **2018**, 20, 19614–19624. DOI: 10.1039/c8cp03753a.
40. Chiu, C. C.; Tsai, S. T.; Hsu P. J.; Huynh, H. T.; Chen, J. L.; Phan, H. T.; Huang,

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0

S. P.; Lin, H. Y.; Kuo, J. L.; Ni, C. K. Unexpected dissociation mechanism of sodiated N-acetylglucosamine and N-acetylgalactosamine. *J. Phys. Chem. A.* **2019,** 123, 3441–3453. DOI: 10.1021/acs.jpca.9b00934.

41. Nguan, H. S.; Tsai, S. T.; Liew, C. Y.; Reddy, N. S.; Hung, S. C.; C. Y.; Ni, C. K. The collision-induced dissociation mechanism of sodiated Hex-HexNAc disaccharides. *Phys. Chem. Chem. Phys.* **2023**, 25, 22179–22194. DOI: 10.1039/D3CP02530F.
42. Hsu, H.C.; Tsai, M. T.; Dyakov, Y. A.; Ni, C. K. Energy transfer of highly vibrationally excited molecules studied by crossed molecular beam/time-sliced velocity map ion imaging. *Int. Rev. Phys. Chem.* **2012**, 31, 201−233. DOI: 10.1080/0144235X.2012.673282
43. Liew, C. Y.; Yen C. C., Chen J. L.; Tsai S. T.; Pawar S.; Wu C. Y.; Ni C. K. Structural identification of N-glycan isomers using logically derived sequence tandem mass spectrometry. *Commun. Chem*. **2021**, 4, 92. DOI: https://doi.org/10.1038/s42004-021-00532-z
44. Liew, C. Y.; Chan, C. K.; Huang, S. P.; Cheng, Y. T.; Hsu, H. C.; Wang, C. C.; Ni C. K. De novo structural determination of oligosaccharide isomers in glycosphingolipids using logically derived sequence tandem mass spectrometry. *Analyst* **2021**, 146, 7345-7357. DOI: 10.1039/d1an01448j.
45. Liew, C. Y.; Chen, J. L.; Ni, C. K. Identification of side-reaction products generated during the ammonia-catalyzed release of N-glycans. *Carbohydr. Res.* **2022**, 552, 108686. DOI: 10.1016/j.carres.2022.108686
46. Stanley, P., Moremen, K. W.; Lewis, N. E.; Taniguchi, N., Aebi, M. *N*-glycans. In *Essentials of Glycobiology*, 4th edition, edited by Varki A. et al. Cold Spring Harbor Laboratory Press, New York (2022).

https://doi.org/10.26434/chemrxiv-2024-hkj8k **ORCID:** https://orcid.org/0000-0001-6503-8905 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC 4.0