Supplementary Information (SI) for Analyst.
This journal is © The Royal Society of Chemistry 2024

Supplementary Information

Structural determination of fructooligosaccharides and raffinose family oligosaccharides using logically derived sequence tandem mass spectrometry

Chia Yen Liew[abc], Wun-Long Li[ad], and Chi-Kung Ni*[ace]

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

[b] International Graduate Program of Molecular Science and Technology, National Taiwan University (NTU-MST), Taipei 10617, Taiwan

[c] Molecular Science and Technology (MST), Taiwan International Graduate Program (TIGP), Academia Sinica, Taipei 10617, Taiwan

[d] Department of Chemistry, National Taiwan Normal University, Taipei, 11677, Taiwan

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

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

(A) Sources of materials

Table S1: sources of materials

Turanose	α-Glc-(1→3)-Fru	Carbosynth
Maltulose	α-Glc-(1→4)-Fru	Carbosynth
Cellobiulose	β-Glc-(1→4)-Fru	LC Scientific Inc. (50 Viceroy Road, Unit 7. Concord, Ontario, L4K 3A7 Canada)
Isomaltulose	α-Glc-(1→6)-Fru	Carbosynth
Levanbiose	β-Fru-(2→6)-Fru	Megazyme (Bray, Ireland)
Trehalulose	α-Glc-(1→1)-Fru	BOC science (Ramsey Road, Shirley, NY 11967, USA)
Inulobiose	β-Fru-(2→1)-Fru	Extract from onion
Leucrose	α-Glc-(1→5)-Fru	BOC science (Ramsey Road, Shirley, NY 11967, USA)
Inulotriose	β-Fru-(2→1)-Fru-β- (2→1)-Fru	Megazyme (Bray, Ireland)
Stachyose hydrate	α-Gal-(1→6)-α-Gal-(1→6)-α-Glc-(1→2)-β-Fru	Sigma-Aldrich (Burlington, Massachusetts, United States)

(B) CID spectra of glucose disaccharides with various linkages

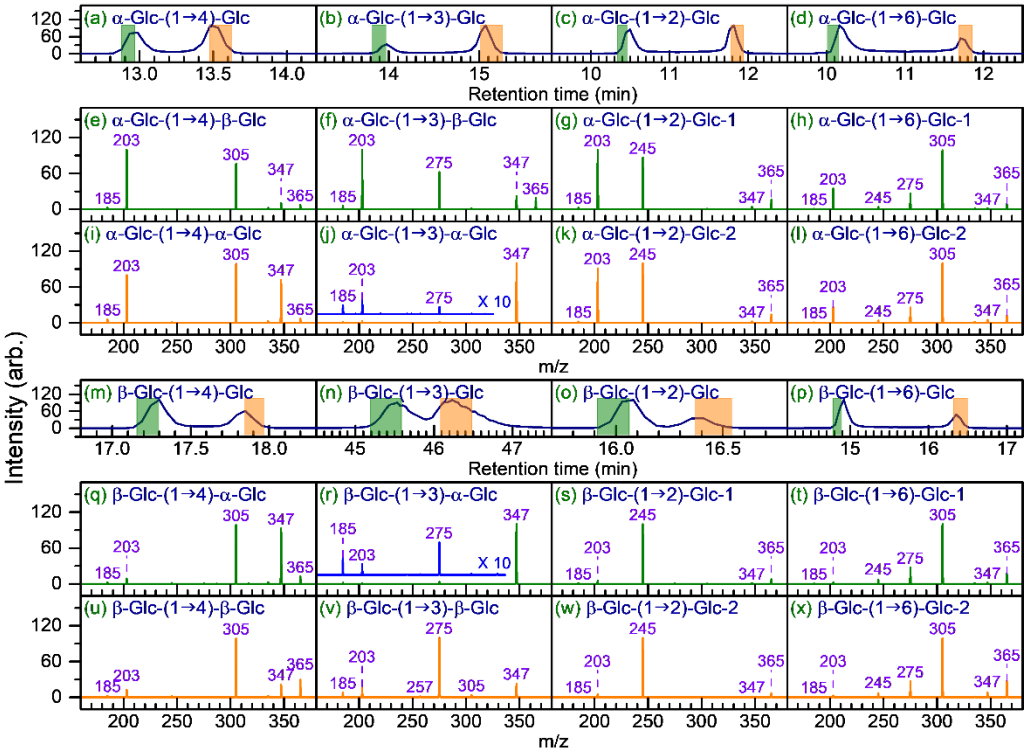

Figure S1. CID spectra of glucose disaccharide with various linkages.

(C) Dissociation mechanism of fructose at the reducing end connecting to other sugars through more than one glycosidic bond (branched linkages)

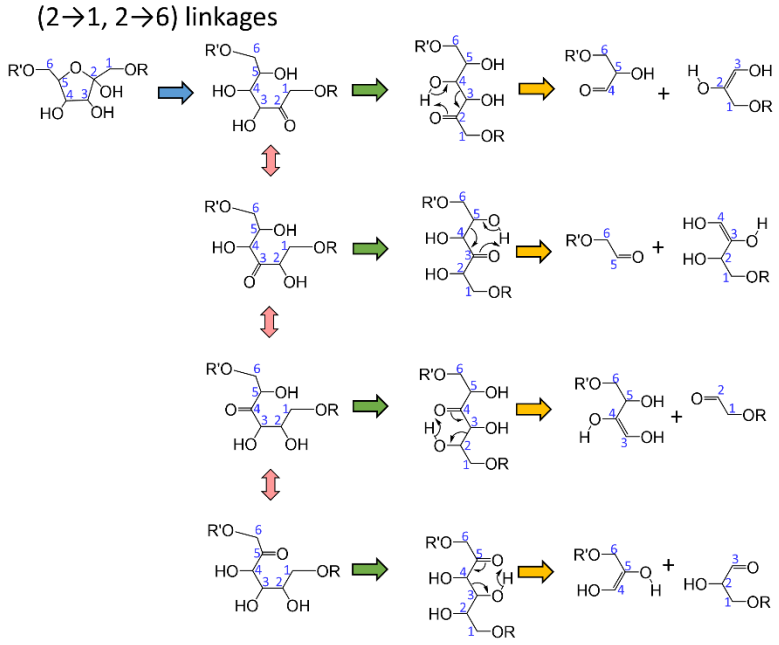

Figure S2. Dissociation mechanism of fructose at the reducing end with (2→1, 2→6) linkages.

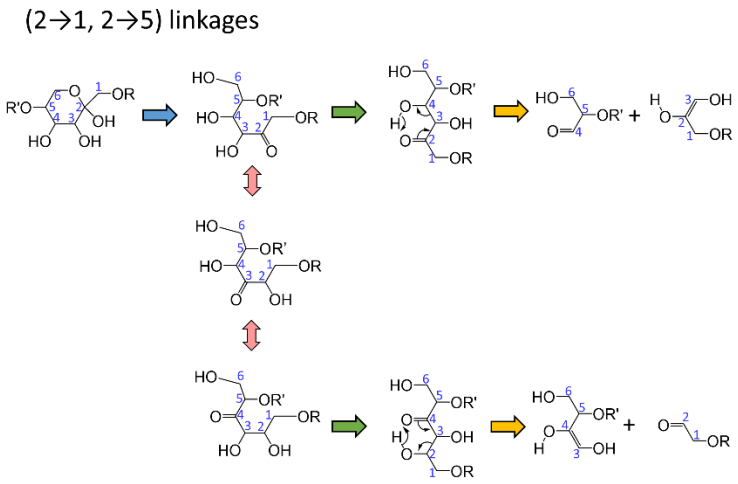

Figure S3. Dissociation mechanism of fructose at the reducing end with (2→1, 2→5) linkages.

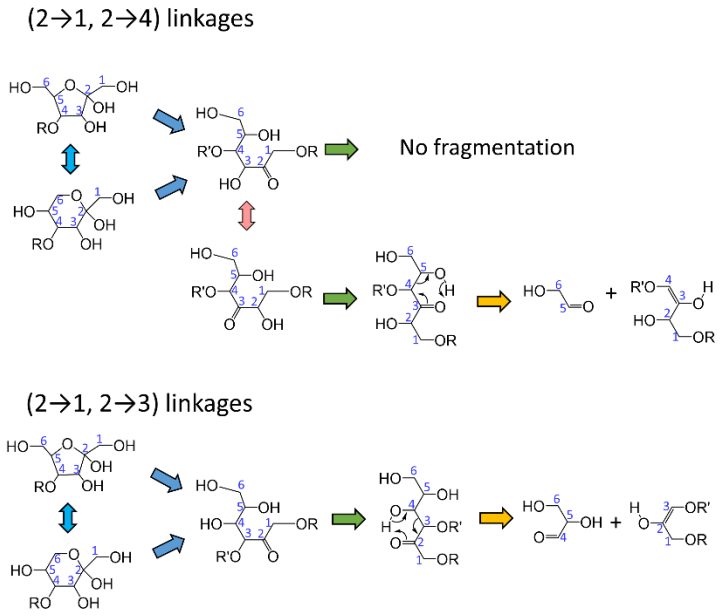

Figure S4. Dissociation mechanism of fructose at the reducing end with (2→1, 2→4) linkages and (2→1, 2→3) linkages,.

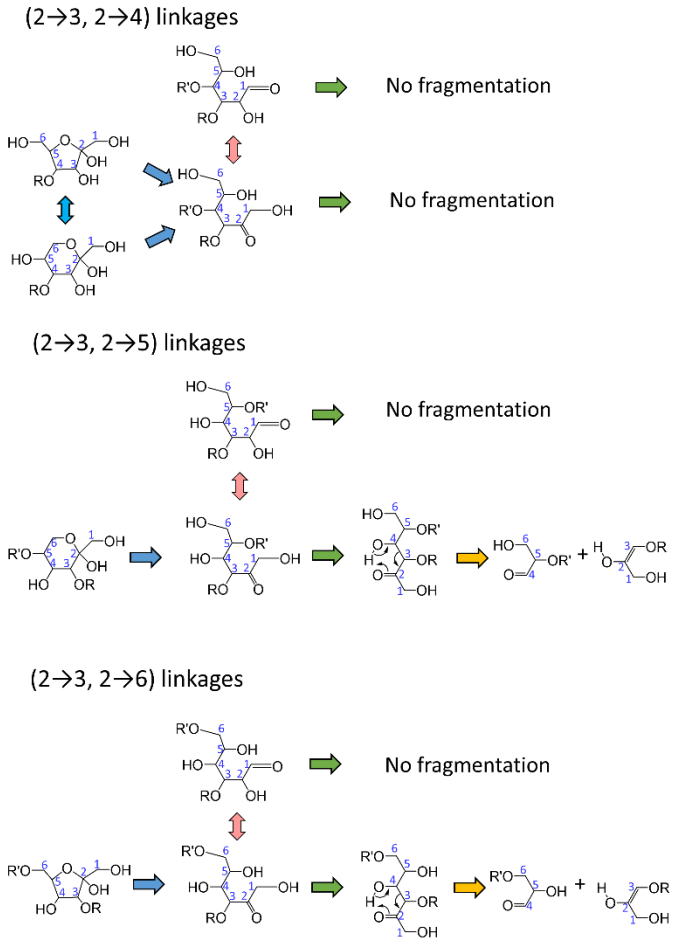

Figure S5. Dissociation mechanism of fructose at the reducing end with (2→3, 2→4) linkages, (2→3, 2→5) linkages, (2→3, 2→6) linkages.

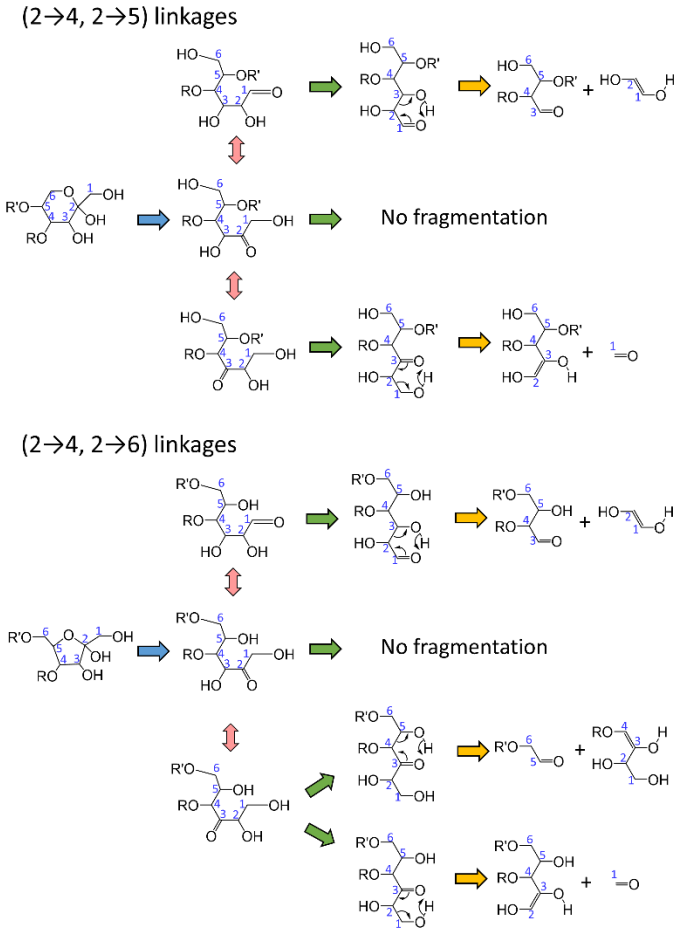

Figure S6. Dissociation mechanism of fructose at the reducing end with (2→4, 2→5) linkages, (2→4, 2→6) linkages.

(D) Step-by-step protocols for trisaccharide and tetrasaccharide structural determination of fructooligosaccharides and raffinose family oligosaccharides

For the method we used, it requires the resonance excitation of sodium ion adducts in an ion trap with low collision energy. Here the collision gas is He and the collision energy is just enough to undergo cross-ring dissociation and dehydration at reducing end, and the glycosidic bond cleavages. We used LTQ XL or Velos linear ion trap mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) in this study. This does not necessary mean they are the most ideal instruments for LODES/MS^n but these instruments are available in our laboratory and they satisfy the requirements for LODES/MS^n. If different instruments are used, one must consider

the differences in the relative ion intensities of the CID spectra due to the differences in the mass-dependent trapping and detection efficiencies, and in particular the difference in collision energy. The definition of collision energy varies between instruments and manufactures, there is not always a calibration of collision energies between different instruments. In this study, we set the collision energy as 25-30% normalized collision energy and activation time 30 ms for disaccharides, and 30%-40% normalized collision energy and activation time 30 ms for large oligosaccharides. At this moment, we have yet to have the opportunity to compare the performance of other instruments. Different instruments may result in different relative intensities, thus our method cannot be used for structural determination.

The structural determination is based on the dissociation mechanisms of sodium ion adducts in the linear ion trap with resonance excitation by collision with He gas. The dissociation mechanisms are summarized as follows. Notably, different ionic species, such as negative ions and protonated ion, have different dissociation mechanisms and they are not applicable to the method described bellows.

- Dehydration occurs mainly at the reducing end of oligosaccharides.
- Cross-ring dissociation mainly occurs at the reducing end of oligosaccharides and follows the rules of retro-aldol reaction.
- Glycosidic bond cleavage occurs at any positons.

The step-by-step protocols for structural determination of fructooligosaccharides and raffinose family oligosaccharides according to the aforementioned dissociation mechanisms are illustrated bellows. In the protocols, the monosaccharide and disaccharide database are used for comparison. The CID spectra in Figure 1 are the monosaccharide database, and the CID spectra in Figure 10 are the disaccharide database.

Trisaccharides

- All the trisaccharides can be divided into three groups, as illustrated in Figure S7. The first step is to identify which group the trisaccharide belongs to. The CID sequence 527→fragments is used to differentiate group 3 from groups 1 and 2. Isomers in group 3 are non-reducing trisaccharides, thus they do not undergo cross-ring dissociation and dehydration. Only groups 1 and 2 generate dehydration fragments m/z 509, and cross-ring fragments m/z 437 and/or 275.

- The cross-ring dissociation patterns in the CID sequence, 527→fragments, are used to differentiate isomers in groups 1 and 2. The structural changes along this CID sequence are illustrated in MS^2→ in Figure S7. According to the retro-aldol reaction which produces cross-ring fragments m/z 437 and 275, group 1 only generates fragment m/z 275 (or the intensity of fragment m/z 275 is much larger than that of m/z 437, where fragment m/z 437 is not produced according to the rule of retro-aldol reaction, thus the intensity is low). In contrast, group 2 only generates fragment m/z 437 (or the intensity of fragment m/z 437 is much larger than that of m/z 275, where fragment m/z 275 is not produced according to the rule of retro-aldol reaction or it requires non-resonance excitation or secondary dissociation, thus the intensity is low). After the differentiation of trisaccharides into groups, the next step is to identify the isomer in each group.

Group 1

- Group 1 only has only one isomer, no more identification is required.

Group 2

- To differentiate isomers in group 2, the CID spectrum, obtained from CID sequence 527→437→275→fragments, is used to determine the linkage of the

glycosidic bond between sugars 1 and 2 by comparing to the CID spectra in disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(4) \rightarrow$ in Figure S7.

- To further differentiate isomers in group 2, the CID spectrum from CID sequence 527→437→365→275→fragments is used to determine the linkage of the glycosidic bond between sugars 2 and 3 by comparing to the CID spectra in disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(5) \rightarrow MS^5(3) \rightarrow$ in Figure S7. This CID sequence start from the cross-ring dissociation, which is very reliable and is selected as the first priority to determine the linkage of the glycosidic bond between sugars 2 and 3
- For group 2, the CID spectrum, obtained from CID sequence 527→509→365→275→fragments, can be used to further crosscheck the linkage of the glycosidic bond between sugars 2 and 3 by comparing to the CID spectra in disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(5) \rightarrow MS^4(6) \rightarrow MS^5(4) \rightarrow$ in Figure S7. This CID sequence starts from dehydration, which is less reliable that the CID sequence starting from cross-ring dissociation. This is because the branching ratio of dehydration not occurs at the O1 atom of the sugar at the reducing end is a little higher than that of cross-ring dissociation. In the fructooligosaccharides and raffinose family oligosaccharides we have studies, the CID sequence starting from dehydration provides the correct structures, but the inaccurate structural assignments have been found in other oligosaccharides. This phenomenon is applicable to the procedure bellows but the description is not repeated in the procedure bellows.

Group 3

- To differentiate isomers in group 3, the CID spectrum, obtained from CID sequence 527→365→fragments, is used differentiate isomers 3a and 3b from isomers 3c and 3d. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(2)$→ in Figure S7. The intensity of fragment m/z 275 is much larger than that of m/z 305 and 245 for isomers 3a and 3b, while the intensity ratio of fragments m/z 305, 275, and 245 for isomers 3c and 3d are 5: 3±1.5: 1.5±1.
- The CID spectrum, obtained from CID sequence 527→365→275→fragments, is used to distinguish isomers between 3a and 3b by comparing to the CID spectra in disaccharide database. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(2)$→$MS^4(1)$→ in Figure S7.
- The CID spectrum, obtained from CID sequence 527→203→fragments is used to differentiate the isomers between 3c and 3d by comparing to the CID spectra in monosaccharide database. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(1)$→ in Figure S7. The CID spectrum, obtained from the CID sequence 527→365→305→203→fragments, can be used to distinguish isomers between 3c and 3d by comparing to the CID spectra in monosaccharide database. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(2)$→$MS^4(2)$→ $MS^5(1)$→ in Figure S7. This CID sequence completely avoids the interference of secondary dissociation, but the signal may be small due to the multi-stage CID.

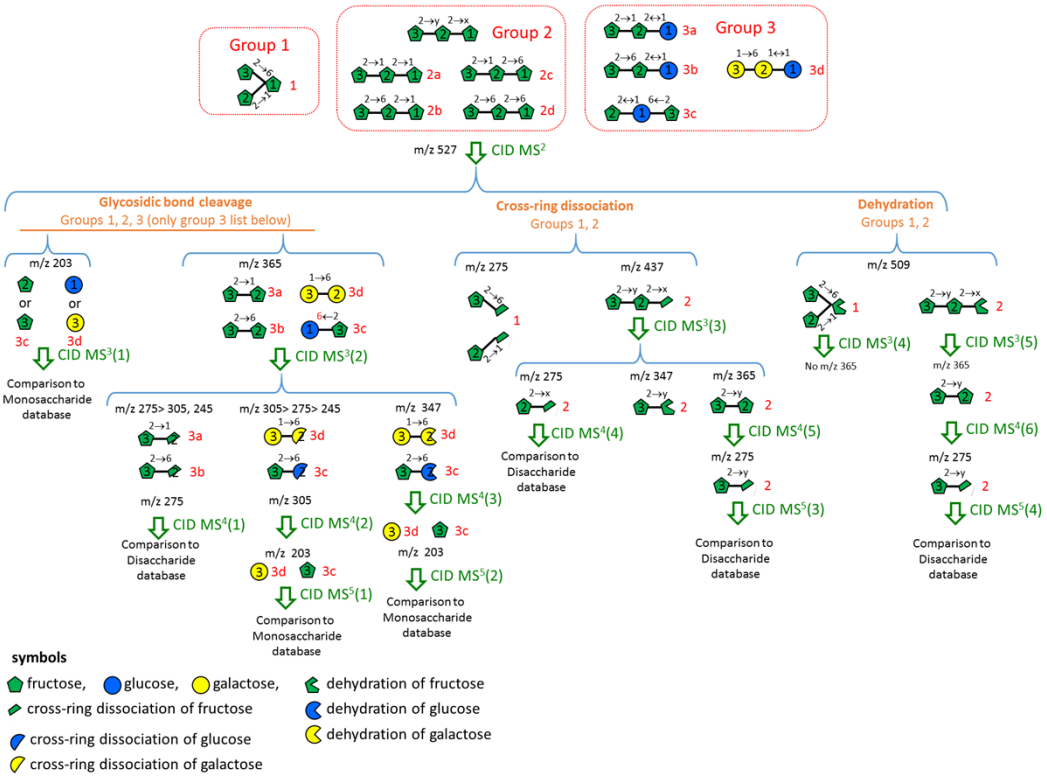

Figure S7. CID sequences and structural changes along the CID sequences for structural determination of trisaccharides.

Tetrasaccharides

- All tetrasaccharides can be divided into four groups, as illustrated in Figure S8. The CID spectrum through the CID sequence 689→fragments is used to classify the tetrasaccharides into different groups. Isomers in group 4 are non-reducing oligosaccharides and do not undergo dehydration and cross-ring dissociation. Only isomers in groups 1, 2, and 3 generate dehydration fragment m/z 671 and cross-ring fragments m/z 599 and 437. Among the cross-ring fragments m/z 599 and 437, isomers in groups 2 and 3 only generate fragment m/z 599 (or the intensity of m/z 599 is much larger than that of m/z 437), while isomers in group 1 only generate fragment m/z 437 (or the intensity of m/z 437 is much larger than that of m/z 599). The structural changes along the CID sequence are

illustrated in $MS^2\rightarrow$ in Figure S8.

- If no fragments m/z 671, 599, and 435 is found in CID spectrum through the CID sequence 689→fragments, the isomers belong to group 4.
- To differentiate groups 2 and 3, CID spectrum through the CID sequence 689→599→527→fragments is used. Isomers in group 2 only generate fragment m/z 275 (or the intensity of m/z 275 is much larger than that of m/z 437), while isomers in group 3 only generate fragment m/z 437 (or the intensity of m/z 437 is much larger than that of m/z 275). The structural changes along the CID sequence are illustrated in $MS^2\rightarrow$ $MS^3(3)\rightarrow$ $MS^4(6)\rightarrow$ in Figure S8.
- After identify which group the isomers belong to, the next step is to identify the isomer in each group.

Group 1

- The CID sequence, 689→437→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 1 and 2 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2\rightarrow$ $MS^3(2)\rightarrow$ $MS^4(1)\rightarrow$ in Figure S8.
- The CID sequence, 689→437→365→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 2 and 3 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2\rightarrow$ $MS^3(2)\rightarrow$ $MS^4(2)\rightarrow$ $MS^5(1)\rightarrow$ in Figure S8.
- The CID sequence, 689→671→365→275→fragments, can be used to crosscheck the linkage of the glycosidic bond between sugars 2 and 3 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2\rightarrow$ $MS^3(4)\rightarrow$ $MS^4(7)\rightarrow$ $MS^5(5)\rightarrow$ in Figure S8.

Group 2

- The CID sequence, 689→599→437→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 1 and 2 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in MS^2→ $MS^3(3)$→ $MS^4(4)$→ $MS^5(2)$→ in Figure S8.

Group 3

- The CID sequence, 689→599→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 1 and 2 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in MS^2→ $MS^3(3)$→ $MS^4(3)$→ in Figure S8.
- The CID sequence, 689→599→437→365→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 2 and 3 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(3)$→$MS^4(4)$→$MS^5(3)$→$MS^6(1)$→ in Figure S8.
- The CID sequence, 689→599→509→365→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 3 and 4 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in MS^2→$MS^3(3)$→$MS^4(5)$→$MS^5(4)$→$MS^6(2)$→ in Figure S8.

Group 4

- The isomers in group 4 can be divided into four subgroups, namely (4a, 4b, 4c, 4d), (4e, 4f, 4j), (4i), and (4g, 4h), according to the MS^3 cross-ring dissociation patterns. The CID spectrum through the CID sequence 689→527→fragments is used to differentiate isomers 4g and 4h from the other isomers. Isomers 4g and 4h do not generate cross-ring fragments 275, 437, or 467. Among these cross-

ring fragments, isomer 4i generates fragment m/z 275 (or the intensity of m/z 275 is much larger than that of m/z 467, 437, and 407). Isomers 4a, 4b, 4c, and 4d generate fragments m/z 467, 437, and 407, but the intensity of m/z 437 is much larger than those of m/z 467, 407 and 275. In contrast, the intensity ratio of fragments m/z 467, 437, and 407 are 5: 3±1.5:1.5±1 for isomers 4e, 4f, and 4j. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow$ in Figure S9.

- To differentiate isomers 4a, 4b, 4c, and 4d, the CID sequence, 689→527→437→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 2 and 3 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(3) \rightarrow MS^5(2) \rightarrow$in Figure S9.
- To further differentiate isomers 4a, 4b, 4c, and 4d,The CID sequence, 689→527→437→365→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 3 and 4 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(3) \rightarrow MS^5(3) \rightarrow MS^6(1) \rightarrow$ in Figure S9.
- To differentiate isomers 4e, 4f, and 4j, the CID sequence, 689→527→467→365→fragments, is used. The intensity of cross-ring fragment m/z 275 from isomers 4e and 4f is much larger than those of m/z 305 and 245, while intensity ratio of fragments m/z 305, 275, and 245 are 5: 3±1.5:1.5±1 for isomers 4j. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(4) \rightarrow MS^5(4) \rightarrow$ in Figure S9.
- To distinguish 4e and 4f, the CID sequence, 689→527→437→365→275→fragments, is used to determine the linkage of the glycosidic bond between sugars 3 and 4 by comparing to disaccharide database.

The CID sequence, 689→509→365→275→fragments, can be used to crosscheck the linkage of the glycosidic bond between sugars 3 and 4 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(3) \rightarrow MS^4(4) \rightarrow MS^5(4) \rightarrow MS^6(2) \rightarrow$ in Figure S9.

- To distinguish 4g and 4h, the CID sequences, 689→365→275→fragments is used to determine the linkage of the glycosidic bond between sugars 3 and 4 by comparing to disaccharide database. The structural changes along the CID sequence are illustrated in $MS^2 \rightarrow MS^3(1) \rightarrow MS^4(1) \rightarrow$ in Figure S3. Notably, the contribution from Fru-(2→6)-Glc also contribute to the spectra, thus interference from Fru-(2→6)-Glc has to be subtracted before comparing to the database. At this moment, we do not have the spectra of Fru-(2→6)-Glc for subtraction.

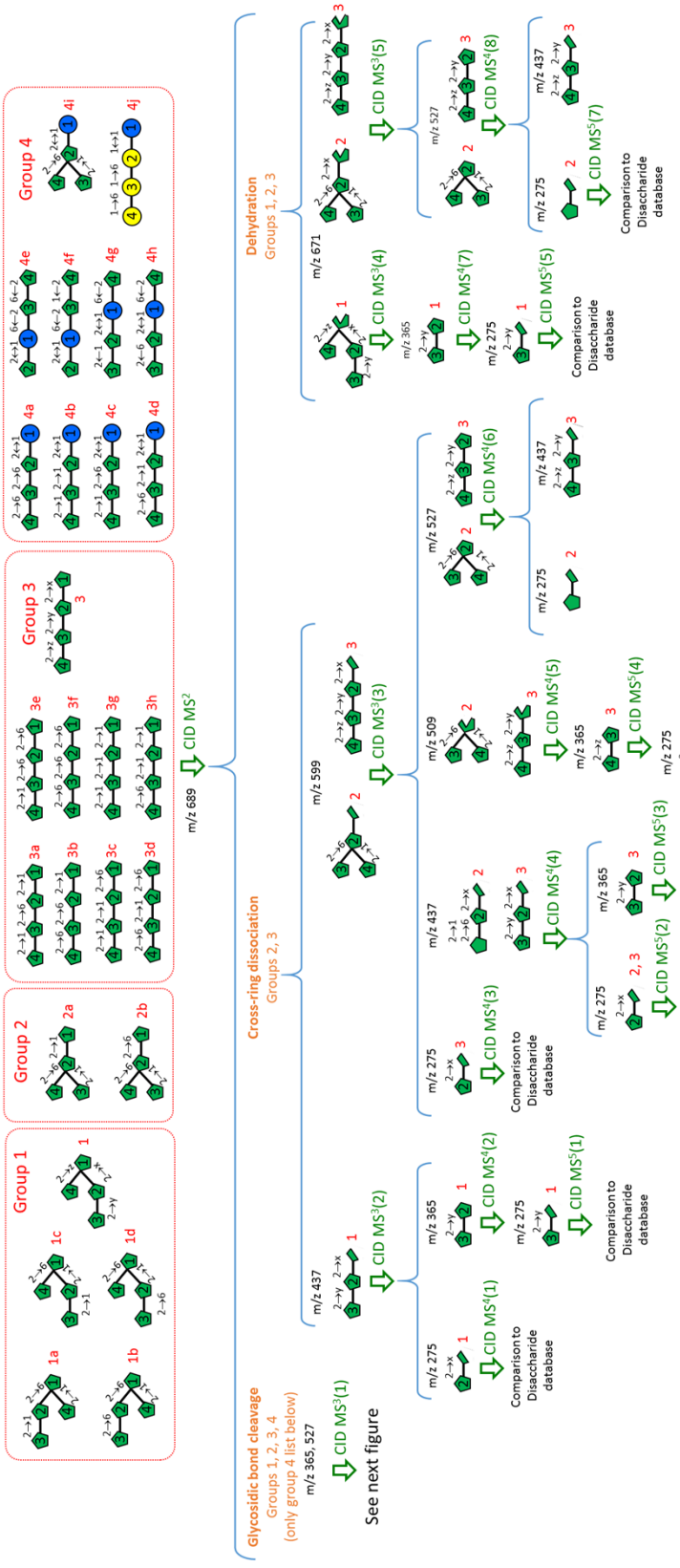

Figure S8. CID sequences and structural changes along the CID sequences for structural determination of tetrasaccharides, part 1.

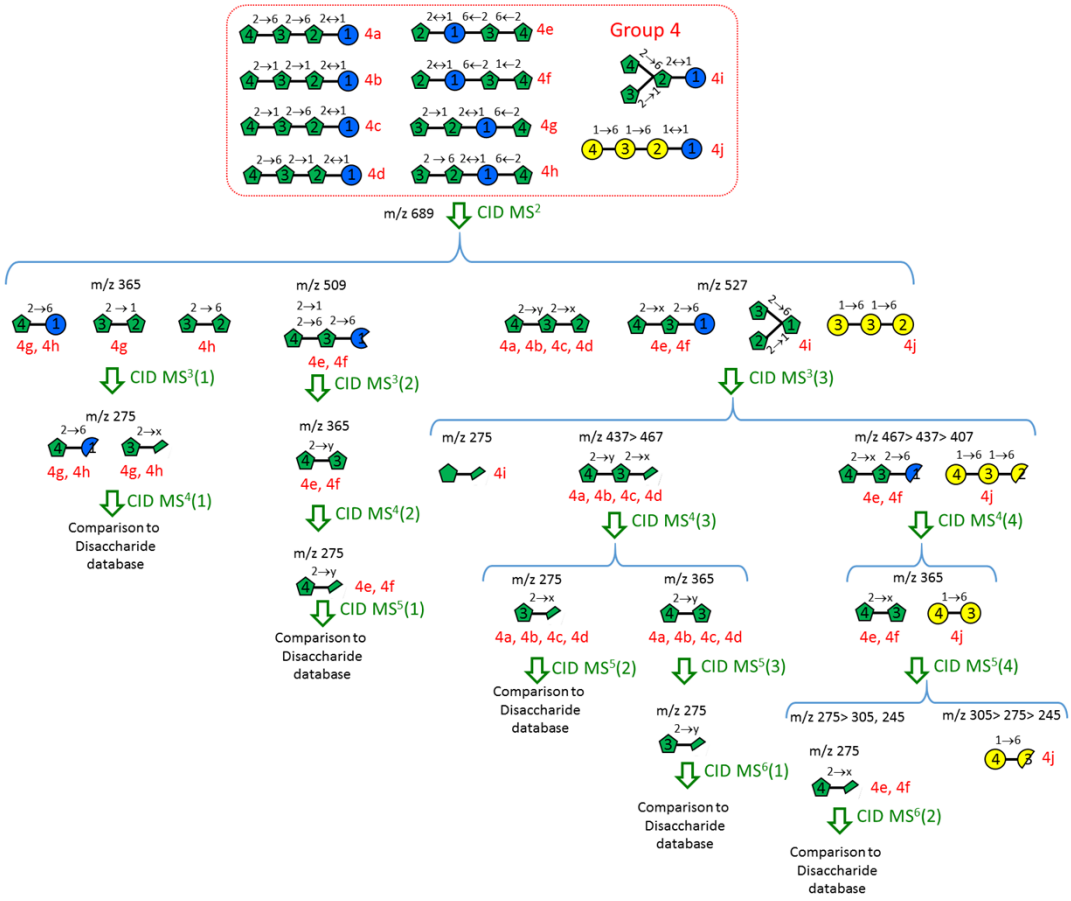

Figure S9. CID sequences and structural changes along the CID sequences for structural determination of tetrasaccharides, part 2.