Manual : Sugar Fragmentation (Version 0.5 : 23 March 2002)

The **purpose of this application** is to calculate and display the main **fragments** (B- and C-, Z- and Y-, A- and X-ions) **of oligosaccharides** that should occur in **MS-spectra**. The extended ASCII character set as recommended by IUPAC is used to input the sequence of complex oligosaccharides.

This web application has been developed by the molecular modeling group of the central spectroscopic department of the German Cancer Research Center (Deutsche Krebsforschungszentrum). Should you find mistakes within the program, or should you have any suggestions about possible changes please contact:

Klaus Lohmann (k.lohmann@dkfz.de) or Willi von der Lieth (w.vonderlieth@dkfz.de)

Other glyco-related web application can be found at:

http://www.dkfz.de/spec/ and http://www.glycosciences.de

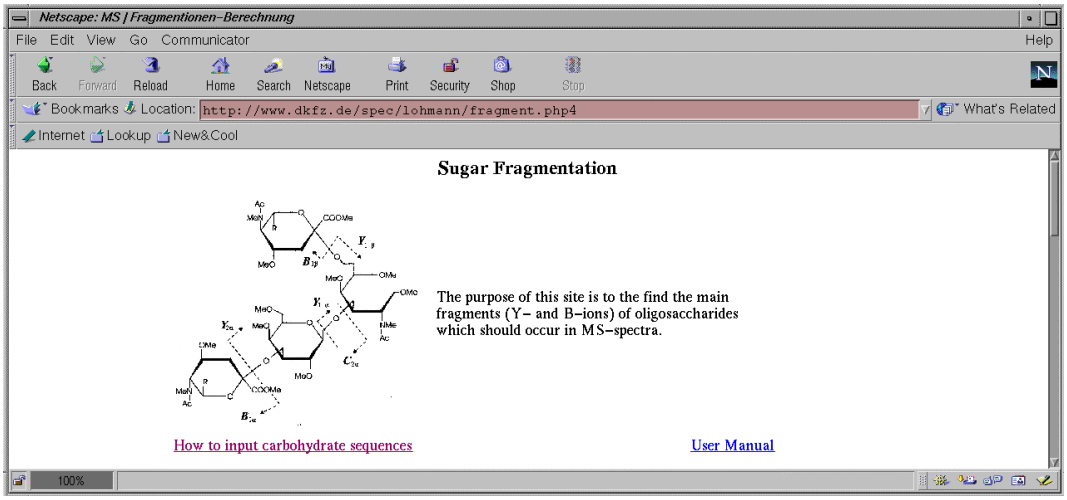
This work was supported by the Deutsche Forschungsgemeinschaft.

1. Start your browser (Netscape or Internet Explorer) and activate the following URL:

http://www.dkfz.de/spec/lohmann/fragment.php4

You will get the following page:

(This site is currently under continuous development. Therefore, its appearance may change from time to time. The figures given below are screenshots of Netscape. If you are using a different browser, like Internet Explorer, some details of the display may be different. The application has been tested for both common browsers. A non-proportional font - like COURIER - is required for the structure input. Please check the settings of your browser if you have problems.)

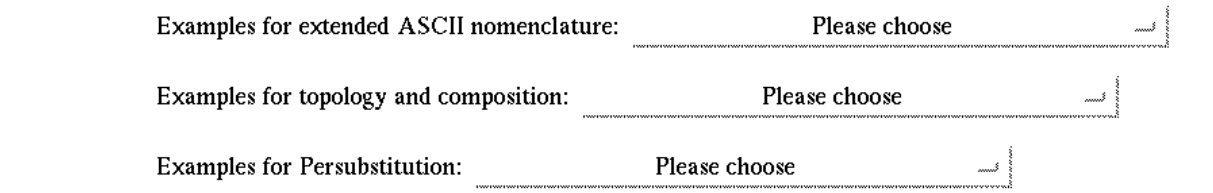


2. Structure Input Window

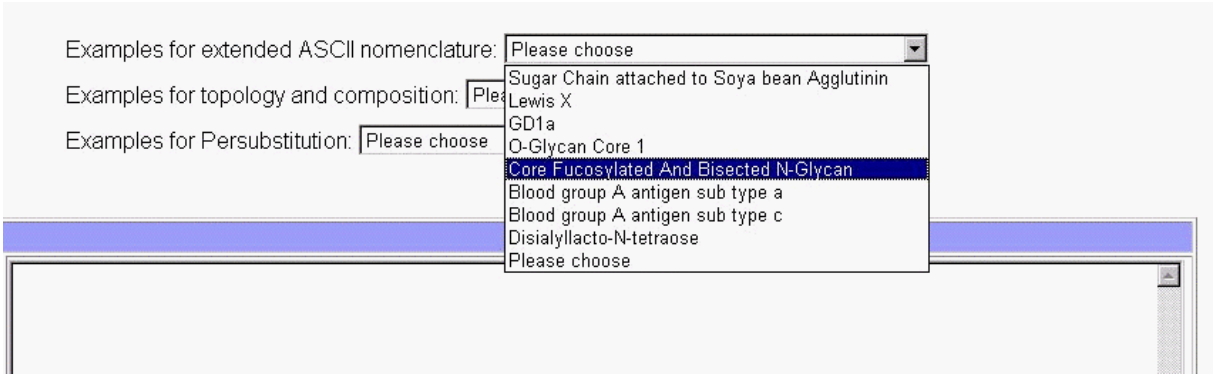
This is a simple ASCII input window where you can type in the sequence of your carbohydrate structure using the extended description. The **extended ASCII character set** (see 2.3) as recommended by IUPAC is used to input the sequence of complex oligosaccharides. However, some additional input rules have to be fulfilled (see 2.3.2.). If only the **topology and composition** of the oligosaccharide is known, a simpler way to input carbohydrate sequences is possible (see 2.1.2). Since the hydroxyl groups of synthetic carbohydrates are often protected we have included a way to indicate if **sugar residue are persubstituted [?]** (see 2.1.3.).

2.1. Examples of valid input structures.

Activating the examples button above the structure input window you can chose between the three example sets:

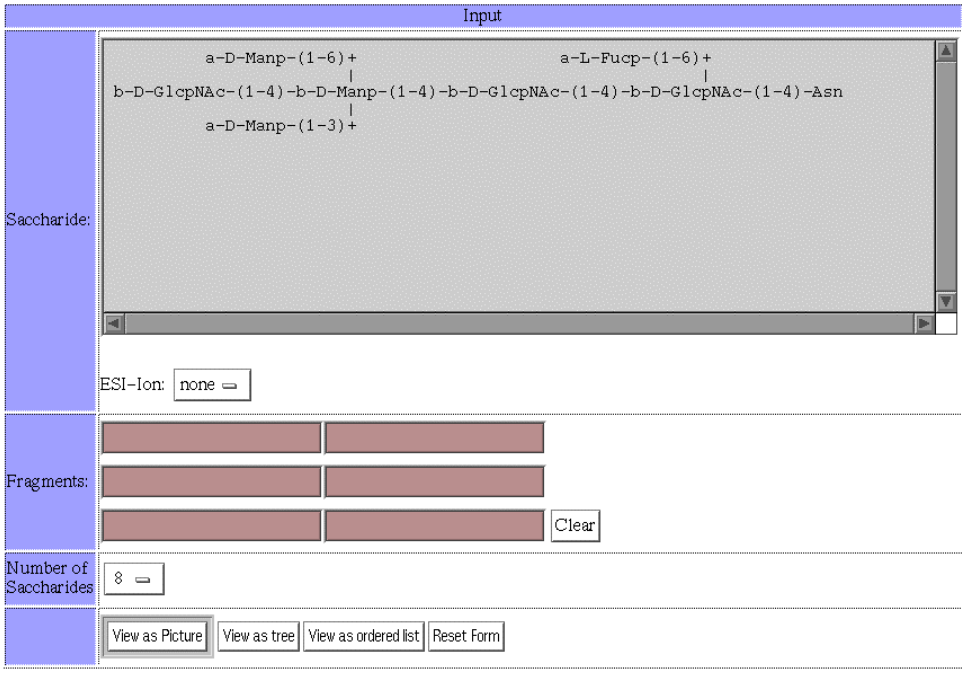


A few typical glycan structures are listed when you activate the '*please choose*' button.

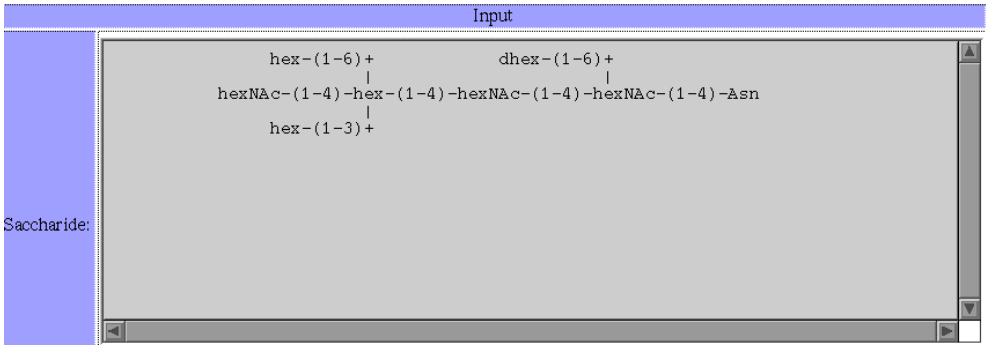


Selecting one of these structures will paste the glycan structure into the input window. The structure of a core fucosylated and bisected N-Glycan will be used throughout this manual as an example.

2.1.1. Extended ASCII character set

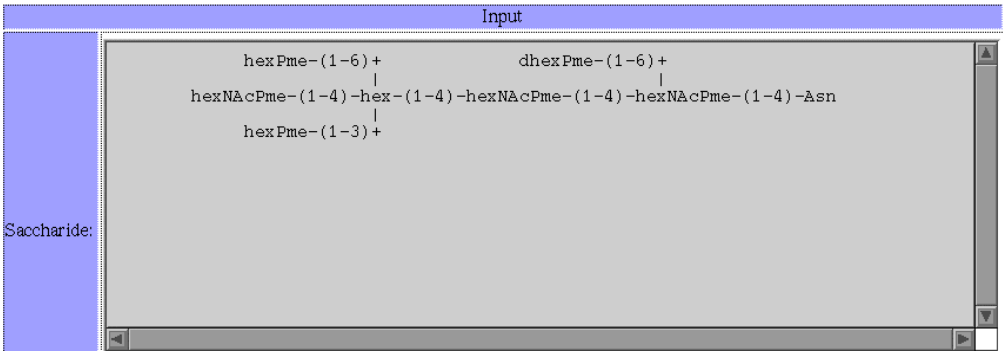


2.1.2. Topology and Composition



Currently supported compositions: Hex, dHex, HexNac, hexAc. Compositions and extended nomenclature can be mixed.

2.1.3. Persubstituted Sugar Residue



Currently supported persubstitutions: PMe and PAc. Persubstitutions are only possible for sugar compositions and cannot be used in combination with the extended nomenclature.

2.2. Copy/Paste your sequences from other sources

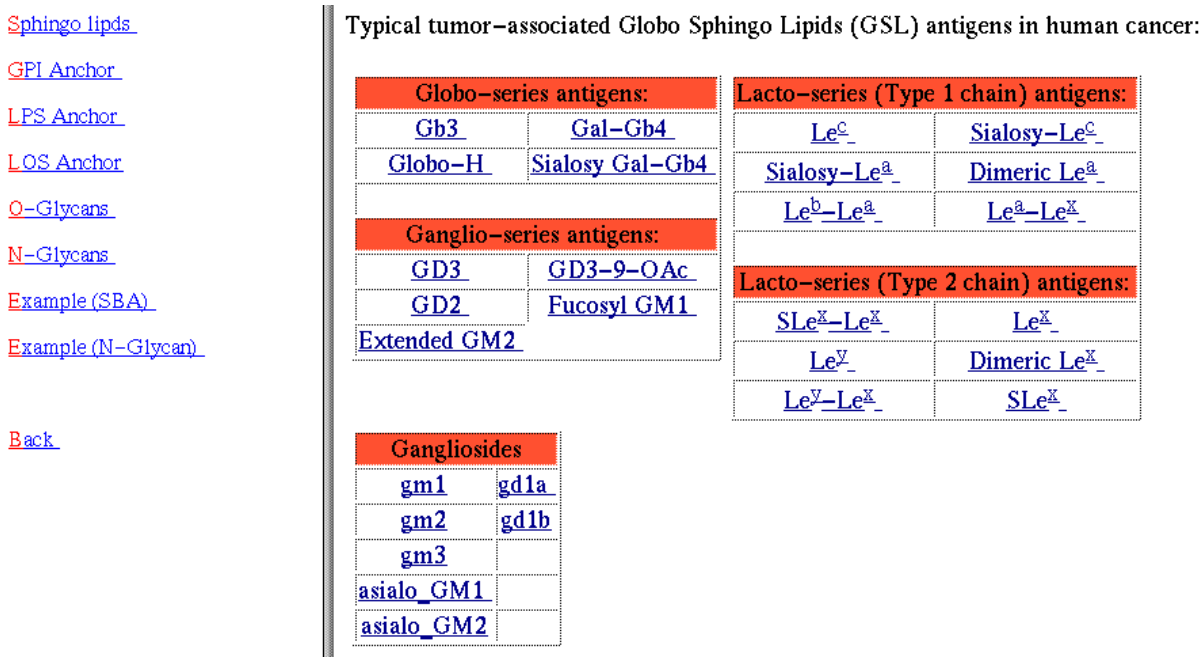
Alternatively you can copy/paste your sequences from predefined templates as discussed below or copy/paste sequences using your favourite editor. In case you copy your sequence from a word processing software like MS-WORD we strongly recommend you create your structure using a non-proportional font like Courier .

2.2.1. Predefined templates for various types of complex carbohydrates

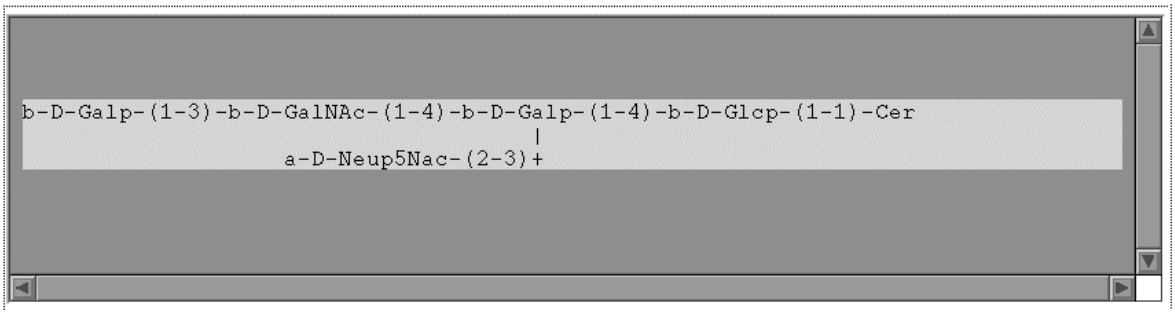
can be found at

http://www.dkfz-heidelberg.de/spec/sweet2/doc/sam/leftframe.html

If you active for example Sphingo lipids on the left hand side you get a list of often discussed Sphingo lipids as display on the right hand side:
For example, if you select Sphingo lipids on the left side you will get a list of common Sphingo lipids on the right side.



Clicking on one of these items you receive the extended description of the corresponding molecule. The example below shows the structure of GM1.



With a copy/paste procedure you can directly copy this structure into the input window. The copy/option works slightly different on different computers. Here the procedure for Windows is described.

a) Indicate the structure you want to transfer as shown above (press left mouse button until the complete structure is highlighted)
b) Call copy (kopieren) from the edit (bearbeiten) menu of your browser window.
c) Click into the Structure input window
d) Call paste (einfügen) from the edit (Bearbeiten) pulldownmenu of the input browser window

Please check that you copied the complete structure and that all vertical connections are at the right position (see Extended Description for Oligosaccharides; additional rules).

The same procedure has to be applied if you have created your carbohydrate structures using your favourite editor. We strongly recommend you use Courier to create your input structures.

Additionally you have to input the number of residues contained in the sequence you input. In case this number and the number of residues identified by the program are not identical, you get a warning.

Wrong number of residues found! I have found 5 residue(s).
Should have found 3 Residue(s).
Structural data is not correct!

Normally this warning indicates that your input sequence was not correct. Most frequently the vertical alignment of monosaccharides in different lines is incorrect. See (2.3.2.) for more details.

2.3. Extended Description for Oligosaccharides

To input your carbohydrate structure, the so-called extended form is used to describe Oligosaccharides . The definition can be found at

http://www.chem.qmw.ac.uk/iupac/2carb/38.html

2.3.1. Example of a N-Glycan structure

```
       α-D-Manp-(1-2)-α-D-Manp-(1-6)+
                                    |
                              α-D-Manp-(1-6)+
                                   |        |
       α-D-Manp-(1-2)-α-D-Manp-(1-3)+        |
                                             |
                                    β-D-Manp-(1-4)-β-D-GlcpNAc-(1-4)-β-D-GlcpNAc-(1-4)-Asn
                                             |
α-D-Manp-(1-2)-α-D-Manp-(1-2)-α-D-Manp-(1-3)+
```

Using the input window a few special input rules have to be fulfilled; these are discussed for the N-Glycan structure given above.

2.3.2. Additional Input rules

a) a,b have to used instead of α,β.
b) the ring configuration (D or L) has to be given always.
c) a hyphen before and after the linkage information is required -(1-3)-
d) Furanoses and Pyranoses have to be indicated.
e) A '+'-sign is required to indicate continuation of the structure in an upper of lower line of the editor.
f) To indicate the vertical connections the '+' ,'|' and the connection to the corresponding monosaccharide have to be in one vertical line. This requirement often causes problems when inputting structures by the copy/paste option when a proportional fond (like Arial or Times) is used.

Correct input

```
a-D-Manp-(1-2)-a-D-Manp-(1-6)+
                              |
                        a-D-Manp-(1-6)
                              |
a-D-Manp-(1-2)-a-D-Manp-(1-3)+
```

Incorrect input:

```
a-D-Manp-(1-2)-a-D-Manp-(1-6)+
                              |
                        a-D-Manp-(1-6)
                              |
a-D-Manp-(1-2)-a-D-Manp-(1-3)+
```

Incorrect input:

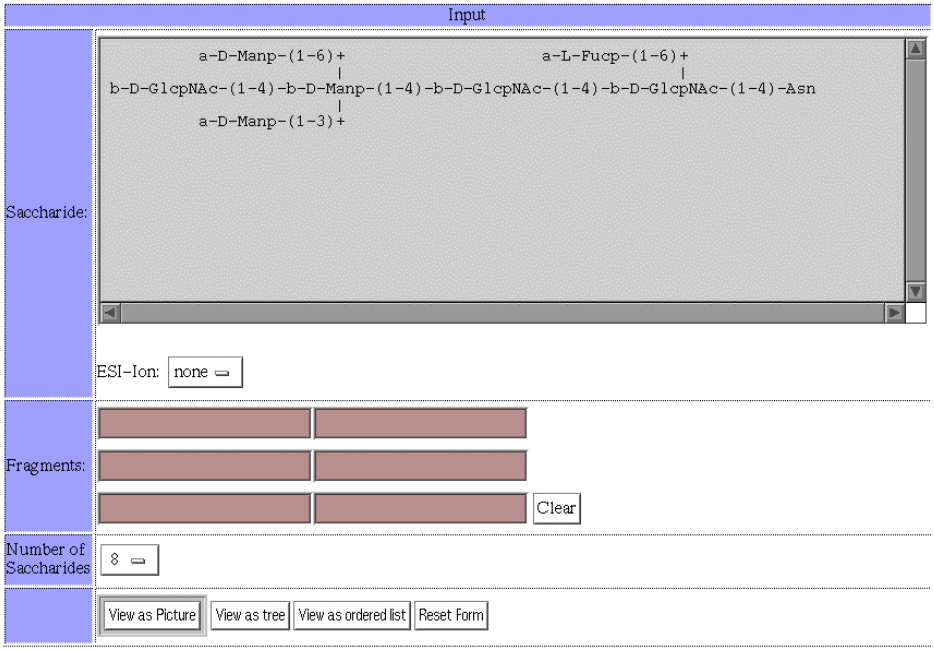
```
a-D-Manp-(1-2)-a-D-Manp-(1-6)+
                              |
                        a-D-Manp-(1-6)
                              |
          a-D-Manp-(1-2)-a-D-Manp-(1-3)+
```

3. Representations of the Fragmentation Scheme

Three different representation of the fragmentation scheme are currently available.

3.1. Ordered list

Press the ‘’ ***view as ordered list*”** **button** at the bottom of the input window to activate this option.



In the first column the fragments are ordered according their increasing mass. The type of ion is given in the second column. The identification of a fragment in carbohydrate sequence is given through the linkage path information. The core fucosylated N-Glycan structure given above contains two a-D-Manp residues. Linkage path information is used to tell the user which of the two residues is specified. Starting from the reducing end of the molecule (right side) and counting only the connections on the reducing side the linkage path 4,4,4,6 indicates the a-D-Manp residue at the top/left. The other a-D-Manp has the linkage path information 4,4,4,3.

View with A and X Ions

Total Mass: 1357.5130 M[+]: 1358.5208 M[Na+]: 1380.5020 M[K+]: 1396.6110

Mass in amu	Ion	Linkage-Path
99.0558	Z(1)-Ion	4
115.0507	Y(1)-Ion	4
147.0657	B(1)-Ion	4, 6
163.0606	C(1)-Ion	4, 6
163.0606	B(4)-Ion	4, 4, 4, 3
	B(2)-Ion	4, 4, 4, 6
179.0556	C(4)-Ion	4, 4, 4, 3
	C(2)-Ion	4, 4, 4, 6
204.0872	B(3)-Ion	4, 4, 4, 4
220.0821	C(3)-Ion	4, 4, 4, 4
448.1931	Z(2)-Ion	4, 4
464.1880	Y(2)-Ion	4, 4
651.2725	Z(3)-Ion	4, 4, 4
667.2674	Y(3)-Ion	4, 4, 4
690.2456	B(5)-Ion	4, 4, 4
706.2406	C(5)-Ion	4, 4, 4
893.3250	B(6)-Ion	4, 4
909.3199	C(6)-Ion	4, 4
1137.4309	Z(5)-Ion	4, 4, 4, 4
1153.4258	Y(5)-Ion	4, 4, 4, 4
1178.4575	Z(4)-Ion	4, 4, 4, 3
	Z(6)-Ion	4, 4, 4, 6
1194.4524	Y(4)-Ion	4, 4, 4, 3
	Y(6)-Ion	4, 4, 4, 6
1194.4524	Z(7)-Ion	4, 6
1210.4473	Y(7)-Ion	4, 6
1242.4623	B(7)-Ion	4
1258.4572	C(7)-Ion	4

Activating the '***view with A- and X- Ions***'**-button** displays an ordered list of all possible Y-,Z-,B-,C- and A Ions, which is normally rather long. Here only the possible assignment of the first five smallest fragments (left) and a cut out of some fragments with higher masses (right) are given. The complete list of fragments and their assignments can be output using the print option of the browser.

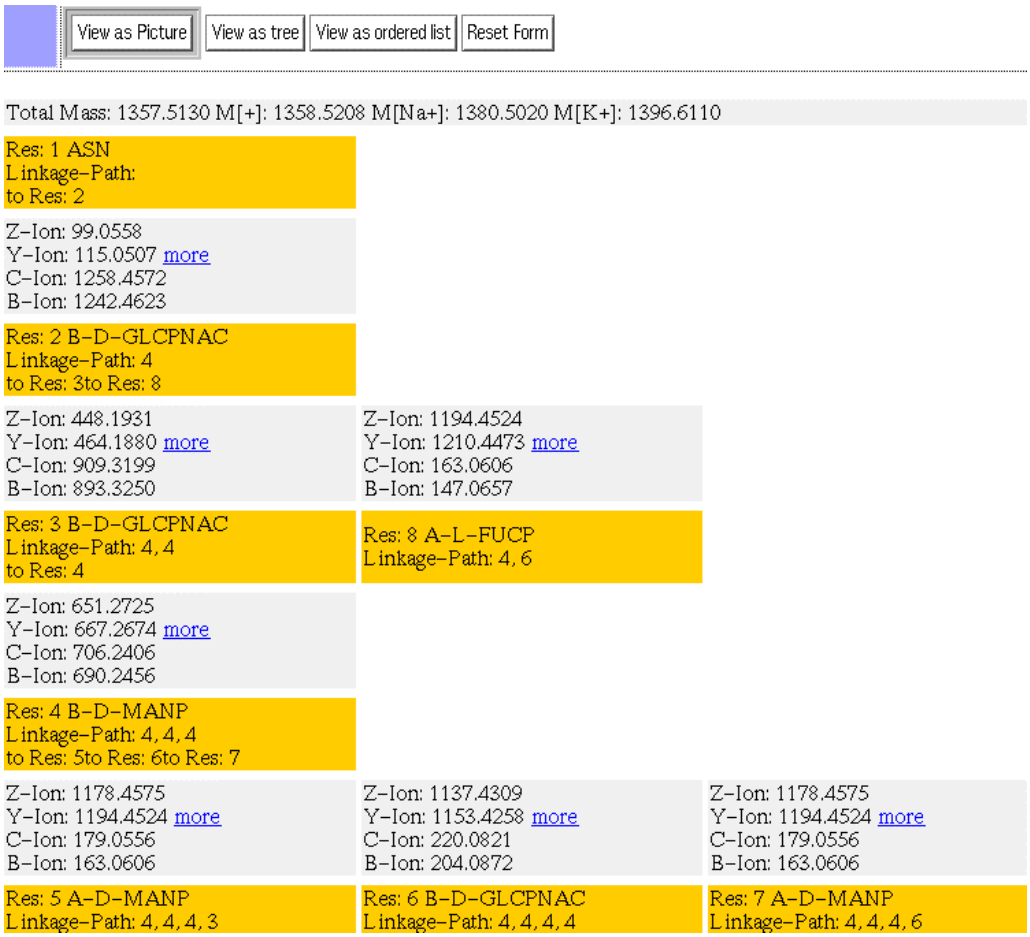
View without A and X Ions

Total Mass: 1357.5130 M[+]: 1358.5208 M[Na+]: 1380.5020 M[K+]: 1396.6110

Mass in amu	Ion	Linkage-Path
13.0078	A(7)0,1-Ion	4
	A(6)0,1-Ion	4, 4
	A(5)0,1-Ion	4, 4, 4
	A(4)0,1-Ion	4, 4, 4, 3
	A(3)0,1-Ion	4, 4, 4, 4
	A(2)0,1-Ion	4, 4, 4, 6
	A(1)0,1-Ion	4, 6
25.0078	X(1)4,5-Ion	4
	X(2)4,5-Ion	4, 4
	X(3)4,5-Ion	4, 4, 4
	X(4)4,5-Ion	4, 4, 4, 3
	X(5)4,5-Ion	4, 4, 4, 4
	X(6)4,5-Ion	4, 4, 4, 6
	X(7)4,5-Ion	4, 6
27.0109	X(1)1,2-Ion	4
	X(2)1,2-Ion	4, 4
	X(5)1,2-Ion	4, 4, 4, 4
28.0313	A(1)4,5-Ion	4, 6
29.0027	X(1)2,3-Ion	4
	X(1)3,4-Ion	4
	X(2)2,3-Ion	4, 4
	X(2)3,4-Ion	4, 4
	X(3)1,2-Ion	4, 4, 4
	X(3)2,3-Ion	4, 4, 4
	X(3)3,4-Ion	4, 4, 4
	X(4)1,2-Ion	4, 4, 4, 3
	X(4)2,3-Ion	4, 4, 4, 3
	X(4)3,4-Ion	4, 4, 4, 3
	X(5)2,3-Ion	4, 4, 4, 4
	X(5)3,4-Ion	4, 4, 4, 4
	X(6)1,2-Ion	4, 4, 4, 6
	X(6)2,3-Ion	4, 4, 4, 6
952.3383	A(7)2,4-Ion	4
1023.3754	A(7)1,4-Ion	4
1036.3832	A(7)0,4-Ion	4
1112.4119	A(7)3,5-Ion	4
1137.4309	Z(5)-Ion	4, 4, 4, 4
1142.4224	A(7)2,5-Ion	4
1153.4258	Y(5)-Ion	4, 4, 4, 4
1166.4337	X(5)0,1-Ion	4, 4, 4, 4
1178.4575	Z(4)-Ion	4, 4, 4, 3
	Z(6)-Ion	4, 4, 4, 6
1193.4446	X(5)0,2-Ion	4, 4, 4, 4
1194.4524	Y(4)-Ion	4, 4, 4, 3
	Y(6)-Ion	4, 4, 4, 6
1194.4524	Z(7)-Ion	4, 6
1207.4602	X(4)0,1-Ion	4, 4, 4, 3
1207.4602	X(6)0,1-Ion	4, 4, 4, 6
1210.4473	Y(7)-Ion	4, 6
1213.4595	A(7)1,5-Ion	4
1222.4473	X(5)0,3-Ion	4, 4, 4, 4
1223.4551	X(7)0,1-Ion	4, 6
1226.4674	A(7)0,5-Ion	4
1236.4629	X(4)0,2-Ion	4, 4, 4, 3
1236.4629	X(6)0,2-Ion	4, 4, 4, 6
1242.4623	B(7)-Ion	4
1251.4500	X(5)0,4-Ion	4, 4, 4, 4
1252.4579	X(7)0,2-Ion	4, 6
1258.4572	C(7)-Ion	4
1265.4657	X(4)0,3-Ion	4, 4, 4, 3
1265.4657	X(6)0,3-Ion	4, 4, 4, 6
1276.4579	X(5)0,5-Ion	4, 4, 4, 4
1281.4606	X(7)0,3-Ion	4, 6
1294.4684	X(4)0,4-Ion	4, 4, 4, 3
1294.4684	X(6)0,4-Ion	4, 4, 4, 6
1310.4633	X(7)0,4-Ion	4, 6
1319.4762	X(4)0,5-Ion	4, 4, 4, 3

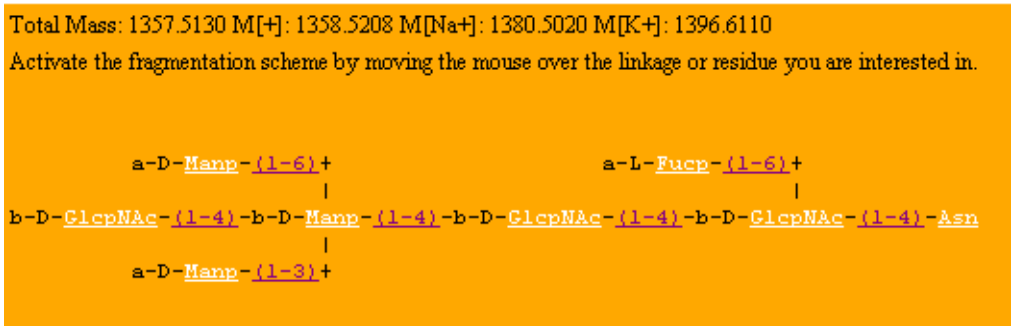
3.2. Tree Structure

Press the **'*view as tree*' button** at the bottom of the input window to activate this option. Ions assigned to a certain linkages are given according the tree spanned by glycans starting from the reducing end. Linkage path information is used to indicate the various branches.

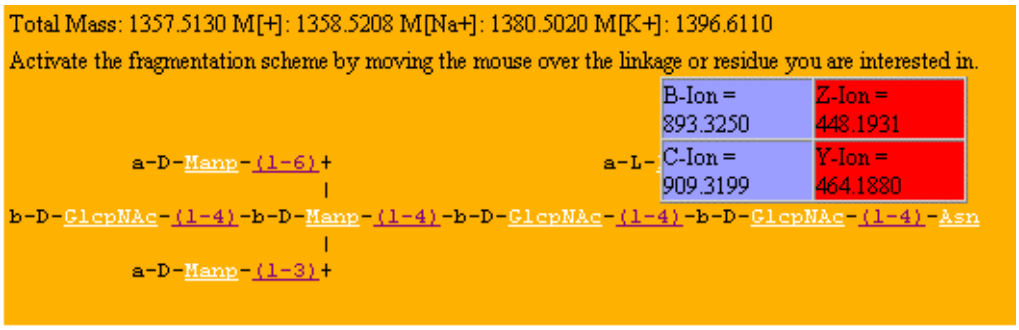


3.3. View as Picture

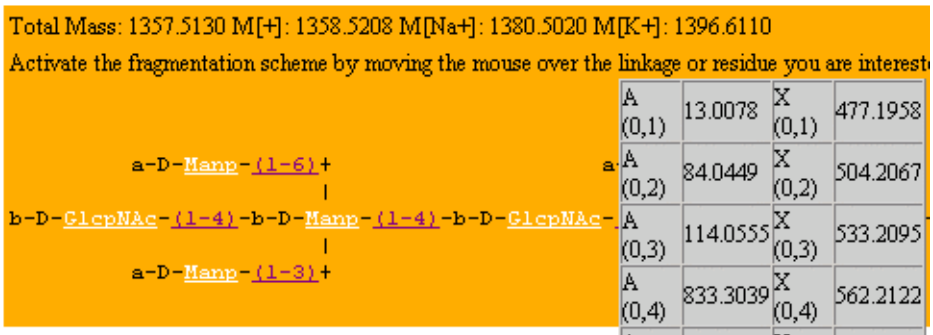
Press the ***"view as picture"*** **button** at the bottom of the input window to activate this option.



When moving the mouse over the linkage of interest (indicated in magenta), the masses of corresponding B- and C-Ions as well as the masses of Z-,Y-Ions will be presented in a small popup window



When moving the mouse over the residue of interest (indicated in white), the masses of corresponding A- and X-Ions will be presented in a popup window. The indices following the ion indicate the type of fragmentation. (A(0,4) means that the bond between the ring oxygen/C1 and the bond between C4 / C5 are broken.



A (0,1)	13.0078	X (0,1)	477.1958
A (0,2)	84.0449	X (0,2)	504.2067
A (0,3)	114.0555	X (0,3)	533.2095
A (0,4)	833.3039	X (0,4)	562.2122
A (0,5)	877.3301	X (0,5)	587.2200
A (1,2)	71.0371	X (1,2)	27.0109
A (1,3)	101.0477	X (1,3)	56.0136
A (1,4)	820.2961	X (1,4)	85.0164
A (1,5)	864.3223	X (1,5)	110.0242
A (2,3)	30.0106	X (2,3)	29.0027
A (2,4)	749.2589	X (2,4)	58.0055
A (2,5)	793.2852	X (2,5)	83.0133
A (3,4)	719.2484	X (3,4)	29.0027
A (3,5)	763.2746	X (3,5)	54.0106
A (4,5)	44.0262	X (4,5)	25.0078

View with A and X Ions

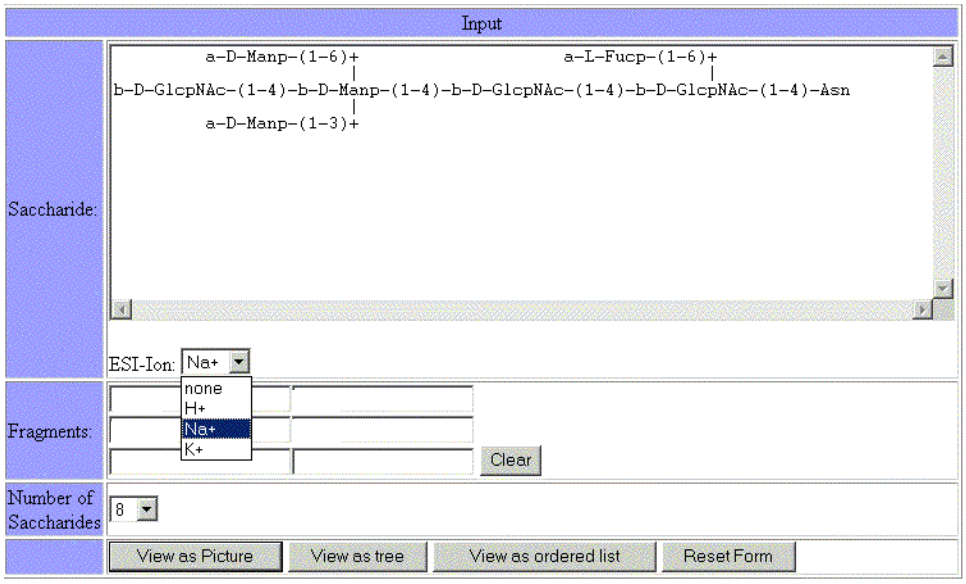
Total Mass: 1357.5130 M[+]: 1358.5208 M[Na+]: 1380.5020 M[K+]: 1396.6110

Mass in amu	Ion	Linkage-Path
123.0534	Z(1)-Ion	4
139.0483	Y(1)-Ion	4
169.0477	B(1)-Ion	4, 6
185.0426	C(1)-Ion	4, 6
185.0426	B(4)-Ion	4, 4, 4, 3
	B(2)-Ion	4, 4, 4, 6
201.0375	C(4)-Ion	4, 4, 4, 3
	C(2)-Ion	4, 4, 4, 6
226.0691	B(3)-Ion	4, 4, 4, 4
242.0640	C(3)-Ion	4, 4, 4, 4
472.1907	Z(2)-Ion	4, 4
488.1856	Y(2)-Ion	4, 4
675.2701	Z(3)-Ion	4, 4, 4
691.2650	Y(3)-Ion	4, 4, 4
712.2276	B(5)-Ion	4, 4, 4
728.2225	C(5)-Ion	4, 4, 4
915.3069	B(6)-Ion	4, 4
931.3019	C(6)-Ion	4, 4
1161.4285	Z(5)-Ion	4, 4, 4, 4
1177.4234	Y(5)-Ion	4, 4, 4, 4
1202.4551	Z(4)-Ion	4, 4, 4, 3
	Z(6)-Ion	4, 4, 4, 6
1218.4500	Y(4)-Ion	4, 4, 4, 3
	Y(6)-Ion	4, 4, 4, 6
1218.4500	Z(7)-Ion	4, 6
1234.4449	Y(7)-Ion	4, 6
1264.4442	B(7)-Ion	4
1280.4391	C(7)-Ion	4

4. Ion mode and adducts

4.1. Electrospray Mass Spectrometry detection mode

The user may enter the masses as neutral ions, positive ions, or as negative ions. Examples of these are [M], $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$



The resulting output (given here as an ordered list) will look like this:

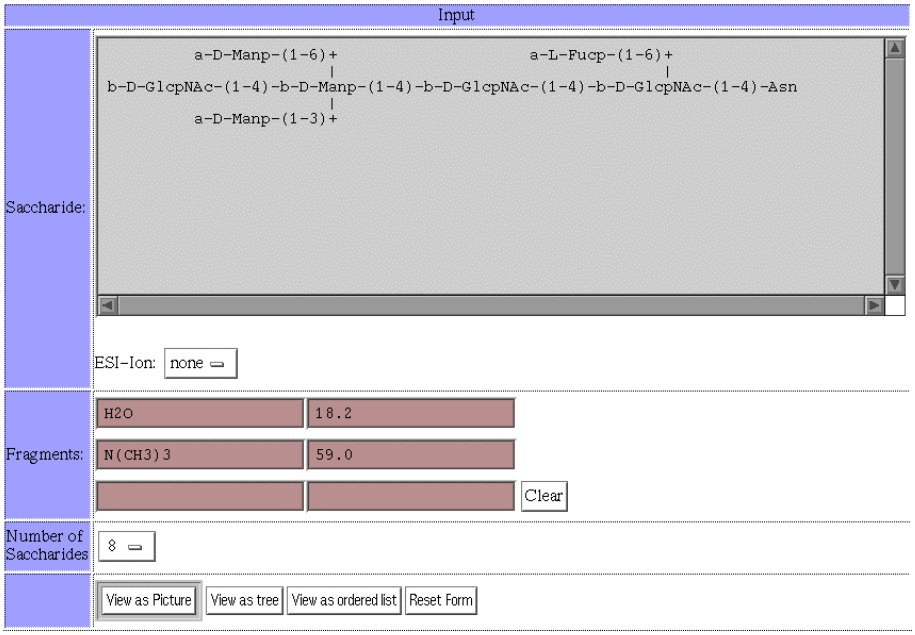
View with A and X Ions

Total Mass: 1357.5130 M[+]: 1358.5208 M[Na+]: 1380.5020 M[K+]: 1396.6110

Mass in amu	Ion	Linkage-Path
123.0534	Z(1)-Ion	4
139.0483	Y(1)-Ion	4
169.0477	B(1)-Ion	4, 6
185.0426	C(1)-Ion	4, 6
185.0426	B(4)-Ion	4, 4, 4, 3
	B(2)-Ion	4, 4, 4, 6
201.0375	C(4)-Ion	4, 4, 4, 3
	C(2)-Ion	4, 4, 4, 6
226.0691	B(3)-Ion	4, 4, 4, 4
242.0640	C(3)-Ion	4, 4, 4, 4
472.1907	Z(2)-Ion	4, 4
488.1856	Y(2)-Ion	4, 4
675.2701	Z(3)-Ion	4, 4, 4
691.2650	Y(3)-Ion	4, 4, 4
712.2276	B(5)-Ion	4, 4, 4
728.2225	C(5)-Ion	4, 4, 4
915.3069	B(6)-Ion	4, 4
931.3019	C(6)-Ion	4, 4
1161.4285	Z(5)-Ion	4, 4, 4, 4
1177.4234	Y(5)-Ion	4, 4, 4, 4
1202.4551	Z(4)-Ion	4, 4, 4, 3
	Z(6)-Ion	4, 4, 4, 6
1218.4500	Y(4)-Ion	4, 4, 4, 3
	Y(6)-Ion	4, 4, 4, 6
1218.4500	Z(7)-Ion	4, 6
1234.4449	Y(7)-Ion	4, 6
1264.4442	B(7)-Ion	4
1280.4391	C(7)-Ion	4

4.2. Input of adducts

The user has the option of specifying any kind of adducts by giving the name of an adduct and the mass associated with this adduct. There is no restriction on the names which may be used for adducts.



The result is an ordered list of masses. The mass of the adducts is added to each fragments of the oligosaccharide.

Total Mass: 1357.5130 M[+]: 1358.5208 M[Na+]: 1380.5020 M[K+]: 1396.6110

Mass in amu	Ion	Linkage-Path
99.0558	Z(1)-Ion	4
115.0507	Y(1)-Ion	4
147.0657	B(1)-Ion	4, 6
163.0606	C(1)-Ion	4, 6
163.0606	B(4)-Ion	4, 4, 4, 3
	B(2)-Ion	4, 4, 4, 6
179.0556	C(4)-Ion	4, 4, 4, 3
	C(2)-Ion	4, 4, 4, 6
204.0872	B(3)-Ion	4, 4, 4, 4
220.0821	C(3)-Ion	4, 4, 4, 4
448.1931	Z(2)-Ion	4, 4
464.1880	Y(2)-Ion	4, 4
651.2725	Z(3)-Ion	4, 4, 4
667.2674	Y(3)-Ion	4, 4, 4
690.2456	B(5)-Ion	4, 4, 4
706.2406	C(5)-Ion	4, 4, 4
893.3250	B(6)-Ion	4, 4
909.3199	C(6)-Ion	4, 4
1137.4309	Z(5)-Ion	4, 4, 4, 4
1153.4258	Y(5)-Ion	4, 4, 4, 4
1178.4575	Z(4)-Ion	4, 4, 4, 3
	Z(6)-Ion	4, 4, 4, 6
1194.4524	Y(4)-Ion	4, 4, 4, 3
	Y(6)-Ion	4, 4, 4, 6
1194.4524	Z(7)-Ion	4, 6
1210.4473	Y(7)-Ion	4, 6
1242.4623	B(7)-Ion	4
1258.4572	C(7)-Ion	4

Appendix:

a) How to Input Substitutions of Sugar Residues

Examples:

a) Extended Description for Oligosaccharides

The numbers following the monosaccharide description indicate the location of the substitution at the corresponding C-Atom of the sugar.

a-D-Galp2OMe, a-D-Galp2OAc3OAc, a-Galp2OMe3OMe4OMe6Ome

Currently supported substitutions: Me, Ac, OMe, OAc, NAc, SO3, PO4, Cl

Substitutent	Input
cPCH	cPCH
PCH	PCH
PEA	PEA
n-S03	NSO3
Phosphate	PO4
Schwefelrest	OSO3
Schwefelrest	SO3
NAC	NAc
NAC	OAc
OMe	OMe
Wasser	H2O
CH	CH
EA	EA
Pyridin	Py
Methyl	Me
Acetate	Ac
Glycolsaeure	Gc
Schwefelwasserstoff	SH
Amine	N

b) Composition of Oligosaccharides

Currently supported compositions: Hex, dHex, HexNac, HexAc. Compositions and extended nomenclature can be mixed.

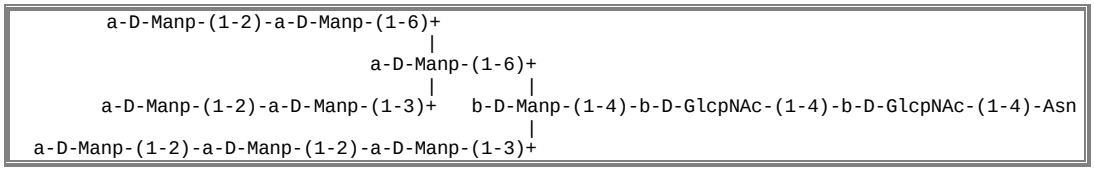
c) Persubstitutions of Oligosaccharides

Currently supported persubstitutions: PMe and PAc. Persubstitutions are only possible for sugar compositions and cannot be used in combination with the extended nomenclature.

b) How to handle secondary fragmentations

To identify the peaks which may result from secondary fragmentations of the complete glycan, the sequences resulting from this fragmentation have to be analyzed separately.

In case you have the following carbohydrate sequence



And you have the impression that some secondary fragmentation takes place than you have to input the sequence of the fragments which remains when the secondary fragmentation takes place. Assuming a fragmentation occurs at the 4,4,4,6 Linkage the following two fragments would result:

Fragment A:

```
                                        b-D-Manp-(1-4)-b-D-GlcpNAc-(1-4)-b-D-GlcpNAc-(1-4)-Asn
                                             |
a-D-Manp-(1-2)-a-D-Manp-(1-2)-a-D-Manp-(1-3)+
```

Fragment B

```
          a-D-Manp-(1-2)-a-D-Manp-(1-6)+
                                        |
                                    a-D-Manp
                                        |
          a-D-Manp-(1-2)-a-D-Manp-(1-3)+
```

Both carbohydrate sequences have to be analyzed separately as shown above. The input of the (sub)sequences can be easily accomplished using copy/paste options.