

Tutorial 1

Aim: To use GlycanAnalyzer to assign N-glycans to the exoglycosidase array in Figure 1. Mass information will be supplied in the form of all m/z intensities for all retention times (Figure 2). In this tutorial we will input the data in Figure 1 and 2 into the software.

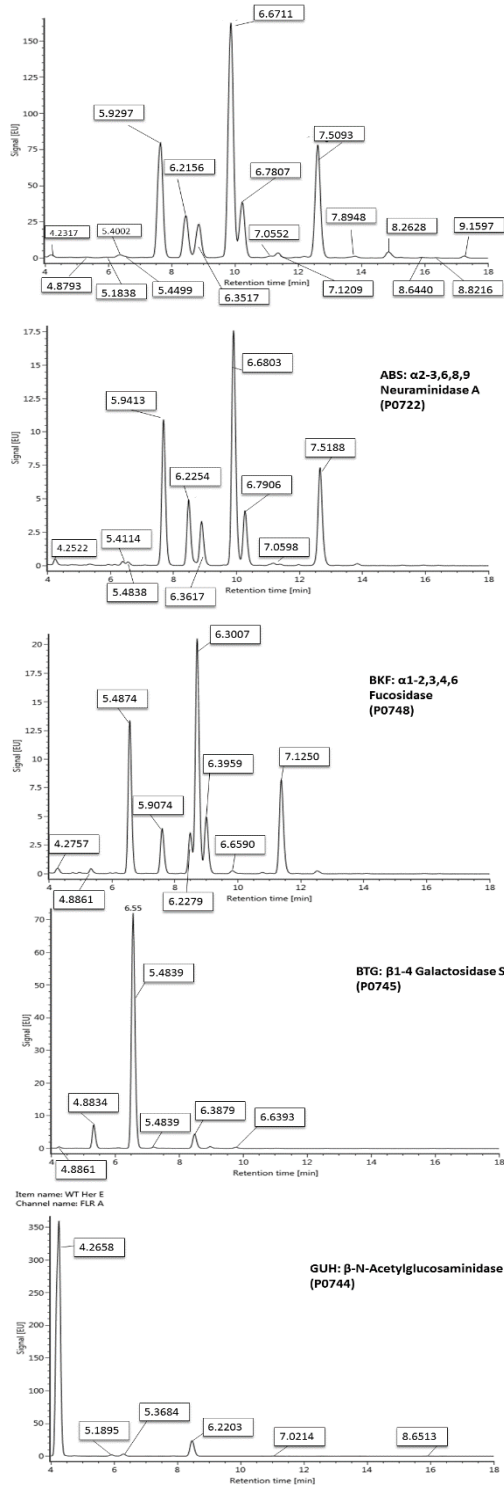


Figure 1. An ABS+BKF+BTG+GUH exoglycosidase array (UPLC). There are three variables: retention time, area of each peak and Glucose Units (boxes). Notice that some peaks have very small size (low abundance) and some are large (high abundance).

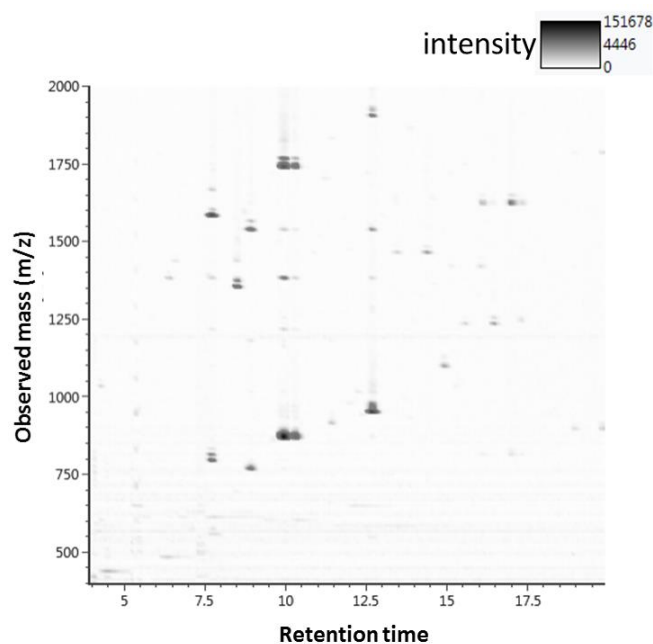


Figure 2. The 3D mass information for an entire UPLC chromatogram. The three dimensions are retention time, m/z and intensity.

Abbreviations:

UPLC: Ultra-performance liquid chromatography

LC: Liquid chromatography

MS: Mass spectrometry

GU: Glucose Units

Step 1: An example dataset is located on the GlycanAnalyzer tutorial webpage called Data_3Dmass.xls. Open the spreadsheet titled Data_3Dmass.xls and examine the data inside each exoglycosidase pane (Note that it is the same data presented in Figure 1 and Figure 2)::

Undigested	ABS	ABS+BKF	ABS+BKF+BTG	ABS+BKF+BTG+GUH
------------	-----	---------	-------------	-----------------

Note: if users are supplying their own datasets then they must have it in exactly the same format as Data_Expert_nomass.xls

Step 2: Notice that the LC data (Fig 1) is tabulated in the spreadsheet under the columns “Observed RT”, “% Amount”, and “Glycan Units”:

Observed RT (min)	% Amount (%)	Glycan Units
4.2	0.39	4.2317
5.32	0.13	4.8793
5.91	0.15	5.1838
6.37	0.5	5.4002
6.48	0.16	5.4499
7.65	18.85	5.9297
8.45	6.67	6.2156

The MS data (Fig 2) is tabulated under the columns “Retention time”, “Observed mass (m/z)”, and “Intensity”:

Retention time (min)	Observed mass (m/z)	Intensity (Counts)
4.04	1531.76806	22506
4.05	420.95744	136955
4.05	422.95481	79376
4.05	488.94466	50222
4.05	774.96025	42541
4.05	796.94088	37528
4.05	556.93316	37378
4.05	490.94139	29755
4.05	1176.92363	29743
4.05	1200.89673	27475
4.05	558.92826	26004
4.05	462.24532	24534
4.05	1197.85683	22941
4.05	1110.97035	22858
4.05	821.92161	22601
4.05	1218.86658	21805

Note: it is vital that all your data is sorted by lowest to highest retention time (i.e. columns ‘Observed RT (min)’ and ‘Retention time (min)’).

The input is summarized in Figure 3.

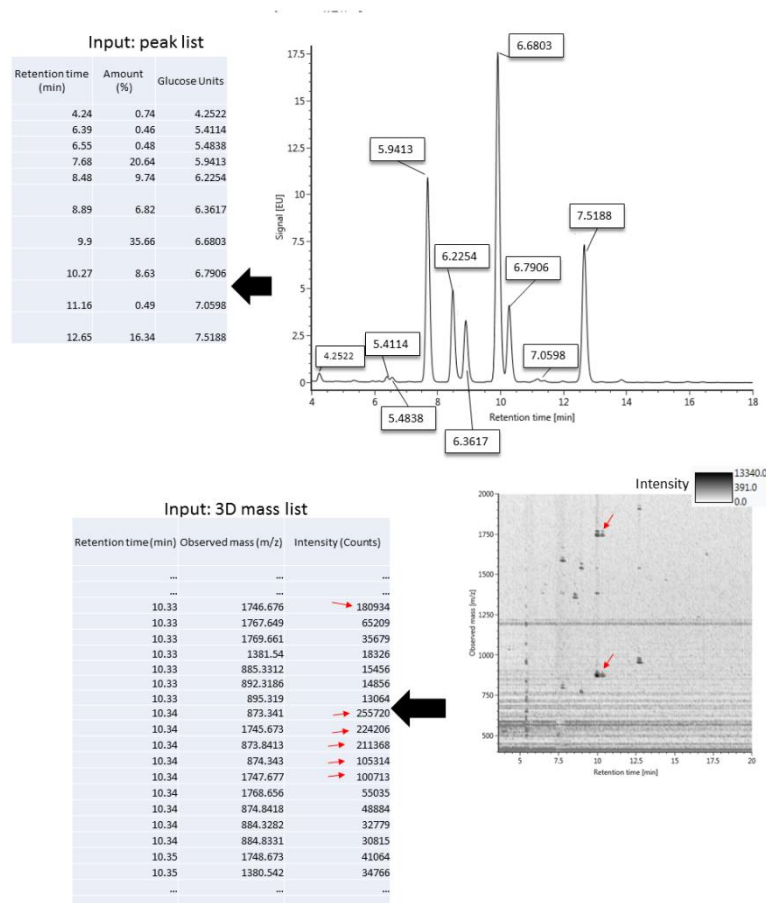
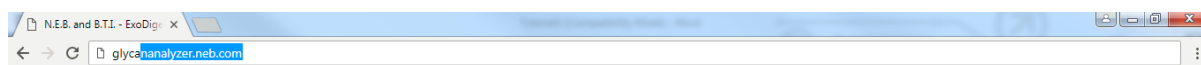


Figure 3. Peak input and 3-dimensional mass input.

Step 3: Load GlycanAnalyzer web application in your browser. Type <http://glycananalyzer.neb.com>



Step 4: We want to generate input for GlycanAnalyzer. Click the link circled on the main page:

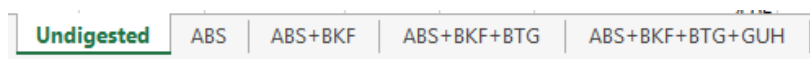
1. Upload exoglycosidase file. ⓘ

Upload a file containing Glucose Units (GUs) and peak areas. Mass and charge columns will increase assignment accuracy.

Generate input file by supplying peaks and mass or merging 3D mass and peak information.

+ Upload file

Step 5: in the undigested data in our spreadsheet:



Copy the 3-dimensional mass list into the left box and the peak list into the text box on the right (see Figure 3 for recap):

1B. Paste the mass (MS1) and peak list. ⓘ

In the text areas below you can enter the MS signal (retention time, m/z and intensity) for the entire chromatogram and all the peaks (retention time, area and GU).
([Click here to paste an example](#))

MAIN profile

MAIN - Mass list (MS1) here

MAIN - Liquid chromatography peak list here

This is the main profile on which exoglycosidase digestions will be applied and glycan annotations will be assigned.

Note that the columns GU, Amount (%), Observed Mass and Observed charge should be separated by tabs. Copy and pasting from spreadsheets automatically separates by tabs.

Step 6: Add data from the spreadsheet by adding exoglycosidases **in order** of application (order ABS, BKF, BTG and GUH – see Figure 1). To add ABS click here:

1A. Add profiles to your exoglycosidase panel.

If you applied a New England Biolab exoglycosidase to your profile select it.

- A main profile must be provided - this is the profile before any digestions take place.
- The **order** of the exoglycosidases in the list is important

Available:

α2-3 Neuraminidase S (NAN1)

α2-3,6,8,9 Neuraminidase A (ABS)

α1-2,3,4,6 Fucosidase (BKF)

β1-3,4 Galactosidase (BTG)

Using:

Main profile (no digestion)

Then add the data in the ABS tab below exactly as in step 5:

Undigested

ABS

ABS+BKF

ABS+BKF+BTG

ABS+BKF+BTG+GUH

Step 7: Repeat for the exoglycosidases as ordered in Figure 1. The final page should resemble the following:

1B. Paste the mass (MS1) and peak list.

In the text areas below you can enter the MS signal (retention time, m/z and intensity) for the entire chromatogram and all the peaks (retention time, area and GU).
(Click here to paste an example)

MAIN profile	Retention time (min)		Observed mass (m/z)		Observed RT (min) % Amount (%) Glycan Units		
	Intensity (Counts)						
	4.04	1531.76806	22506		4.2	0.39	4.2317
	4.15	420.95744	136955		5.32	0.13	4.8793
	4.05	422.95481	79376		5.91	0.15	5.1838
main+ ABS profile	Retention time (min)		Observed mass (m/z)		Observed RT (min) % Amount (%) Glycan Units		
	Intensity (Counts)						
	4.05	527.1582	27984		4.24	0.74	4.2522
	4.08	420.95739	42822		6.39	0.46	5.4114
	4.08	422.9553	25821		6.55	0.48	5.4838
main+abs+ BKF profile	Retention time (min)		Observed mass (m/z)		Observed RT (min) % Amount (%) Glycan Units		
	Intensity (Counts)						
	4	1723.68331	11691		7.68	20.64	5.9413
	4.02	1220.69287	28544		4.27	0.76	4.2757
	4.05	1318.6707	29238		5.33	0.65	4.8861
main+abs+bkf+ BTG profile	Retention time (min)		Observed mass (m/z)		Observed RT (min) % Amount (%) Glycan Units		
	Intensity (Counts)						
	4	1629.63496	20937		6.56	21.91	5.4874
	4	1340.65208	14946		7.59	6.89	5.9074
	4.04	403.22915	23294		4.24	0.46	4.2547
main+abs+bkf+btg+ GUH profile	Retention time (min)		Observed mass (m/z)		Observed RT (min) % Amount (%) Glycan Units		
	Intensity (Counts)						
	4	1619.71839	26539		5.32	7.76	4.8834
	4	1717.69659	18300		6.55	84.26	5.4839
	4.01	1629.63073	19547		7.2	0.48	5.7554



Click the button and save the file to your PC.

After clicking 'Merge data to input file' the user is asked if their MS data contains a fluorescent label. Users should select one of three things: (i) select a label supported by GlycanAnalyzer (2AB, RFMS, procainamide), (ii) select 'other' and enter the mass labels explicitly or (iii) select label free. In this tutorial our dataset was generated using the 2AB label and we therefore should make sure this button is clicked:

Additional mass.

In order to detect the correct mass GlycanAnalyzer needs to take additional mass into account (e.g. florescent labels). The other textbox can be used to enter the mass of other labels and/or mass of peptides. Click 'no additional mass' if N-glycans are unmodified.

☒ **2-aminobenzamide (2AB)**

☐ **RapiFluor-MS™ (RFMS)**

☐ **Procainamide**

☐ **Other**

☐ **No additional mass (e.g. your mass data does not consider the mass of a fluorescent label)**

OK I got it

Step 8: After the input generation is complete you will be directed back to the main input page. On the main input page upload the file generated in step 7 by clicking the button circled:

1. Upload exoglycosidase file. ⓘ

Upload a file containing Glucose Units (GUs) and peak areas. Mass and charge columns will increase assignment accuracy.

Generate input file by [supplying peaks and mass](#) or [merging 3D mass and peak information](#).

+ Upload file

2. Select a peak to analyze. ⓘ

Select the peak that you want to determine glycan structures in the **undigested** profile.

Select peak ▲

3. Select a glycoprotein (optional) ⓘ

Supplying glycoprotein information will reduce the types of possible glycan structures and increase the accuracy.

Select glycoprotein ▲

Get Glycan List

Step 9: Click the “Select peak” button shown below to select the peak you would like to analyze.


2. Select a peak to analyze. ⓘ

Select the peak that you want to determine glycan structures in the **undigested** profile.

Select peak ▲

Select peak 13 (note that this is a large peak with GU 7.5093 in the undigested profile in Figure 1)

Note: it is recommended to start annotating the larger peaks first as these contain the most fluorescent and mass signal and therefore are easier to annotate. This will allow users to build a picture of possible digests to other peaks before smaller peaks are processed.

Step 10: Click the button  and wait for the assignment to complete. Note that your PC is not being used for the calculation. All computations are taking place on our state-of-the-art servers.

Step 11: Figure 4 shows the returned list of N-glycans for peak 13. Figure 4 part 1: The returned N-glycans are ranked by a score (closer to zero the better). 2: The user has the option to reject any candidate N-glycan. 3&4: N-glycans are displayed in oxford notation and drawn in SNFG notation - click the ‘+’ symbol for higher resolution glycan image. 5: The details column gives some clues on the correctness of the assignment might be, for example how close the theoretical masses and GU values are to the isotopic average mass and the GU. 6: Links to other databases. 7: If the user

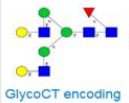

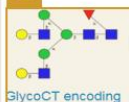

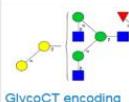

chooses to 'accept structures into chromatogram' the number of completed peaks will increase here. Clicking 'view chromatogram' gives a summary of currently completed peaks (see step 14). 8: More visualization is available when these links are clicked this is explained in more detail in step 13.

Ranked list of N-glycans for peak 13

Below are the N-glycan ratings, scored by **Observed** mass changes and GU shifts. Remove glycans, as needed (e.g. α -galactose in human), and assign them to the chromatogram.

7 Peak analysis status:
0 of 18 peaks (0%).

[View chromatogram](#)

Keep?	Score	Oxford notation	Diagram	Details	More information	graphical analysis
<input checked="" type="checkbox"/>	0.16	F(6)A2G2	 GlycoCT encoding	5 Mass (Observed / Expected): 1908.23 / 1907.7 GU (Observed / Expected): 7.5093 / 7.54 Δ GU : 0.0307	 1 2 3 4 5 1	Peak shift graph Score calculation Main panel mass evidence
<input checked="" type="checkbox"/>	0.19	F(6)A2G(4)2	 GlycoCT encoding	4 Mass (Observed / Expected): 1908.23 / 1907.7 GU (Observed / Expected): 7.5093 / 7.571 Δ GU : 0.0617	 1 2 3 4 5 1	8 Peak shift graph Score calculation Main panel mass evidence
<input checked="" type="checkbox"/>	1.91	F(6)A2G(4)1Ga(3)1	 GlycoCT encoding	Mass (Observed / Expected): 1908.23 / 1907.71 GU (Observed / Expected): 7.5093 / 7.4257 Δ GU : 0.0836	 1 2 3 1	Peak shift graph Score calculation Main panel mass evidence

[Accept structures into chromatogram](#)

[Analyze New Peak](#)

Figure 4. The single peak assignment output.

Step 12: If the user is finished rejecting candidates and are happy with the current set of candidates we can click [Accept structures into chromatogram](#). In this case the completed peaks will be updated as follows:

Peak analysis status:
1 of 18 peaks (5.56%).

[View chromatogram](#)

, the user can view a summary of their current accepted peaks by clicking 'view chromatogram'. This summary is the topic of step 14.

Step 13: The output in Figure 4 part 8 can be broken into three stages

'Peak shift graph' (Figure 5): This presents a directed graph showing the estimated peak movement. The size of a circle is proportional to the area of the peak. Clicking the circles reveals the glycan structure assigned to the peak in each digestion profile.

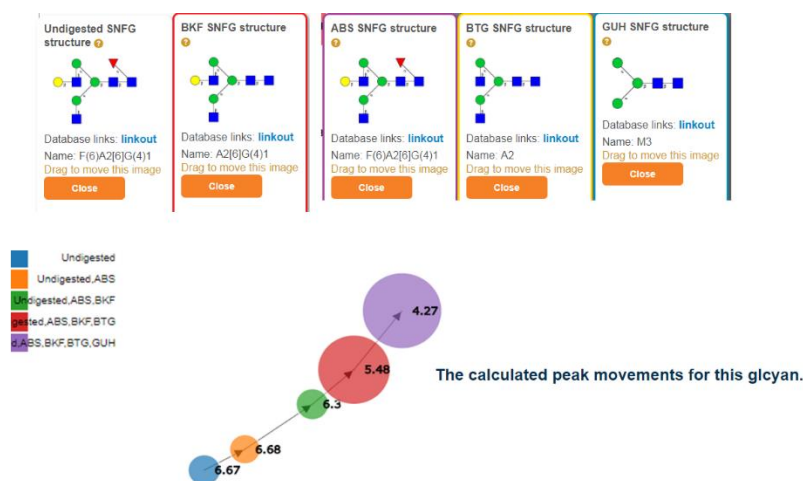


Figure 5. The directed graph representation of the peak movements.

'Score calculation' (Figure 6): This is perhaps the most important information revealed by GlycanAnalyzer. It reveals the following information:

- Expected m/z is present in all digestions.
- The peak shift on the H/UPLC chromatograms.

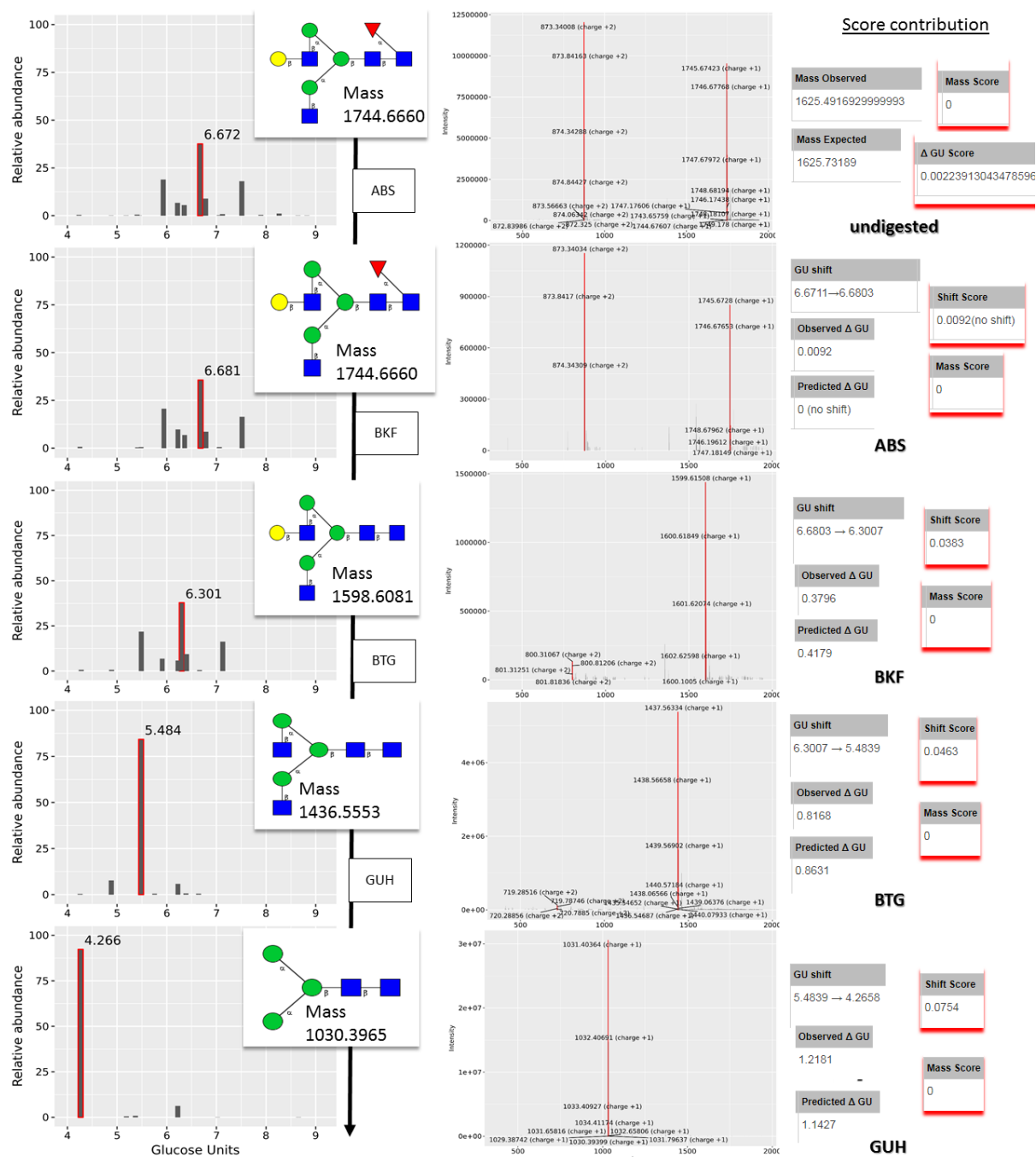


Figure 6. Score calculation for peak 9 of the monoclonal antibody. The total score is 0.1714 coming from the following contributions: $\Delta GU = 0.0022$, Mass score = 0 (all masses found; this increments by 1 if mass not found in any profile), Shift score = $0.0092 + 0.0383 + 0.0463 + 0.0754$. Total score is $0.0022 + 0 + 0.0092 + 0.0383 + 0.0463 + 0.0754$.

‘Main panel mass evidence’ (Figure 7): This is simply the mass spectra of the N-glycan(s) which were assigned to each peak. The red bar indicates possible m/z for the N-glycan in all charge states.

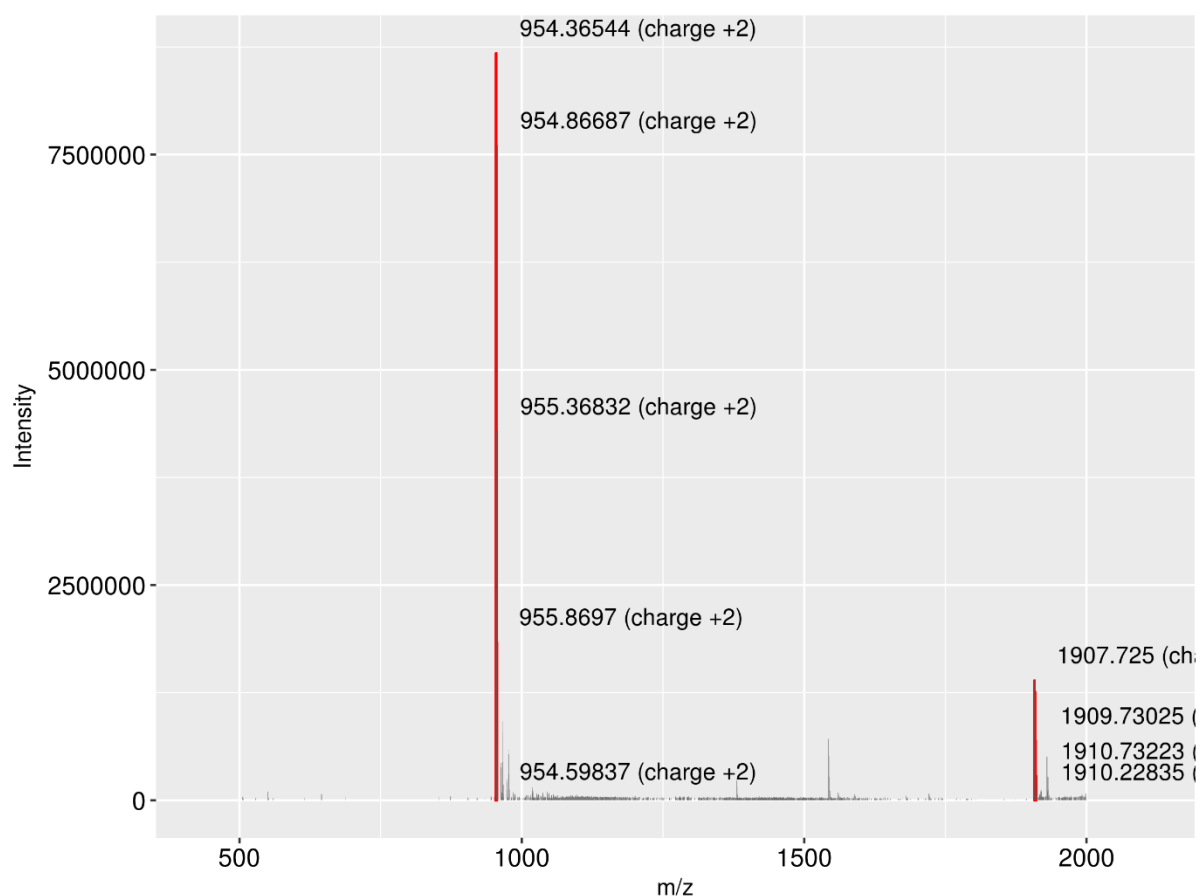


Figure 7. The mass spectra for peak 13 in the monoclonal antibody. Two peaks were found to be linked to the N-glycan assignment (red) corresponding to the +2H and +1H charged states. Other peaks are coloured in grey.

Step 14: Once the peaks are accepted (Figure 4) the user can view a summary of all the completed assignments by clicking the 'View chromatogram' button in Figure 4 part 7. The resulting summary page is shown in Figure 8 for accepted peaks 13 and 14.

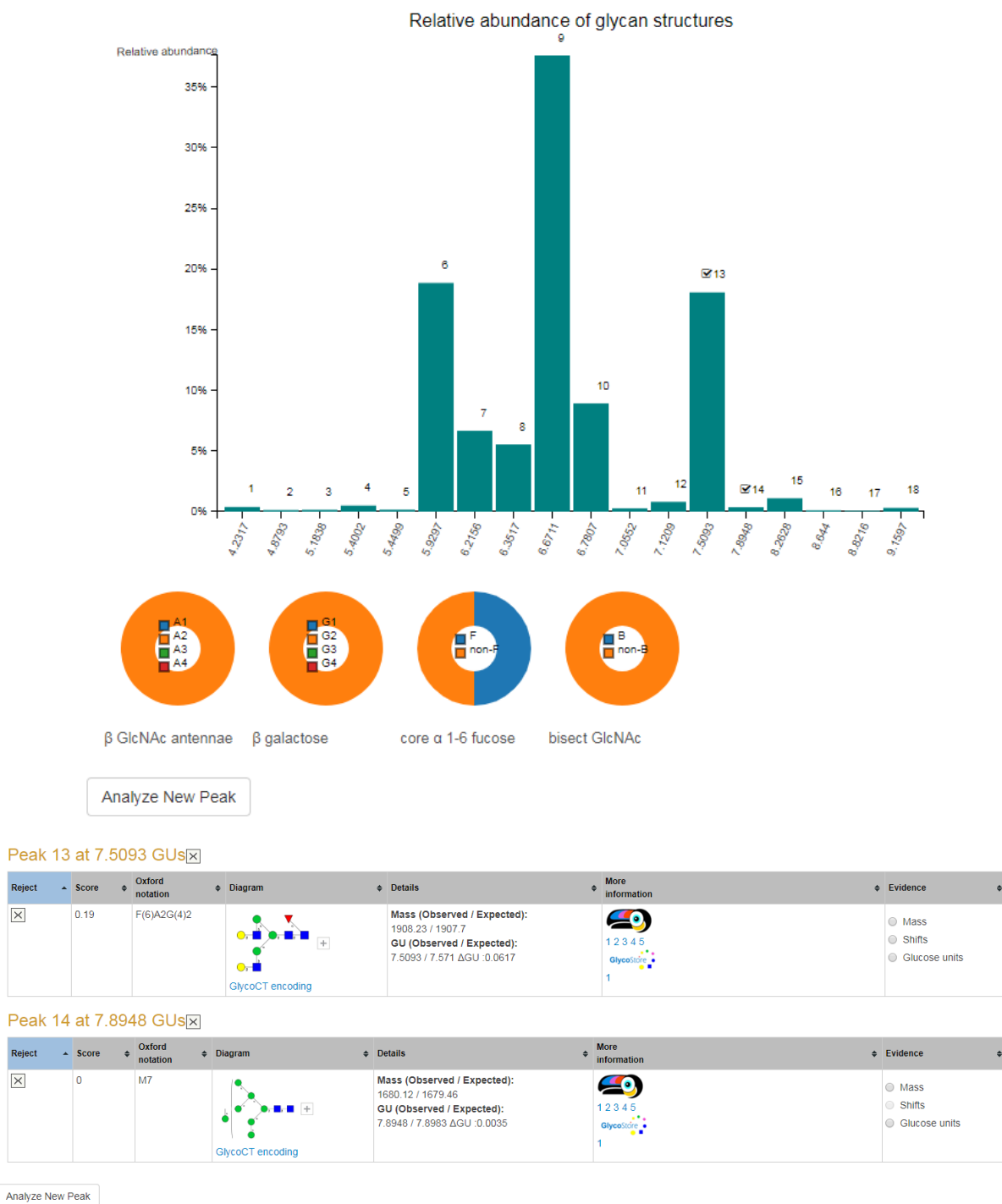


Figure 8. The summary output page for the monoclonal antibody. This view is presented when clicking ‘view chromatogram’ in Figure 4. The bar chart shows each peaks relative abundance where a tick indicates user acceptance of the peak. The pie charts give the distribution of sialic acids, GlcNAc antennae, galactose and other monosaccharides. The tables present the N-glycans in the accepted peaks. Peak 13 contains the best possible sources of information: matching mass, matching GU and mass and GU shifts. Peak 14 has only two pieces of evidence: matching mass and matching GU, there were no shifting peaks found due to its small size. Peak 13’s N-glycan assignment can be considered to have strong support while peak 14’s has medium support. The weakest level of supporting evidence is GU similarity alone. The Mass, Shifts and Glucose unit buttons can be clicked to visualize the evidence.