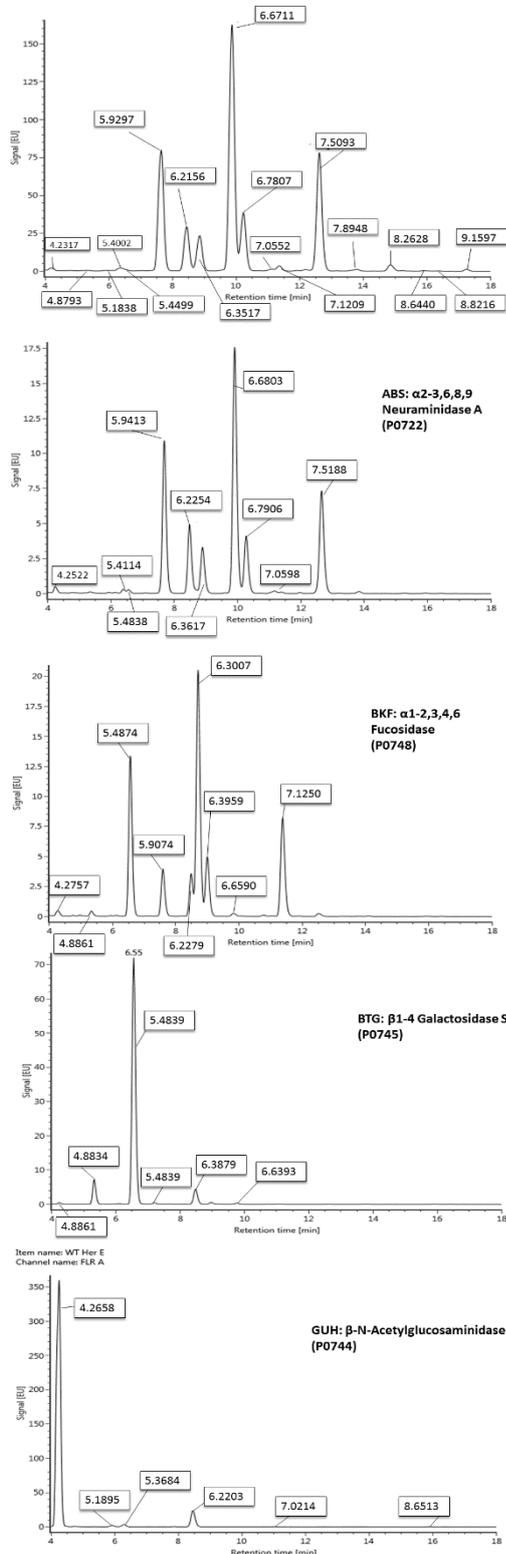


# Tutorial 2

**Aim:** Guides the user through the input process when they have determined the mass and charge states themselves. In other words mass is not extracted from the raw data but instead provided by the user. Shows how we can assign all peaks with one button click. Figure 1 shows the monoclonal antibody exoglycosidase array which we will work with.



**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).

### 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\_Expert\_mass.xls. Open the spreadsheet titled Data\_Expert\_mass.xls and examine the data inside each exoglycosidase pane (Note that it is the same data presented in Figure 1):

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 "GU", "Amount (%)", and the MS is summarized in the tabs "Observed Mass" and "Observed charge":

GU	Amount (%)	Observed Mass	Observed charge
4.2317	0.39	1030.39653	1
4.8793	0.13	1233.4759	1
5.1838	0.15	1192.44935	1
5.4002	0.5	1379.53381	1
5.4499	0.16	NA	NA
5.9297	18.85	1582.61318	2
6.2156	6.67	1354.50218	1
6.3517	5.54	1541.58663	2
6.6711	37.6	1744.66601	2
6.7807	8.92	1744.66601	1
7.0552	0.27	1516.555	1
7.1209	0.81	1703.63946	1
7.5093	18.07	1906.71883	2
7.8948	0.37	1678.60782	1
8.2628	1.1	2197.81425	2
8.644	0.11	1840.66065	1
8.8216	0.07	1840.66065	1
9.1597	0.31	2488.90966	2

Each peak has now been annotated with a mass and a charge. The user has determined these themselves.

Note that in this case the 2AB label is added to the masses. Also note that one peak in the undigested profile with a very low relative abundance could not be determined ('NA' in Observed Mass and Observed charge columns).

Each peak is therefore defined using four variables as showing in Figure 3.

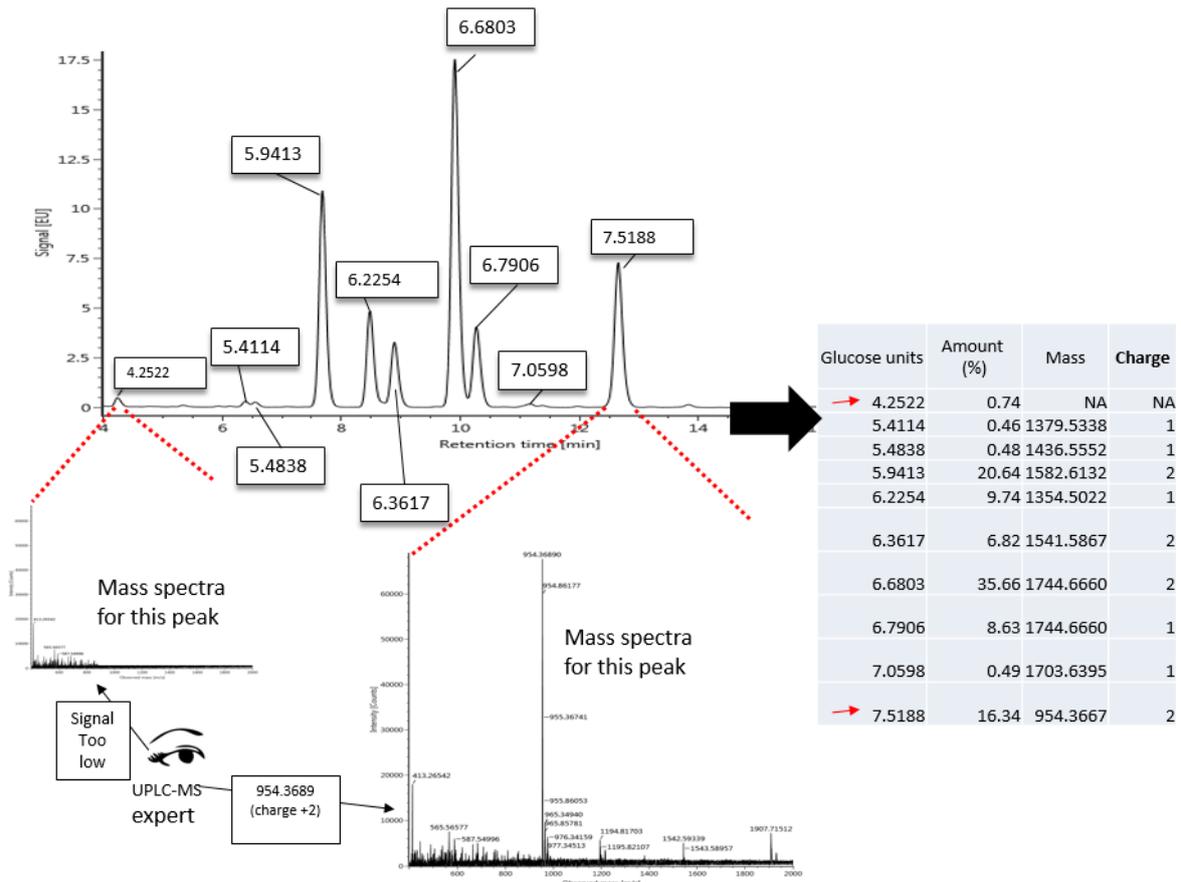


Figure 2. Peak input with mass observed by the user.

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



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

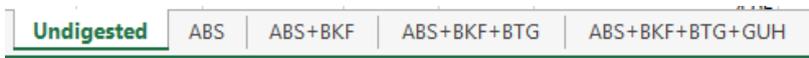
### 1. Upload exoglycosidase file. ⓘ

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.



Step 5: in the undigested data tab in our spreadsheet:



Copy the undigested data into the main profile.:

## 1B. Paste the UPLC-MS profiles as a list.

In the text areas below you can enter UPLC-MS profiles (GU, % Area, Mass, and Charge) for each exoglycosidase used.  
([Click here to paste an example](#))

MAIN profile

MAIN - GU, % area, mass, charge 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

and repeat for BKF, BTG and

GUH enzymes as ordered in Figure 1.

**Step 7:** The final data page should resemble the following:

### 1B. Paste the UPLC-MS profiles as a list.

In the text areas below you can enter UPLC-MS profiles (GU, % Area, Mass, and Charge) for each exoglycosidase used.

[\(Click here to paste an example\)](#)

<b>MAIN profile</b>	<table border="1"><thead><tr><th>GU</th><th>Amount (%)</th><th>Observed Mass</th><th>Observed charge</th></tr></thead><tbody><tr><td>4.2317</td><td>0.39</td><td>1030.39653</td><td>1</td></tr><tr><td>4.8793</td><td>0.13</td><td>1233.4759</td><td>1</td></tr><tr><td>5.1838</td><td>0.15</td><td>1192.44935</td><td>1</td></tr><tr><td>5.4002</td><td>0.5</td><td>1379.53381</td><td>1</td></tr></tbody></table>	GU	Amount (%)	Observed Mass	Observed charge	4.2317	0.39	1030.39653	1	4.8793	0.13	1233.4759	1	5.1838	0.15	1192.44935	1	5.4002	0.5	1379.53381	1
GU	Amount (%)	Observed Mass	Observed charge																		
4.2317	0.39	1030.39653	1																		
4.8793	0.13	1233.4759	1																		
5.1838	0.15	1192.44935	1																		
5.4002	0.5	1379.53381	1																		
main+ <b>ABS profile</b>	<table border="1"><thead><tr><th>GU</th><th>Amount (%)</th><th>Observed Mass</th><th>Observed charge</th></tr></thead><tbody><tr><td>4.2522</td><td>0.74</td><td>1030.39653</td><td>1</td></tr><tr><td>5.4114</td><td>0.46</td><td>1379.53381</td><td>1</td></tr><tr><td>5.4838</td><td>0.48</td><td>1436.55528</td><td>1</td></tr><tr><td>5.9413</td><td>20.64</td><td>1582.61318</td><td>1</td></tr></tbody></table>	GU	Amount (%)	Observed Mass	Observed charge	4.2522	0.74	1030.39653	1	5.4114	0.46	1379.53381	1	5.4838	0.48	1436.55528	1	5.9413	20.64	1582.61318	1
GU	Amount (%)	Observed Mass	Observed charge																		
4.2522	0.74	1030.39653	1																		
5.4114	0.46	1379.53381	1																		
5.4838	0.48	1436.55528	1																		
5.9413	20.64	1582.61318	1																		
main+abs+ <b>BKF profile</b>	<table border="1"><thead><tr><th>GU</th><th>Amount (%)</th><th>Observed Mass</th><th>Observed charge</th></tr></thead><tbody><tr><td>4.2757</td><td>0.76</td><td>1030.39653</td><td>1</td></tr><tr><td>4.8861</td><td>0.65</td><td>1233.4759</td><td>1</td></tr><tr><td>5.4874</td><td>21.91</td><td>1436.55528</td><td>1</td></tr><tr><td>5.9074</td><td>6.89</td><td>1395.52873</td><td>1</td></tr></tbody></table>	GU	Amount (%)	Observed Mass	Observed charge	4.2757	0.76	1030.39653	1	4.8861	0.65	1233.4759	1	5.4874	21.91	1436.55528	1	5.9074	6.89	1395.52873	1
GU	Amount (%)	Observed Mass	Observed charge																		
4.2757	0.76	1030.39653	1																		
4.8861	0.65	1233.4759	1																		
5.4874	21.91	1436.55528	1																		
5.9074	6.89	1395.52873	1																		
main+abs+bkf+ <b>BTG profile</b>	<table border="1"><thead><tr><th>GU</th><th>Amount (%)</th><th>Observed Mass</th><th>Observed charge</th></tr></thead><tbody><tr><td>4.2547</td><td>0.46</td><td>1030.39653</td><td>1</td></tr><tr><td>4.8834</td><td>7.76</td><td>1233.4759</td><td>1</td></tr><tr><td>5.4839</td><td>84.26</td><td>1436.55528</td><td>1</td></tr><tr><td>5.7554</td><td>0.48</td><td>1395.52873</td><td>1</td></tr></tbody></table>	GU	Amount (%)	Observed Mass	Observed charge	4.2547	0.46	1030.39653	1	4.8834	7.76	1233.4759	1	5.4839	84.26	1436.55528	1	5.7554	0.48	1395.52873	1
GU	Amount (%)	Observed Mass	Observed charge																		
4.2547	0.46	1030.39653	1																		
4.8834	7.76	1233.4759	1																		
5.4839	84.26	1436.55528	1																		
5.7554	0.48	1395.52873	1																		
main+abs+bkf+btg+ <b>GUH profile</b>	<table border="1"><thead><tr><th>GU</th><th>Amount (%)</th><th>Observed Mass</th><th>Observed charge</th></tr></thead><tbody><tr><td>4.2658</td><td>92.29</td><td>1030.39653</td><td>1</td></tr><tr><td>5.1895</td><td>0.5</td><td>1192.44935</td><td>1</td></tr><tr><td>5.3684</td><td>0.84</td><td>1192.44935</td><td>1</td></tr><tr><td>6.2203</td><td>6.21</td><td>1354.50218</td><td>1</td></tr></tbody></table>	GU	Amount (%)	Observed Mass	Observed charge	4.2658	92.29	1030.39653	1	5.1895	0.5	1192.44935	1	5.3684	0.84	1192.44935	1	6.2203	6.21	1354.50218	1
GU	Amount (%)	Observed Mass	Observed charge																		
4.2658	92.29	1030.39653	1																		
5.1895	0.5	1192.44935	1																		
5.3684	0.84	1192.44935	1																		
6.2203	6.21	1354.50218	1																		



Click the button Generate input file 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 GylanAnalyzer (2AB, RFMS, procanamidine), (ii) select 'other' and enter the labels mass 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. fluorescent 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:** Once the input is generated 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 ▾

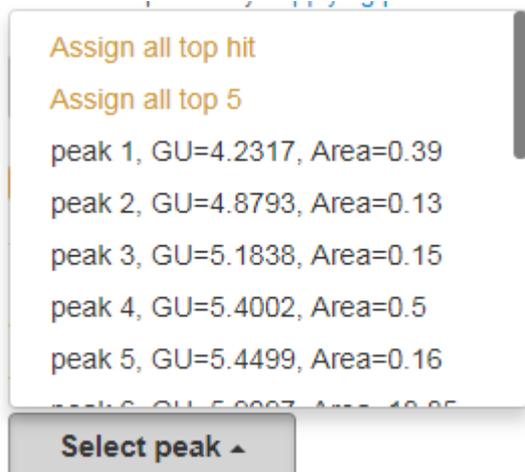
## 3. Select a glycoprotein (optional) ⓘ

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

Select glycoprotein ▾

Get Glycan List

**Step 9:** In this tutorial we will assign all peaks to the chromatogram. Select 'Assign all top 5' in the select peak drop up menu seen below.



**Step 10:** Click the button  and wait for the assignment to complete. This computation is quite demanding and it may take 10-20minutes to return assignments. However, your PC is not being used the calculations are taking place on our servers at new England Biolabs. Note that we did not 'Select glycoprotein' on the input page.

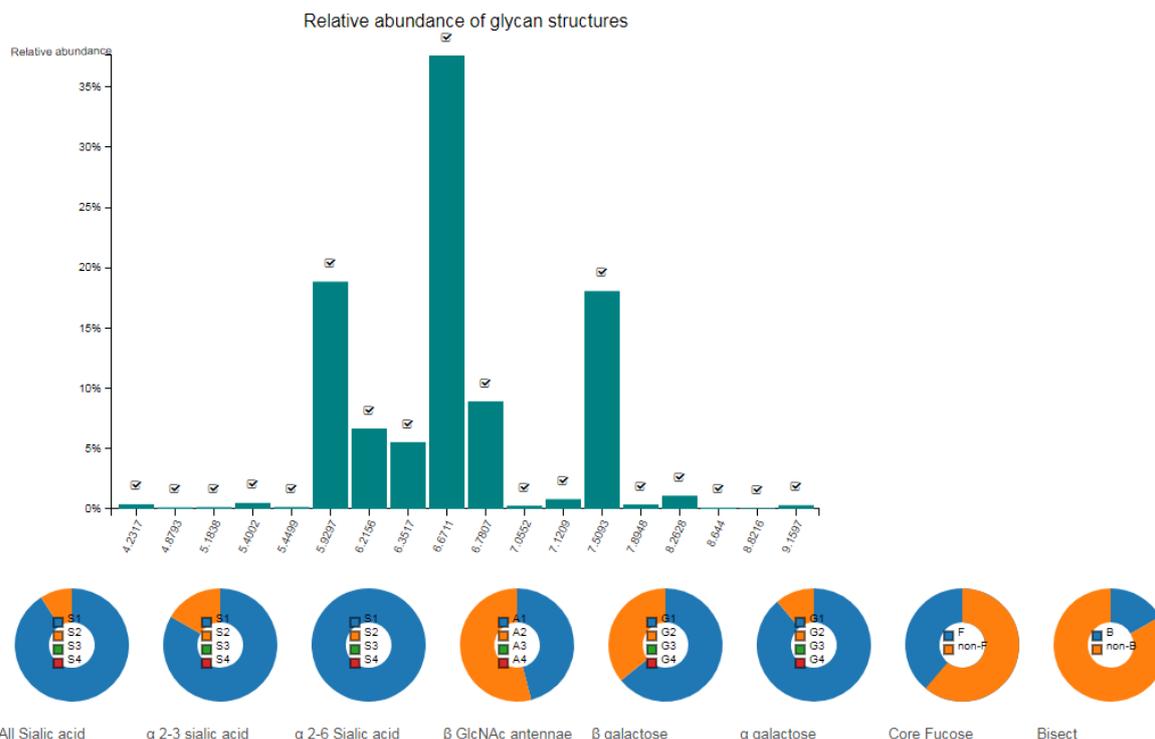
**Step 11:** Figure 4 shows assigned N-glycans summary page.

## Chromatogram Summary

This page summarizes the currently assigned N-glycan peaks.

Score threshold:  Include:  High confidence only  High abundant peaks only Update assignments

Peak analysis status:  
 18 of 18 peaks (100%).  
 N-glycan abundance accounted for: 100.02 %.



**Figure 4.** . The summary output page for the monoclonal antibody. This view is returned when all peaks are computed. 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 annotated in the accepted peaks. If we take as an example Peak 13 and 14. 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.

**Step 12:** Each peak now has an assignment and a corresponding table of ranked glycans. In fact, there should be at most 5 assignments all ranked by the score in each peak. However, not all 5 should be accepted by the user – see step 13.

**Step 13:** The glycans ranked in the bottom 2 of peak 15 in (Figure 5) should definitely be rejected as the mass was not matched very well. Once they are rejected by clicking the 'x' button all the

statistics are updated. Since there is no mass spectra uploaded no mass spectra images can be generated the best we can do is match the mass supplied by the user. This can be seen in the 'Details' column (Figure 5). Note that the masses are not exactly matched and may differ slightly because the mass in our dataset is the isotopic average mass. Secondly, the M7 mannose structure ranked 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup>, in Figure 5, are all isomers and perhaps one can accept that it is either M7 or M7 D1 since the GU values are very close to the ones observed in the database. We could reject M7 D3.

Peak 14 at 7.8948 GUs

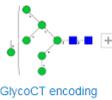
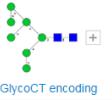
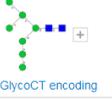
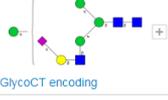
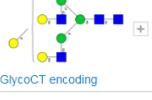
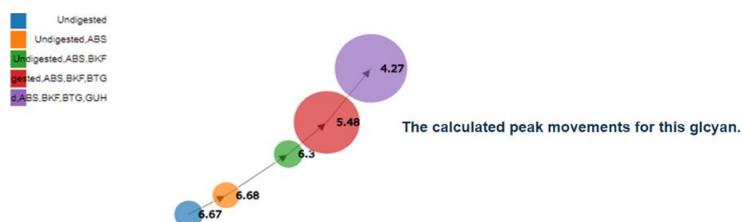
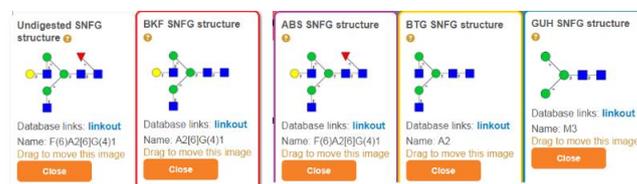
Reject	Score	Oxford notation	Diagram	Details	More information	Evidence
<input type="checkbox"/>	0	M7		Mass (Observed / Expected): 1678.61 / 1679.46 GU (Observed / Expected): 7.8948 / 7.8983 ΔGU :0.0035	 1 2 3 4 5 GlycoSide 1	<input type="radio"/> Mass <input type="radio"/> Shifts <input type="radio"/> Glucose units
<input type="checkbox"/>	0.07	M7 D1		Mass (Observed / Expected): 1678.61 / 1679.46 GU (Observed / Expected): 7.8948 / 7.9685 ΔGU :0.0737	 1 2 3 4 5 GlycoSide 1	<input type="radio"/> Mass <input type="radio"/> Shifts <input type="radio"/> Glucose units
<input type="checkbox"/>	0.13	M7 D3		Mass (Observed / Expected): 1678.61 / 1679.46 GU (Observed / Expected): 7.8948 / 7.763 ΔGU :0.1318	 1 2 3 4 5 GlycoSide 1	<input type="radio"/> Mass <input type="radio"/> Shifts <input type="radio"/> Glucose units
<input type="checkbox"/>	1.02	M4A1G(4)1S(6)1		Mass (Observed / Expected): 1678.61 / 1849.63 GU (Observed / Expected): 7.8948 / 7.916 ΔGU :0.0212	 1 2 3 GlycoSide 1	<input type="radio"/> Mass <input type="radio"/> Shifts <input type="radio"/> Glucose units
<input type="checkbox"/>	1.02	A2G(4)2Ga(3)1		Mass (Observed / Expected): 1678.61 / 1923.71 GU (Observed / Expected): 7.8948 / 7.9122 ΔGU :0.0174	 1 2 GlycoSide 1	<input type="radio"/> Mass <input type="radio"/> Shifts <input type="radio"/> Glucose units

Figure 5. Assignments for peak 14 in our data example. All the glycan masses in peak 14 match except for the bottom ranked two.

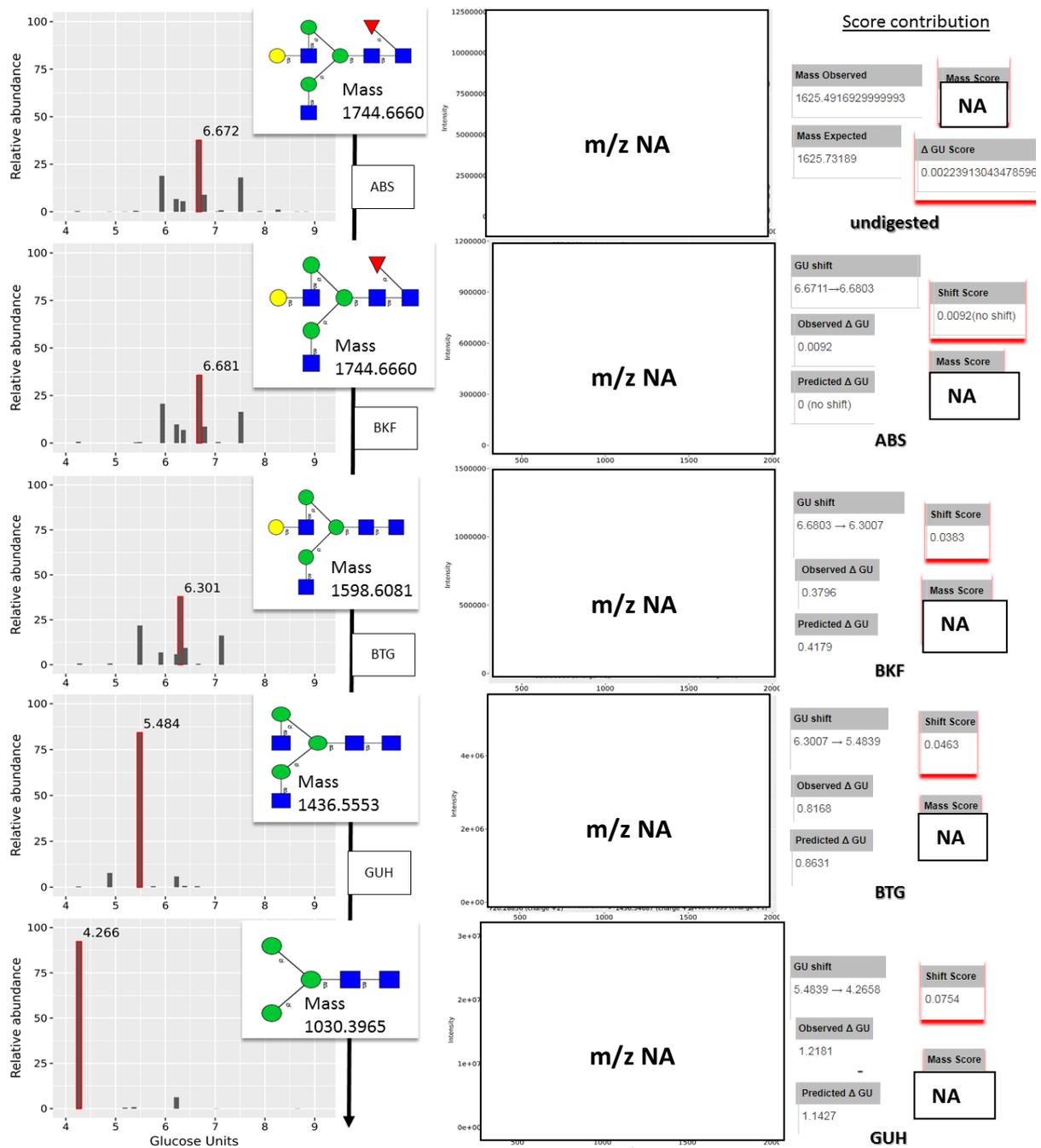
Step 14: The following information is available.

1. If a user clicks on a bar (in Figure 4) they will be presented with a peak shift graph. This presents a directed graph showing the estimated peak movement (Figure 6). 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 6.** The directed graph representation of the peak movements.

2. 'Score calculation' (Figure 6): This is perhaps the most important information revealed by GlycanAnalyzer. It reveals the peak shift in the H/UPLC chromatograms. However, because the user supplied their own observed masses then there will be no images generated for the m/z shifts. Also, the score does not take mass into account.



**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 = NA, Shift score= 0.0092 + 0.0383 + 0.0463 + 0.0754. Total score is 0.0022 + 0 + 0.0092 + 0.0383 + 0.0463 + 0.0754.

