Tutorial 3

Aim: Guides the user though the input process when they have no mass information. The only information is the peaks (GU and area) from the UPLC. Accuracy will be lost in this situation. Figure 1 shows the monoclonal antibody exoglycosidase array which we will work with.

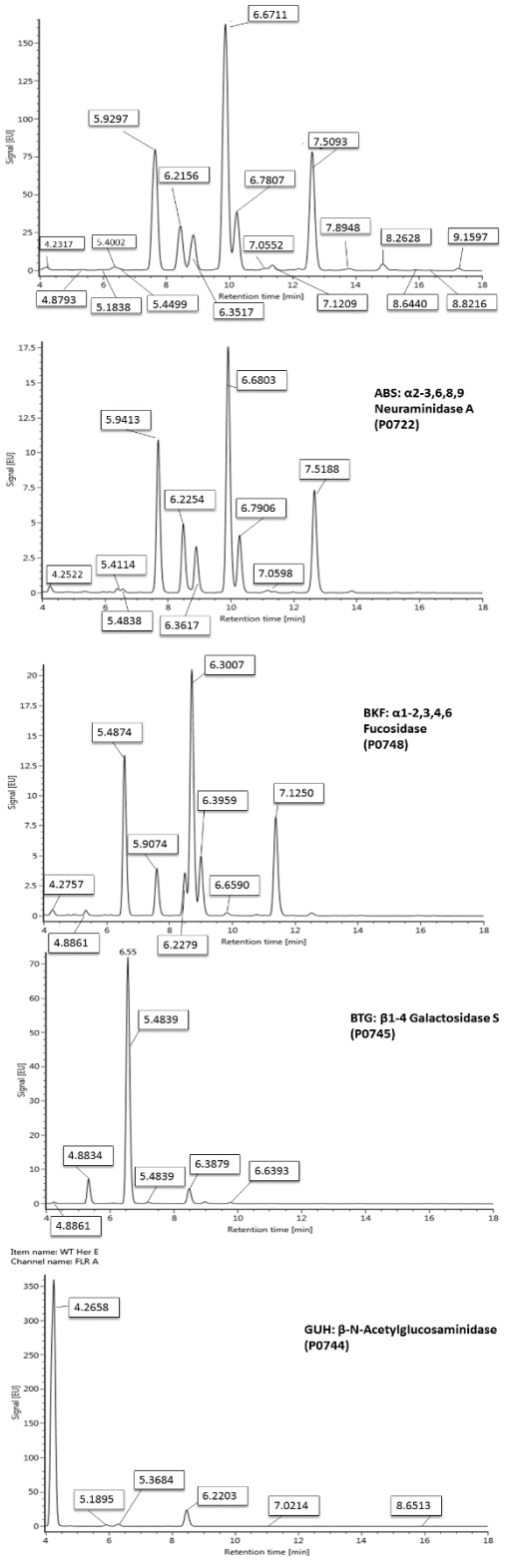


Figure 1. An ABS+BKF+BTG+GUH exoglycosidase array (UPLC). There are three variables: retention time, area of each peak and Glucose Units (boxes). Notice that some peaks have very small size (low abundance) and some are large (high abundance).

Abbreviations:

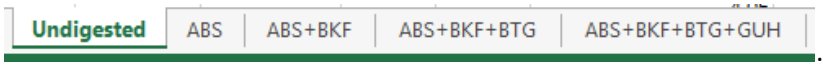
UPLC: Ultra-performance liquid chromatography

LC: Liquid chromatography

MS: Mass spectrometry

GU: Glucose Units

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



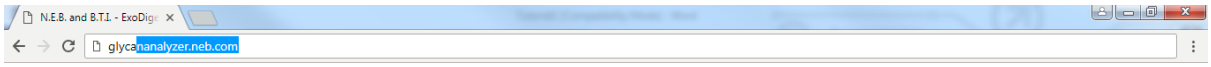
Note: if users are supplying their own datasets then they must have it in exactly the same format as Data_Expert_nomass.xls

Step 2: Notice that the LC data (Fig 1) is tabulated in the spreadsheet under the columns "GU", "Amount (%)", and the MS summarized in the tabs "Observed Mass" and "Observed charge" are not available (NA):

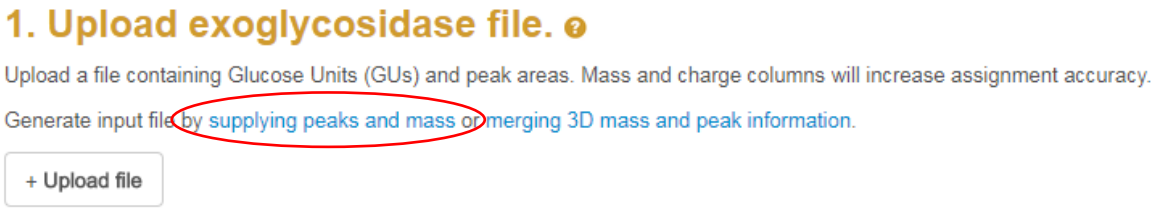
GU	Amount (%)	Observed Mass	Observed charge
4.2317	0.39	NA	NA
4.8793	0.13	NA	NA
5.1838	0.07	NA	NA
5.4002	0.5	NA	NA
5.4499	0.16	NA	NA
5.9297	16.42	NA	NA
5.9297	16.42	NA	NA
6.2156	6.67	NA	NA
6.3517	5.54	NA	NA
6.6711	37.61	NA	NA
6.7807	8.92	NA	NA
7.0552	0.27	NA	NA
7.1209	0.55	NA	NA
7.1209	0.55	NA	NA
7.5093	18.08	NA	NA
7.8607	0.11	NA	NA
7.8948	0.26	NA	NA
8.2628	1.1	NA	NA
8.644	0.09	NA	NA
8.6902	0.05	NA	NA
8.8216	0.07	NA	NA
9.1597	0.31	NA	NA

Each peak is therefore defined using two variables Glucose Units and % area.

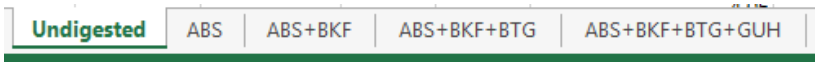
Step 3: Load GlycanAnalyzer web application in your browser. Type http://glycananalyzer.neb.com



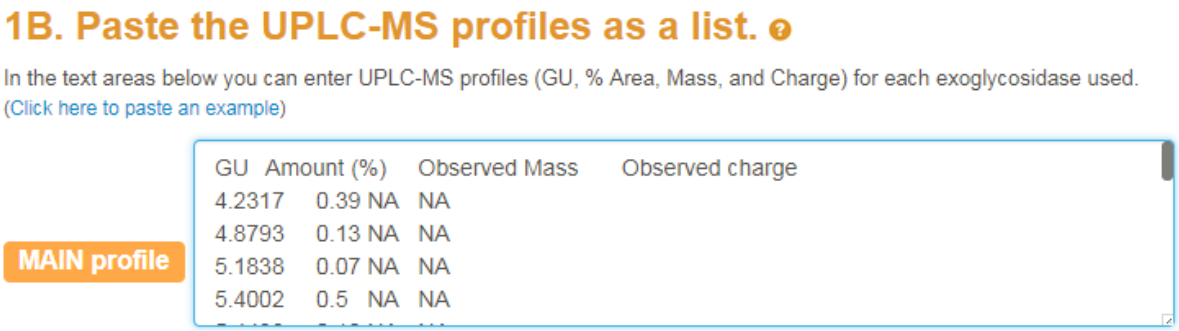
Step 4: We want to generate input for GlycanAnalyzer. Click the link circled on the main page:



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



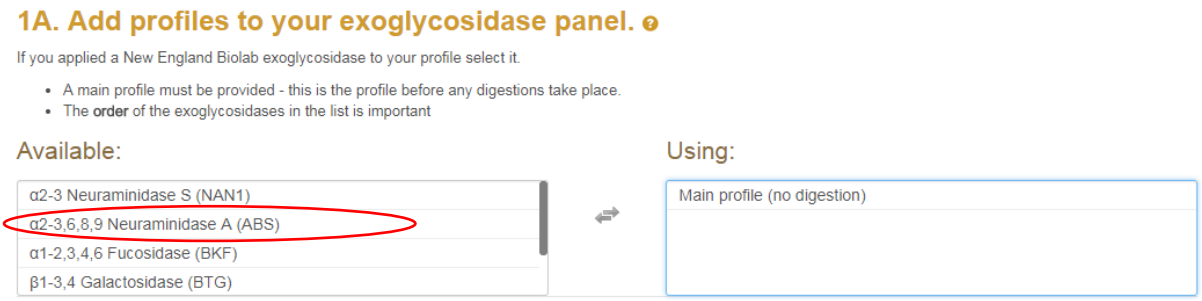
Copy the undigested data into the main profile. Note that the columns GU, %area, mass and charge should be separated by tabs:

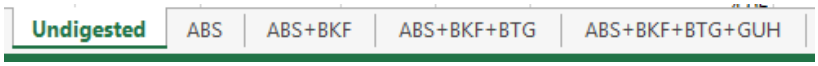


This is the main profile on which exoglycosidase digestions will be applied and glycan annotations will be assigned.

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.

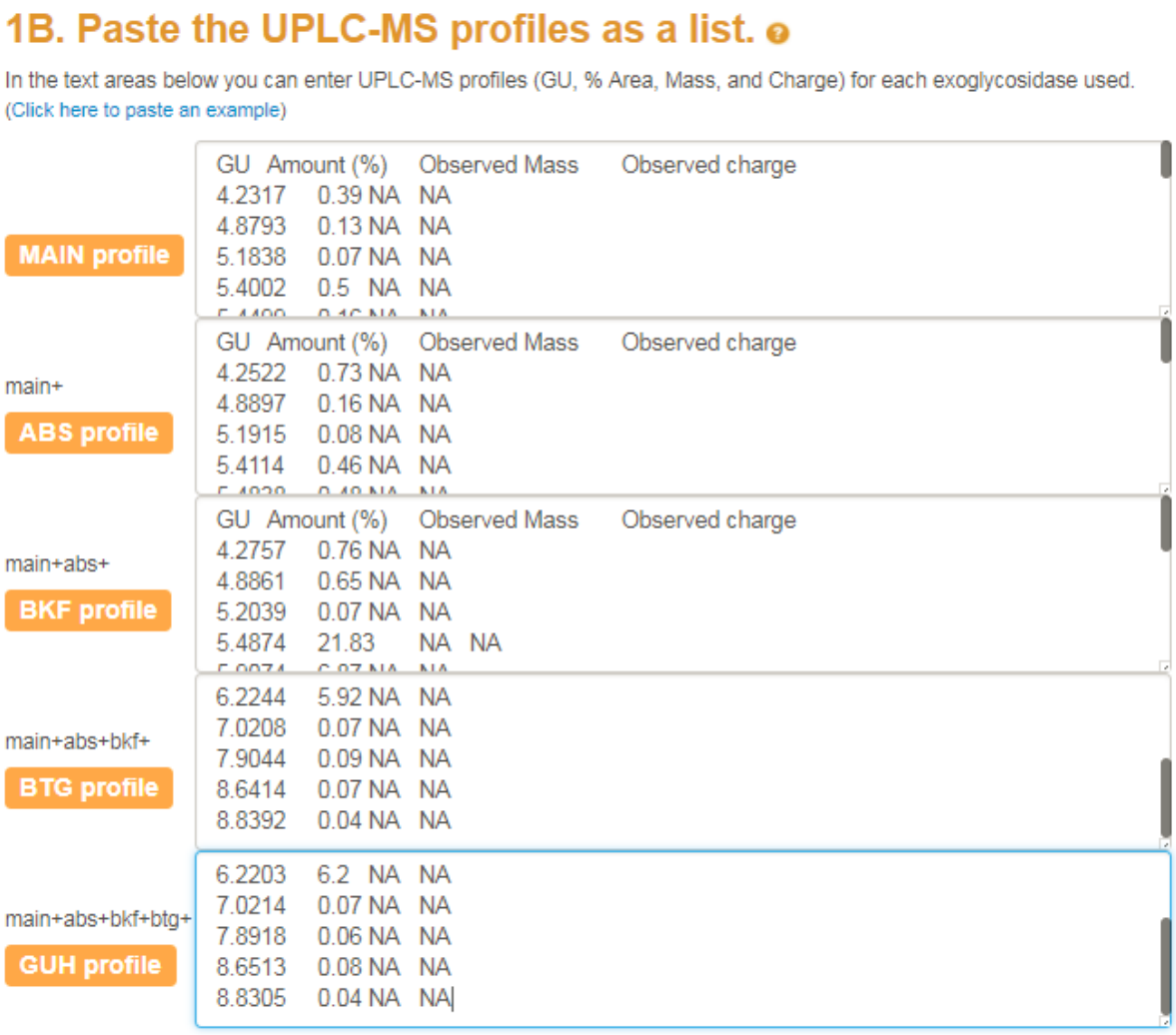
Step 6: Add data from the spreadsheet by adding exoglycosidases **in order** of application shown in Figure 1 (order ABS, BKF, BTG and GUH). To add ABS click here:

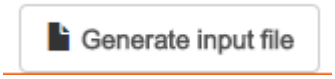


Then add the data (see step 5) from the spreadsheet in the ABS panel: 

. Copy and paste from the spreadsheet and repeat for BKF, BTG and GUH enzymes as ordered in Figure 1.

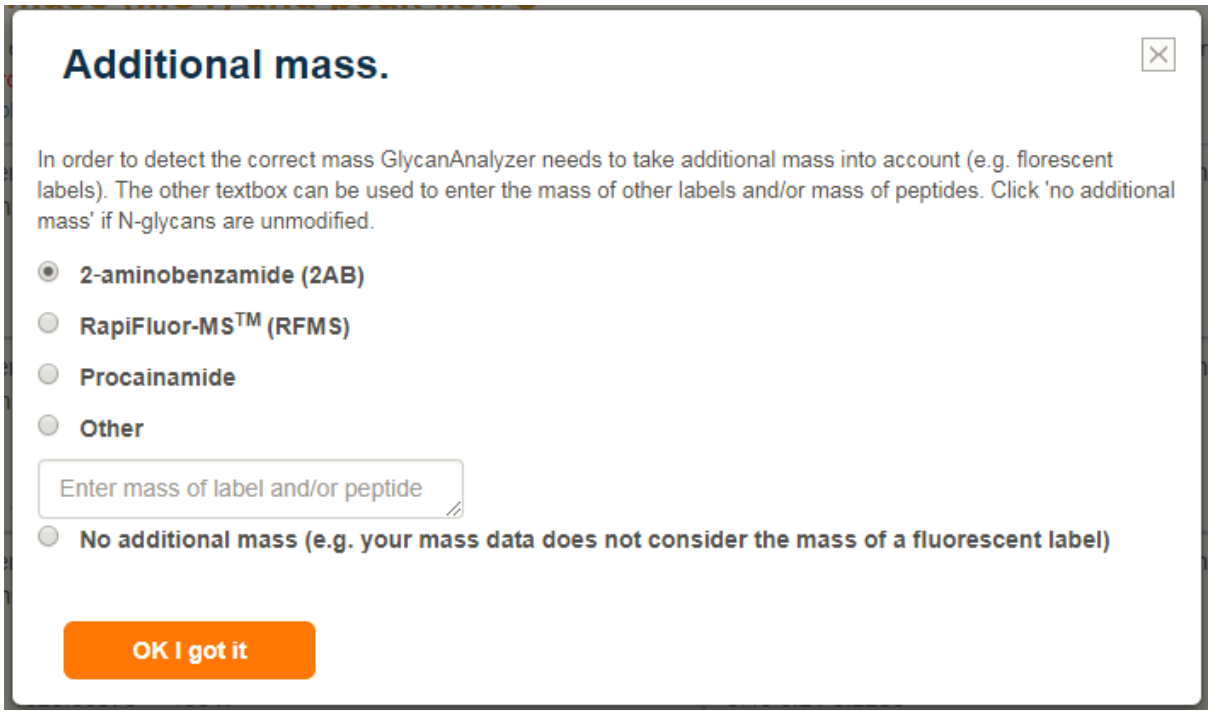
Step 7: The final data page should resemble the following:



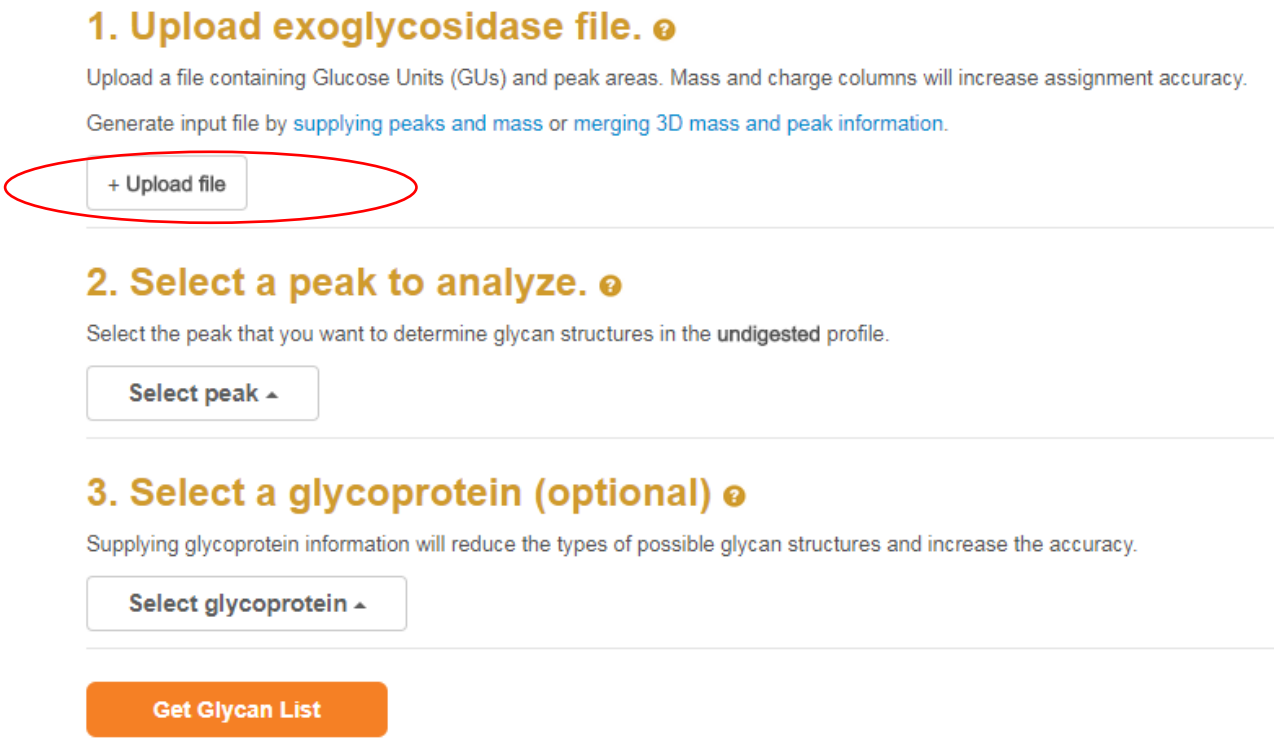
Click the button 

and save the file to your PC.

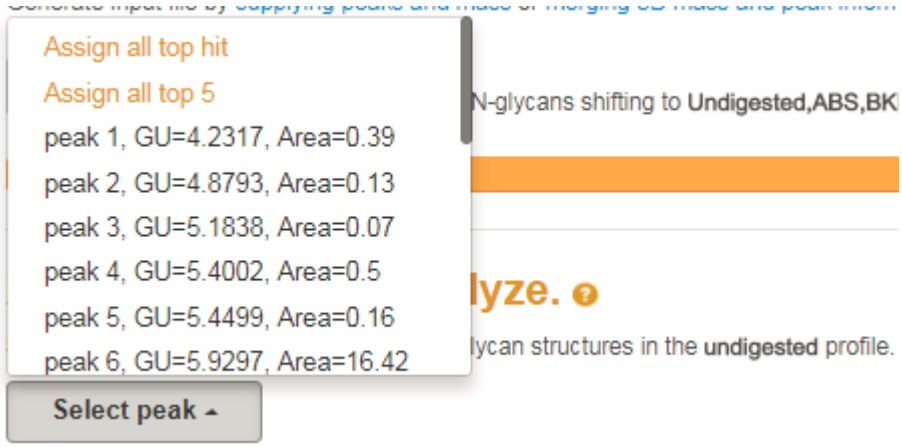
After clicking ‘Merge data to input file’ the user is asked if their MS data contains a florescent label. Users should select one of three things: (i) select a label supported by GlyanAnalyzer (2AB, RFMS, procanaminde), (ii) select ‘other’ and enter the labels mass explicitly or (iii) select label free.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



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



Step 9: In this tutorial we will assign N-glycans to peak. Select 'peak 9' in the select peak drop up menu seen below.



Step 10: Don't select any glycoprotein and Click the button 

and wait for the assignment to complete. This computation should only take 3-4 minutes. However, your PC is not being used the calculations are taking place on our servers at new England Biolabs.

Step 11: Figure 4 shows the top 3 N-glycans returned for peak 9. This step will show the user how to interpret the output. The returned N-glycans are ranked by a score (closer to zero the better – Figure 4a). The user has the option to reject any candidate N-glycan (Figure 4b). N-glycans are displayed in oxford notation (Figure 4c) and drawn in SNFG notation (Figure 4d – click the '+' symbol for higher resolution glycan image). The details column gives some clues to how correct the assignment might

be (Figure 4e), for example how close the theoretical mass are and GU are to the predicted mass and observed GU values. Figure 4g is described in step 12 and Figure 4h is explained in step 13.

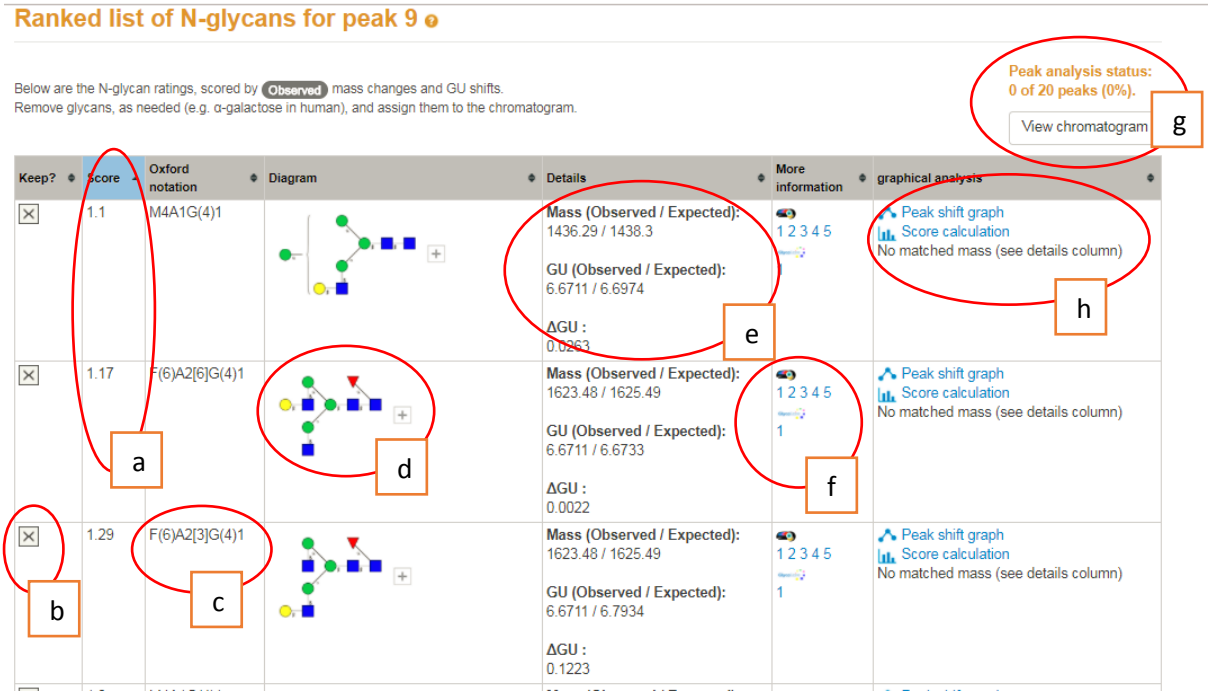


Figure 4. The single peak assignment output.

Step 12: If we are 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 4h can be broken into three stages

1. '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.

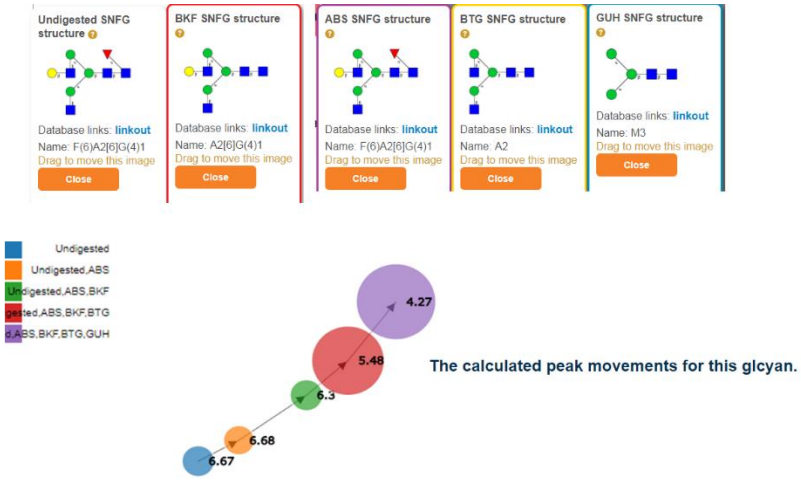


Figure 5. The directed graph representation of the peak movements.

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

- The peak shift on the digestion array chromatograms.

However, because the user supplied no observed masses then there will be no information on the observed m/z shifts. Also, the score does not take mass into account.

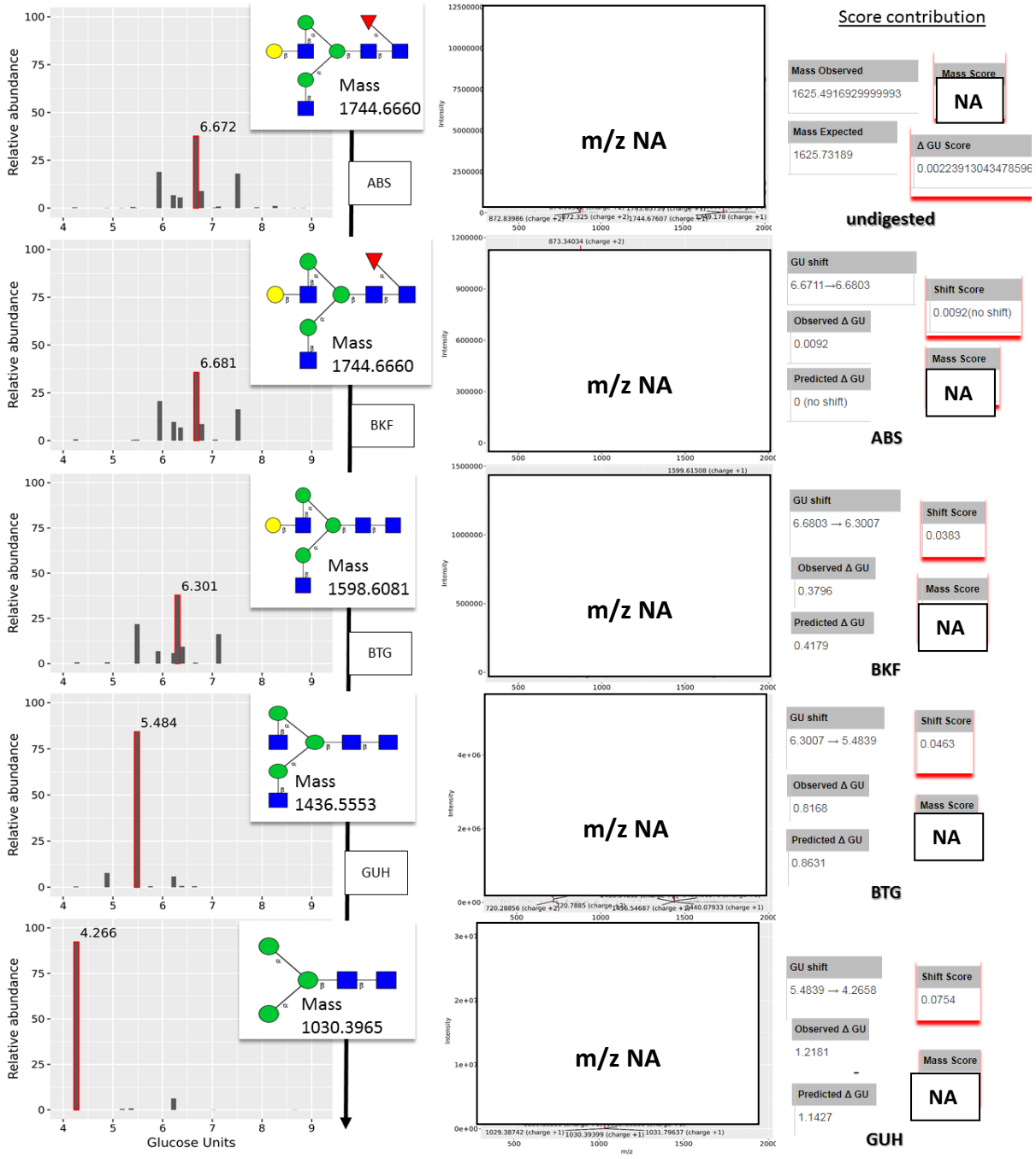


Figure 6. Score calculation for peak 9 of the monoclonal antibody. The total score is 0.1714 coming from the following contributions: ΔGU = 0.0022, Mass score = NA, Shift score= 0.0092 + 0.0383 + 0.0463 + 0.0754. Total score is 0.0022 + 0 + 0.0092 + 0.0383 + 0.0463 + 0.0754.