

Tutorial 3

Aim: Guides the user through 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.

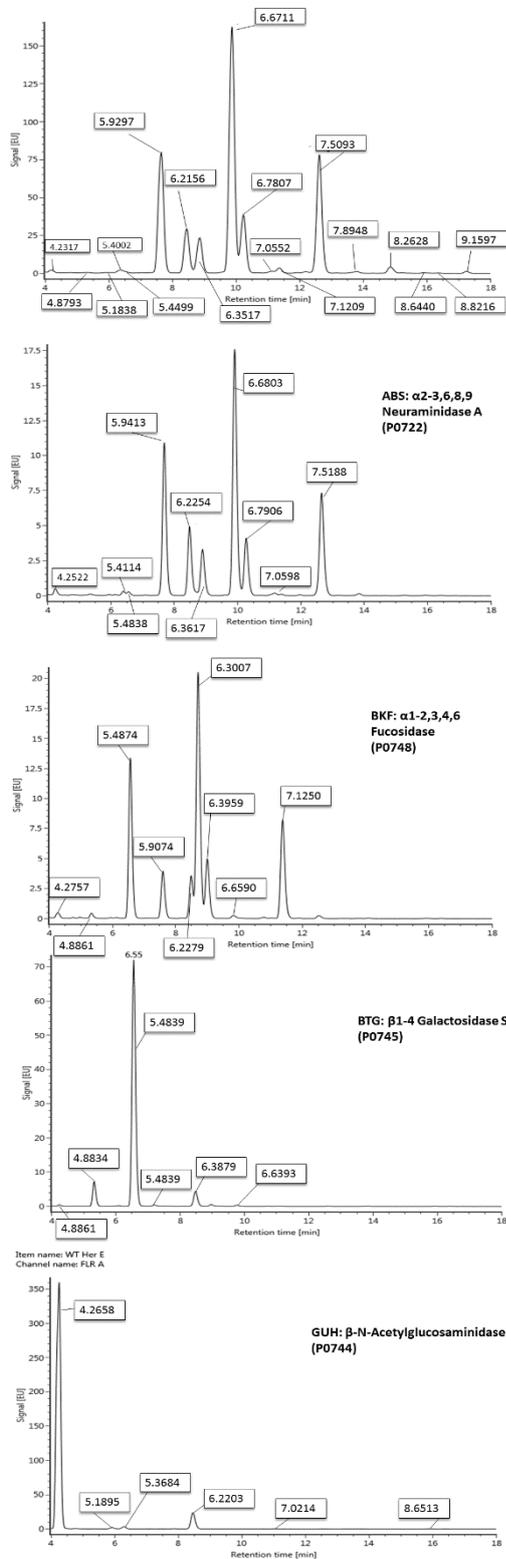


Figure 1. An ABS+BKF+BTG+GUH exoglycosidase array (UPLC). There are three variables: retention time, area of each peak and Glucose Units (boxes). Notice that some peaks have very small size (low abundance) and some are large (high abundance).

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:

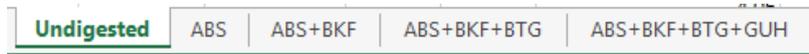
1. Upload exoglycosidase file. ⓘ

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 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:

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

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

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:

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:

<input type="checkbox"/> α2-3 Neuraminidase S (NAN1)
<input checked="" type="checkbox"/> α2-3,6,8,9 Neuraminidase A (ABS)
<input type="checkbox"/> α1-2,3,4,6 Fucosidase (BKF)
<input type="checkbox"/> β1-3,4 Galactosidase (BTG)

Using:

Main profile (no digestion)

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

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

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

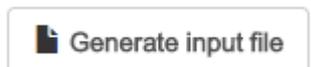
Step 7: The final data page should resemble the following:

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)

	GU	Amount (%)	Observed Mass	Observed charge
MAIN profile	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
main+ ABS profile	4.2522	0.73	NA	NA
	4.8897	0.16	NA	NA
	5.1915	0.08	NA	NA
	5.4114	0.46	NA	NA
	5.4829	0.19	NA	NA
main+abs+ BKF profile	4.2757	0.76	NA	NA
	4.8861	0.65	NA	NA
	5.2039	0.07	NA	NA
	5.4874	21.83	NA	NA
	5.9974	0.07	NA	NA
main+abs+bkf+ BTG profile	6.2244	5.92	NA	NA
	7.0208	0.07	NA	NA
	7.9044	0.09	NA	NA
	8.6414	0.07	NA	NA
	8.8392	0.04	NA	NA
main+abs+bkf+btg+ GUH profile	6.2203	6.2	NA	NA
	7.0214	0.07	NA	NA
	7.8918	0.06	NA	NA
	8.6513	0.08	NA	NA
	8.8305	0.04	NA	NA



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 GlycanAnalyzer (2AB, RFMS, procainamide), (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: 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:

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 N-glycans to peak. Select 'peak 9' in the select peak drop up menu seen below.

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

Assign all top hit
Assign all top 5
peak 1, GU=4.2317, Area=0.39
peak 2, GU=4.8793, Area=0.13
peak 3, GU=5.1838, Area=0.07
peak 4, GU=5.4002, Area=0.5
peak 5, GU=5.4499, Area=0.16
peak 6, GU=5.9297, Area=16.42

Select peak ▾

N-glycans shifting to Undigested,ABS,BK

lyze. ?

glycan structures in the **undigested** profile.

Step 10: Don't select any glycoprotein and Click the button [Get Glycan List](#) 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.

Ranked list of N-glycans for peak 9

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.

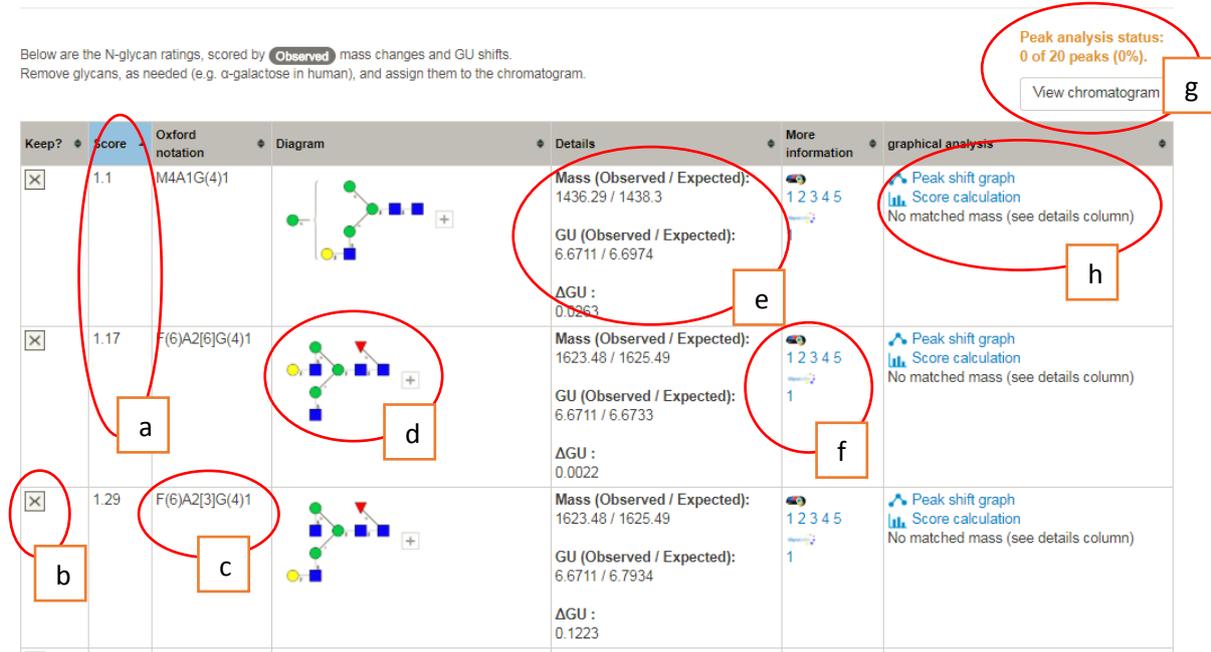


Figure 4. The single peak assignment output.

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.

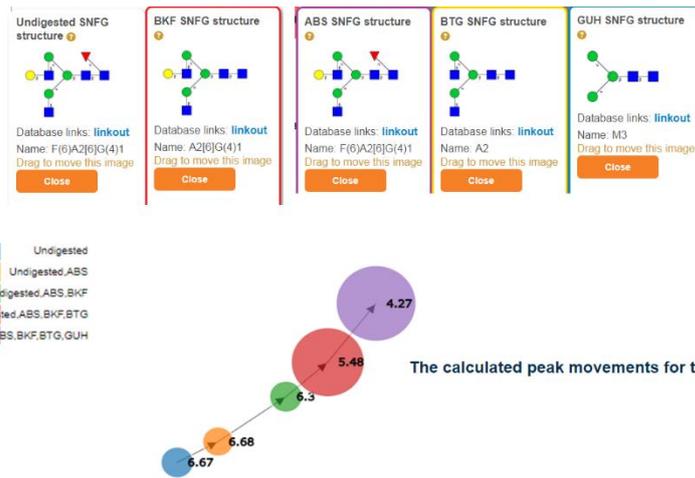


Figure 5. The directed graph representation of the peak movements.

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.

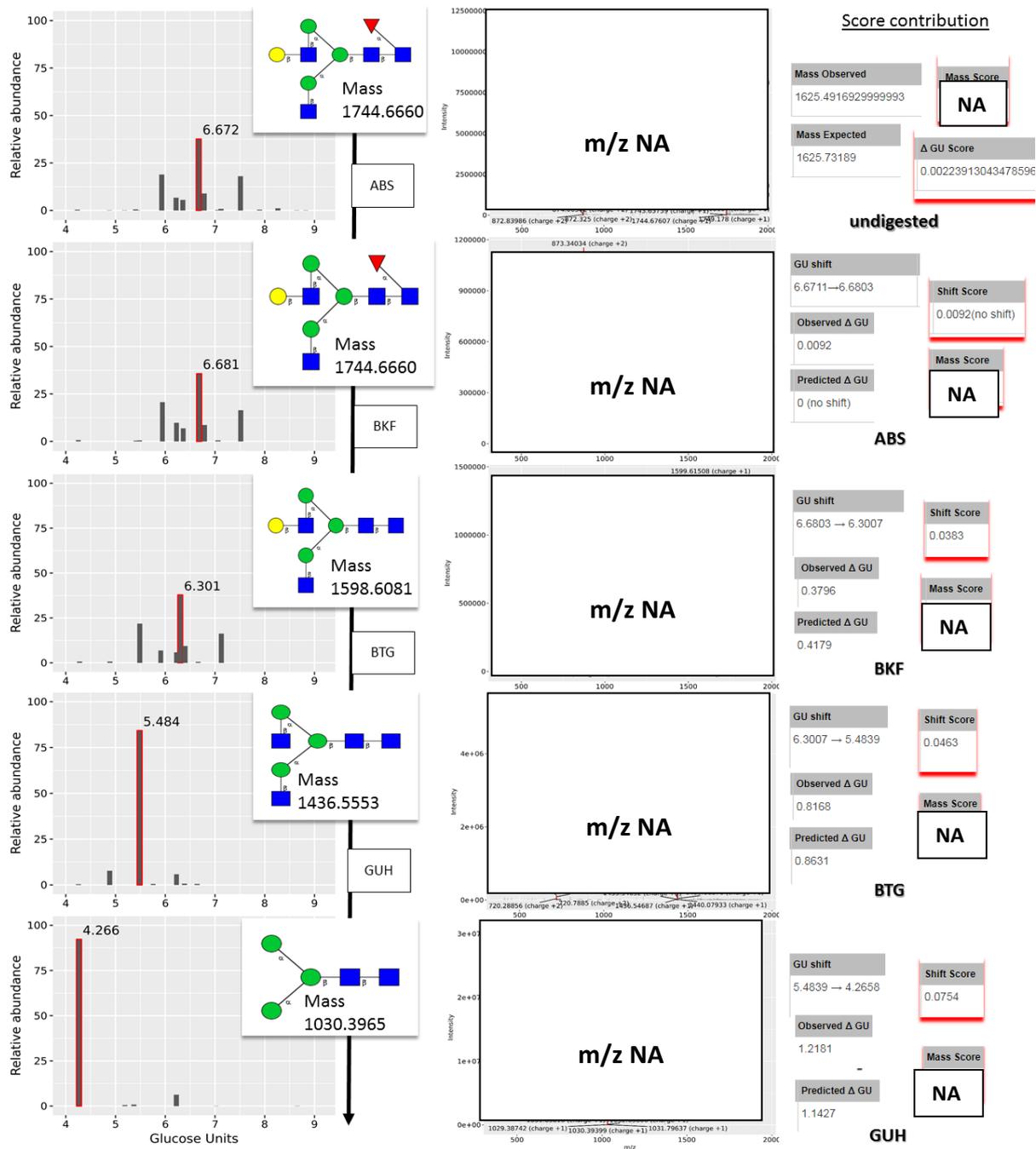


Figure 6. Score calculation for peak 9 of the monoclonal antibody. The total score is 0.1714 coming from the following contributions: Δ GU = 0.0022, Mass score = NA, Shift score = 0.0092 + 0.0383 + 0.0463 + 0.0754. Total score is 0.0022 + 0 + 0.0092 + 0.0383 + 0.0463 + 0.0754.