GPC reader is an R Shiny application for routine calculations of Mp, average molecular weights (Mn and Mw), and PDI from GPC chromatograms (Gel permeation chromatography/ size exclusion chromatography (SEC)). You can upload multiple files, inducing calibration files, and save results as csv file.

All calculations are based on the formulas given in [1].

$$Mn = \frac{\sum H_i}{\sum \frac{H_i}{M_i}} \quad Mw = \frac{\sum H_i M_i}{\sum H_i} \quad PDI = \frac{Mw}{Mn}$$

Mn number-average molecular weight;
Mw weight-average molecular weight;
PDI polydispersity index;
H_i height at the i-th point (detector signal intensity);
M_i molecular weight at the i-th point.

Mp is calculated as

$$Mp = 10^{(t * slope + intercept)}$$ for a linear calibration curve

$$Mp = 10^{(t^3 * coef\,3 + t^2 * coef\,2 + t * coef\,1 + intercept)}$$ for a polynomial calibration curve

t is retention time (or retention volume).

App plots the molecular weight distribution (MWD), specifically differential MWD.

dW/d(log M_i) is calculated as

$$\frac{dW_i}{d(\log M_i)} = \frac{dW_i}{dV_i} \frac{dV_i}{d(\log M_i)}$$

Where

$$\frac{dW_i}{dV_i} = \frac{H_i}{\sum \frac{H_i}{\triangle V}}$$

And △V is constant time (volume) interval.

Sidebar and Tabs

The application has a sidebar and two tabs.

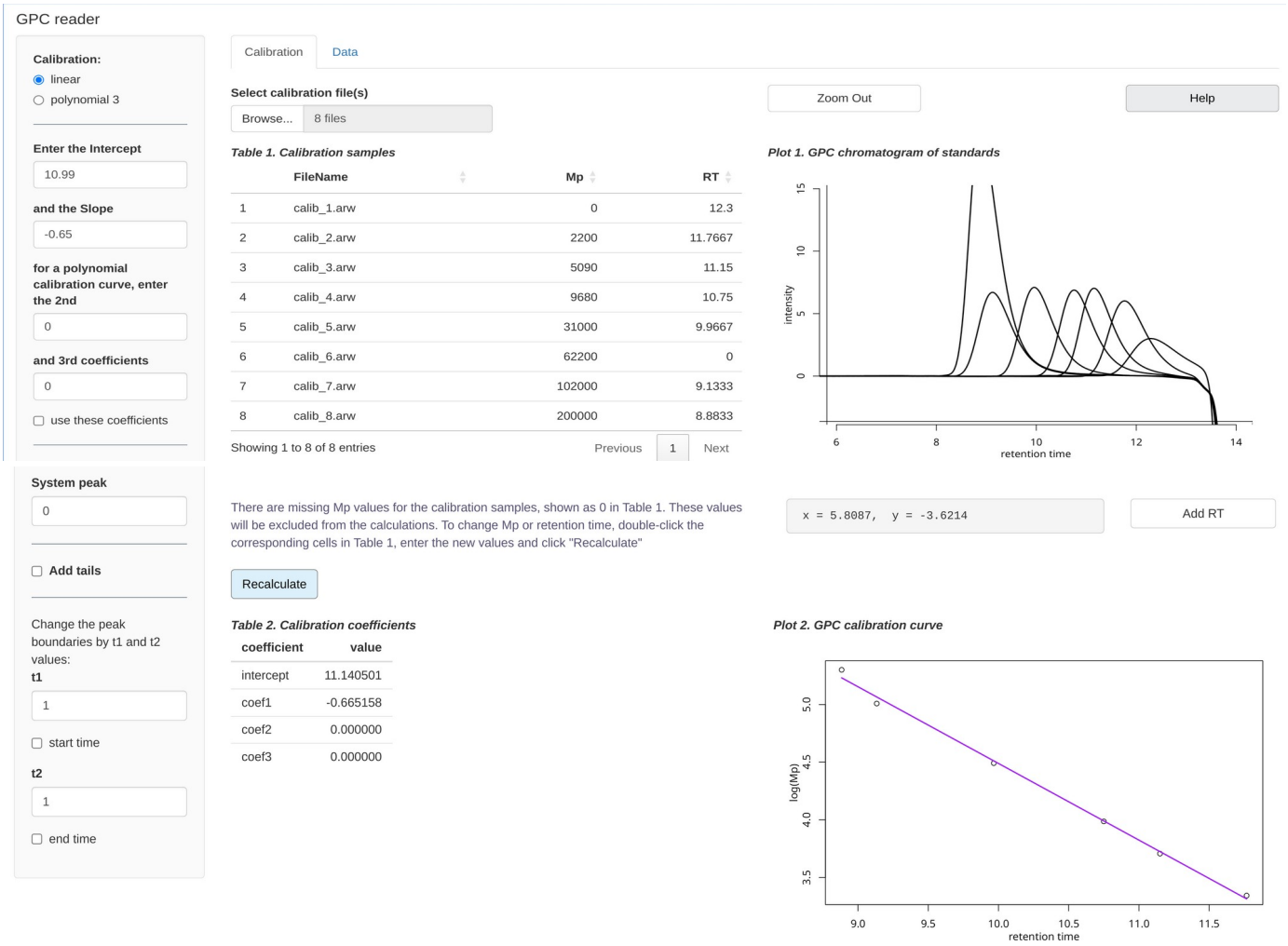
The sidebar panel and the first tab "Calibration" are intended for setting the coefficients of the calibration curve. On the first tab, you will have two tables and two plots:

1) ***Table 1. Calibration samples.*** This table shows the names of the corresponding files, Mp values and peak retention times for each calibration standard.

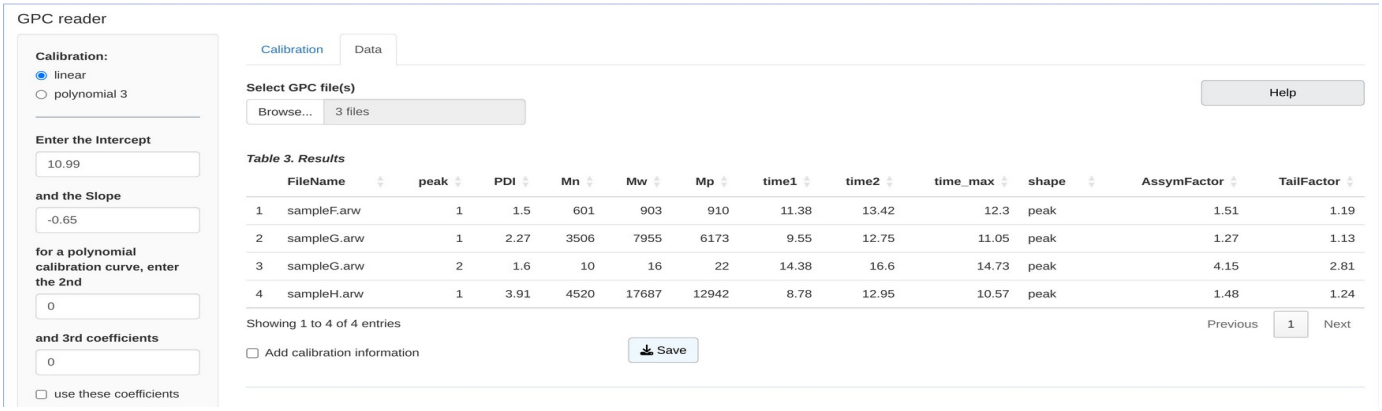
2) ***Plot 1. GPC chromatogram of standards.*** Detector signal intensity versus time.

3) ***Table 2. Calibration coefficients.*** It shows the calibration plot coefficients calculated from calibration files or entered in the sidebar

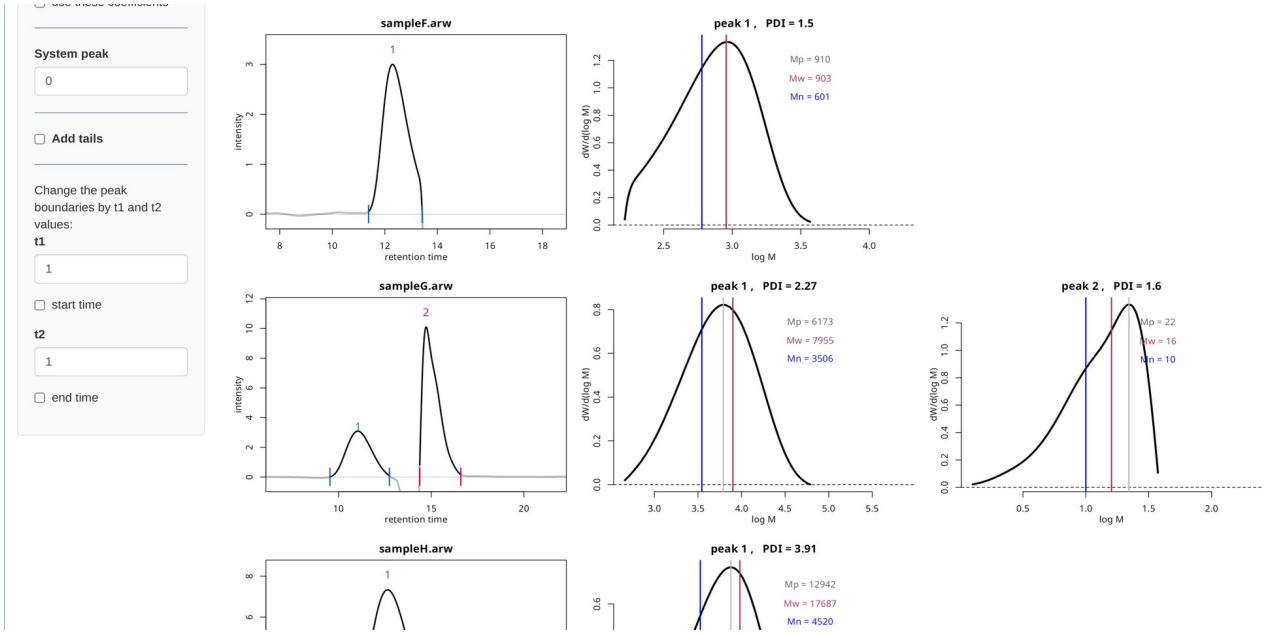
4) ***Plot 2. GPC calibration curve.*** The logarithm of the molecular weight of polymer standards, vs the retention time (the elution volume).



The second tab "Data" is used to load GPC files and display the results (***Table 3 Results***).

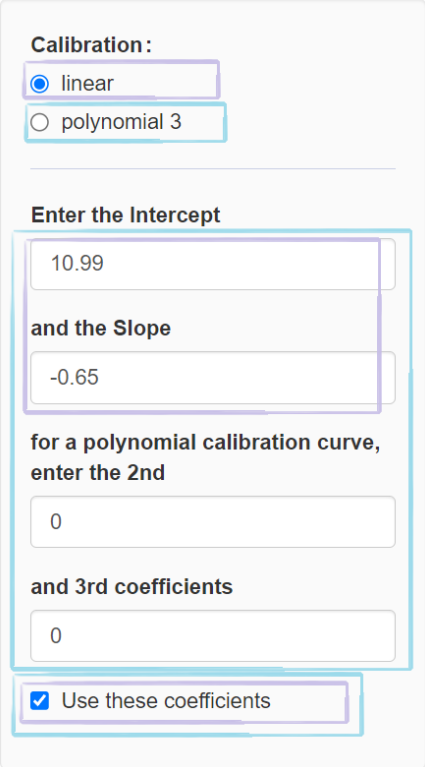


It shows chromatograms and molecular weight distribution (MWD) graphs.



Calibration

There are two ways to set up calibration.



1. If you know the *intercept* and *slope* coefficients, you can enter these values in the corresponding fields in the sidebar. Then check the **Use these coefficients** checkbox.

For linear calibration:

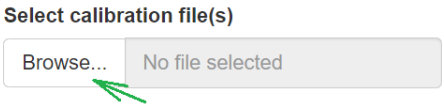
- set the Intercept (it is greater than 0);
- the Slope (it is negative number);
- leave the second and the third coefficients equal to 0.

For a polynomial fit:

- switch the **Calibration** to "***polynomial 3***";
- enter *intercept* and other coefficients;
- note that the coefficients must be very precise;

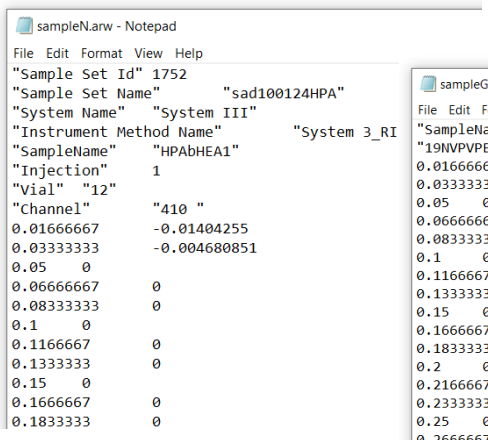
do not round them for polynomial fit.

2. The second way is to load GPS data files with calibration standards on the "Calibration" tab.



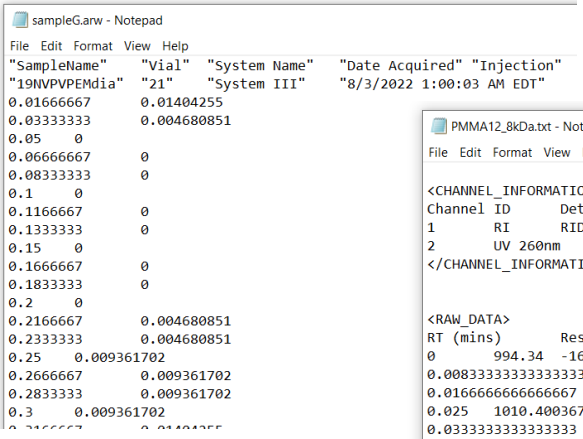
The application does NOT read raw .dat files. You need to export .dat files to .txt, .arw, excel or similar files.

Typical GPC files have tab-separated format with some metadata at the beginning of the file. If you have more than one detector, the application will only use data from the first one.


sampleN.arw - Notepad

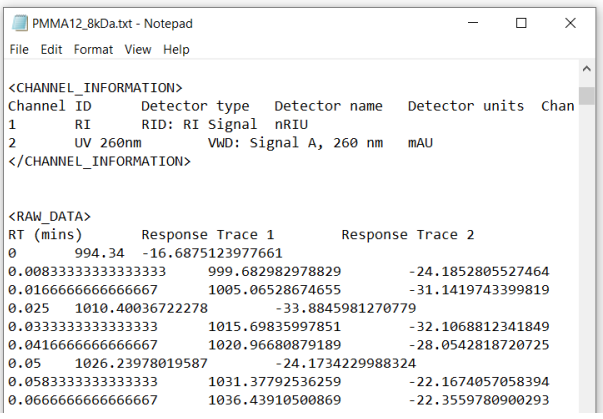
File Edit Format View Help

```
"Sample Set Id" 1752
"Sample Set Name"       "sad100124HPA"
"System Name"   "System III"
"Instrument Method Name"        "System 3_RI
"SampleName"    "HPAbHEA1"
"Injection"     1
"Vial"  "12"
"Channel"       "410 "
0.01666667      -0.01404255
0.03333333      -0.004680851
0.05    0
0.06666667      0
0.08333333      0
0.1     0
0.1166667       0
0.1333333       0
0.15    0
0.1666667       0
0.1833333       0
```


sampleG.arw - Notepad

File Edit Format View Help

```
"SampleName"    "Vial"  "System Name"   "Date Acquired" "Injection"
"19NVPVPEMdia"  "21"    "System III"    "8/3/2022 1:00:03 AM EDT"
0.01666667      0.01404255
0.03333333      0.004680851
0.05    0
0.06666667      0
0.08333333      0
0.1     0
0.1166667       0
0.1333333       0
0.15    0
0.1666667       0
0.1833333       0
0.2     0
0.2166667       0.004680851
0.2333333       0.004680851
0.25    0.009361702
0.2666667       0.009361702
0.2833333       0.009361702
0.3     0.009361702
```


PMMA12_8kDa.txt - Notepad

File Edit Format View Help

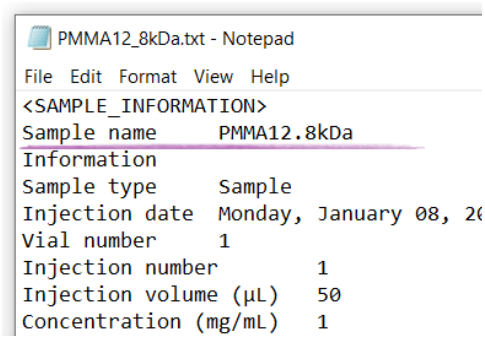
```
<CHANNEL_INFORMATION>
Channel ID      Detector type   Detector name   Detector units  Chan
1       RI      RID: RI Signal  nRIU
2       UV 260nm        VWD: Signal A, 260 nm   mAU
</CHANNEL_INFORMATION>


<RAW_DATA>
RT (mins)       Response Trace 1        Response Trace 2
0       994.34  -16.6875123977661
0.00833333333333333     999.682982978829        -24.1852805527464
0.0166666666666667      1005.06528674655        -31.1419743399819
0.025   1010.40036722278        -33.8845981270779
0.0333333333333333      1015.69835997851        -32.1068812341849
0.0416666666666667      1020.96680879189        -28.0542818720725
0.05    1026.23978019587        -24.1734229988324
0.0583333333333333      1031.37792536259        -22.1674057058394
0.0666666666666667      1036.43910500869        -22.3559780900293
```

There are two options.

2.1. Load multiple files at once if each calibration sample is in a separate file (each calibration standard has been injected into the GPC separately).

Application extracts Mp values for each calibration sample from the ***SampleName*** or ***FileName*** (if no sample name).



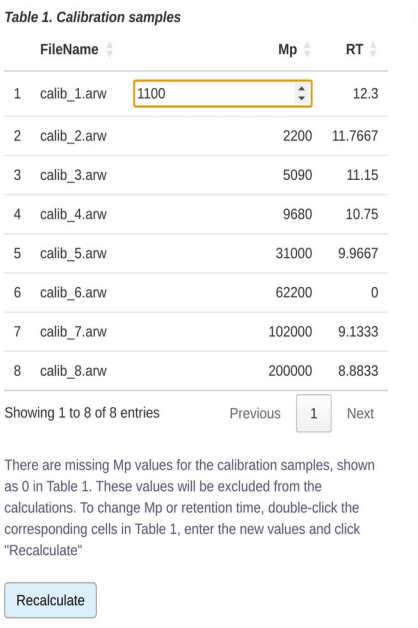
Examples of the sample name and Mp (in Da units):

PMMA12.8kDa	12800
PMMA1100	1100
PMMA31000	31000
PS2.55K	2550
PS200k	200000

These values are listed in the second column of **Table 1**.

If these values are incorrect or missing, double-click the appropriate cells in the **Table 1** and enter the new Mp(s) values.

The retention time values are given in the third column of **Table 1**. If the peak intensity is too low, the peak may not be detected, in which case the RT is set to zero.



If necessary, you can also adjust the peak retention time values (double click on table 1 cells).

Any missing values or values equal to zero will not be used to calculate calibration coefficients.

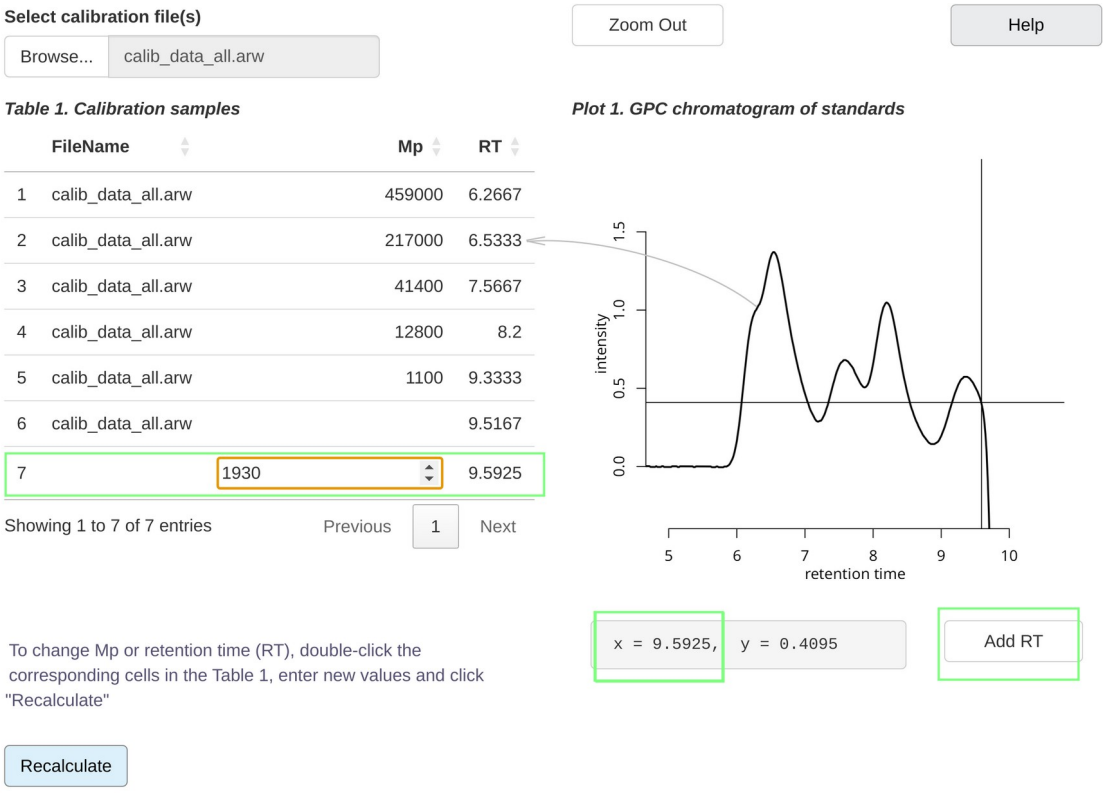
To update the calibration curve and calibration coefficients, click the **"Recalculate"** button.

2.2. Load one file with the one chromatogram if a mixture of calibration standards has been injected into the GPC.

Add the Mp values to the 2nd column of **Table 1** (double-click the corresponding cells) and click **"Recalculate"** button. If there are shoulder peaks, these are also included in **Table 1**. So if a false peak has been detected and you do not want it to be used for the calibration, simply leave the value for Mp blank.

The retention time values can also be added from **Plot 1**. If you click on this plot, the x and y values will appear in the text box below. To add this RT value to the **Table1**, click the **"Add RT"** button and add the Mp value (double-click the table cell).

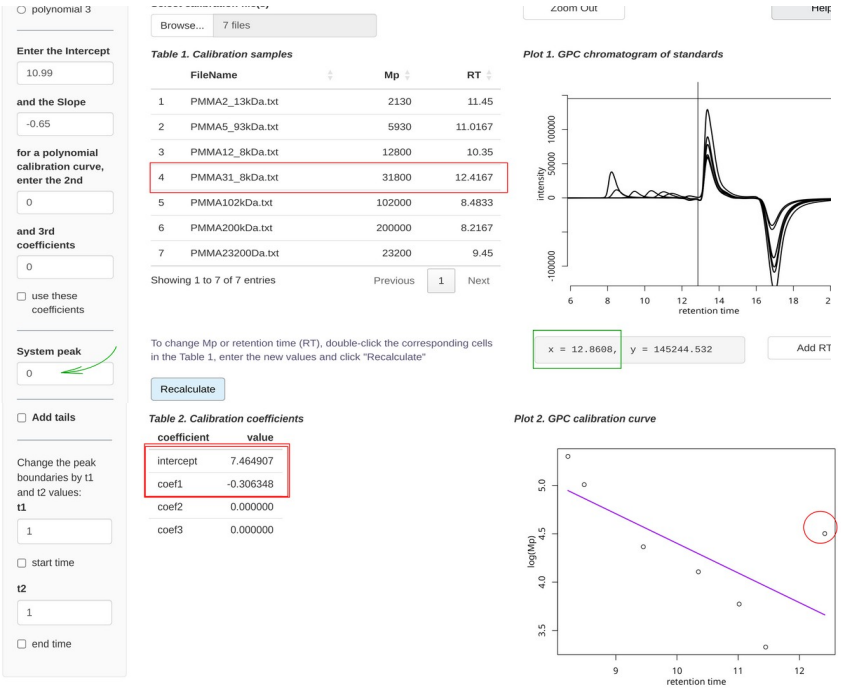
To zoom in on this graph, drag the mouse from top left to bottom right across the graph area. To zoom out click on **"Zoom Out"**.

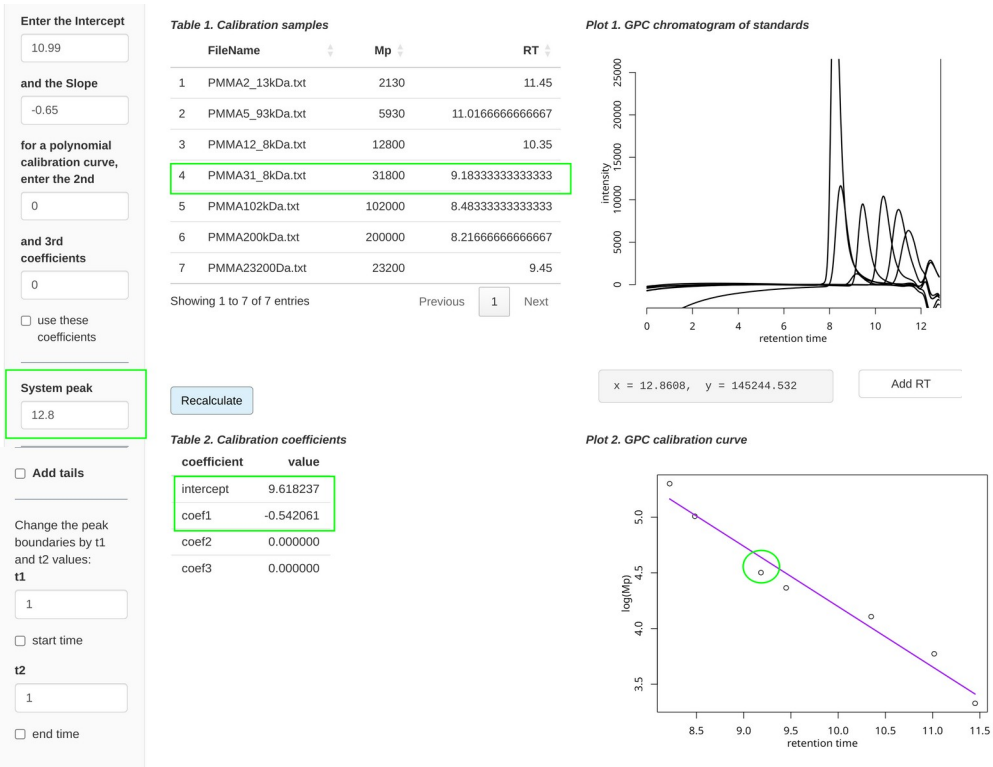


To update the calibration curve and calibration coefficients, click the **"Recalculate"** button.

You can also enter the time of the system peak to improve the calibration if the analyte peak is much smaller than the system peak. When a system peak time is added, all chromatograms, both calibration and sample data, will be clipped at this value. To estimate where the system peak starts, click on **Plot 1** area and type an x value in the box **System peak** in the sidebar.

In this example, sample 4 has low intensity and the system peak was detected instead of the sample peak.





The application will automatically recalculate the calibration. However, if Mp values were missed, you will need to re-enter these values in **Table 1**.

Choose linear or polynomial3 fit for the calibration curve.

Calibration:

- ○ linear
- ◉ polynomial 3

Enter the Intercept

10.99

and the Slope

-0.65

for a polynomial calibration curve, enter the 2nd

0

and 3rd coefficients

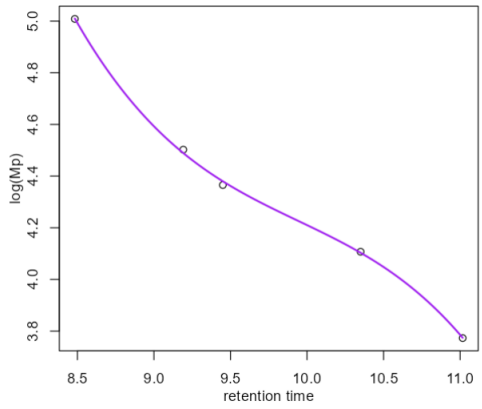
0

☐ Use these coefficients

Table 2. Calibration coefficients

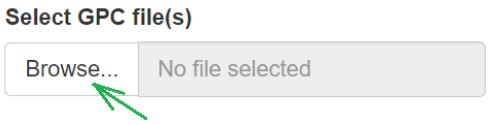
coefficient	value
intercept	124.701962
coef1	-35.790605
coef2	3.571755
coef3	-0.119761

Plot 2. GPC calibration curve



Data tab

Load the GPC file(s) in the second tab.



Results table

Results are summarized in the **Table 3**. For each GPC file (***FileName***) and each peak (***peak***) there are:

- calculated *Mp*, average molecular weights (*Mn* and *Mw*), and *PDI*;
- ***time1*** and ***time2*** are the start and end times of the peak, respectively; these values can be changed with the **t1** and **t2** values on the sidebar;
- ***time_max*** is the retention time at the peak's maximum;
- if more than 10% of a peak's tail is missing, it is labeled as a "shoulder" in the ***shape*** column;
- asymmetry factor (***AssymFactor***) and tailing factor (***TailFactor***) are calculated as given in [2];
- Check **Add calibration information** box to add it to the **Table 3**

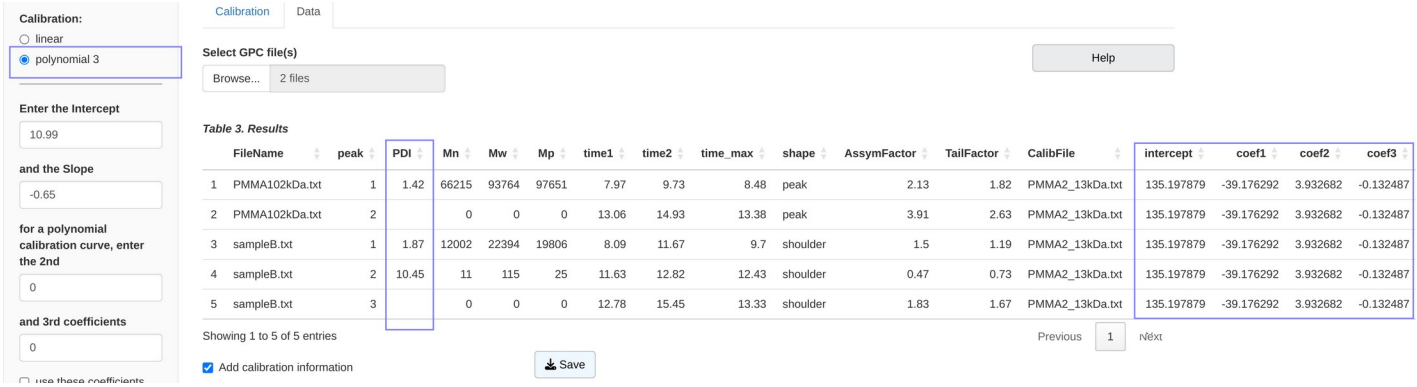
Table 3. Results

	FileName	peak	PDI	Mn	Mw	Mp	time1	time2	time_max	shape	AssymFactor	TailFactor	CalibFile	intercept	coef1
1	PMMA102kDa.txt	1	1.17	86987	102067	113817	7.97	9.73	8.48	peak	2.13	1.82	PMMA2_13kDa.txt	9.739714	-0.552083
2	PMMA102kDa.txt	2	1.22	151	184	227	13.06	14.93	13.38	peak	3.91	2.63	PMMA2_13kDa.txt	9.739714	-0.552083
3	sampleB.txt	1	2.3	12135	27961	24239	8.09	11.67	9.7	shoulder	1.5	1.19	PMMA2_13kDa.txt	9.739714	-0.552083
4	sampleB.txt	2	1.13	769	872	751	11.63	12.82	12.43	shoulder	0.47	0.73	PMMA2_13kDa.txt	9.739714	-0.552083
5	sampleB.txt	3	1.32	160	211	239	12.78	15.45	13.33	shoulder	1.83	1.67	PMMA2_13kDa.txt	9.739714	-0.552083

Showing 1 to 5 of 5 entries Previous 1 Next

Add calibration information Save

In the sidebar you can switch between linear and polynomial calibration. The results will be updated automatically.



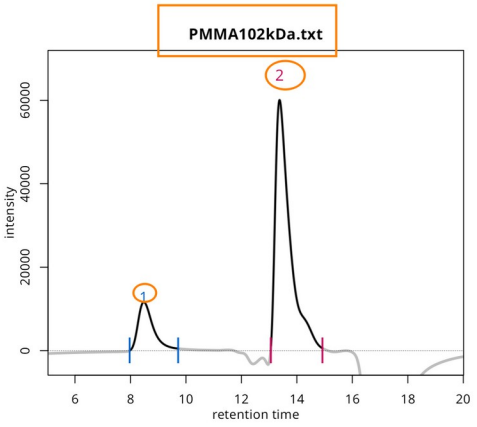
Click on **"Save"** button to save results as *.csv* file

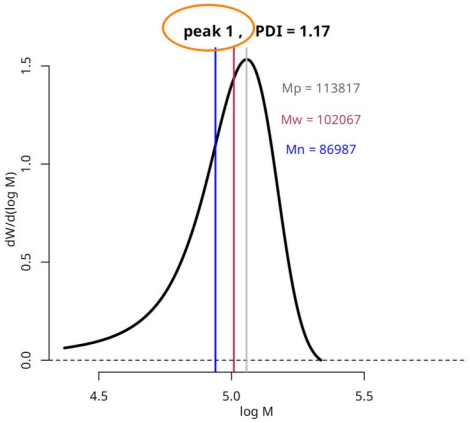
Plots

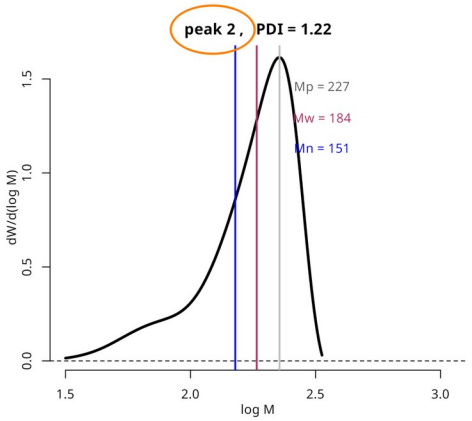
There will be two types of plots for each file.

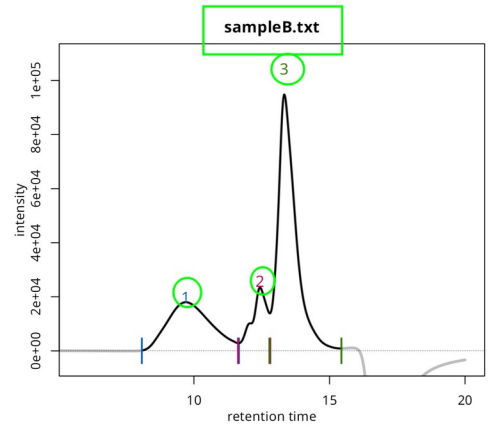
◦ The first is the chromatogram. One graph for each file, where the graph name matches the GPC file name and the peaks are labeled 1, 2, 3, etc. Short vertical lines marking the beginning and end of each peak.

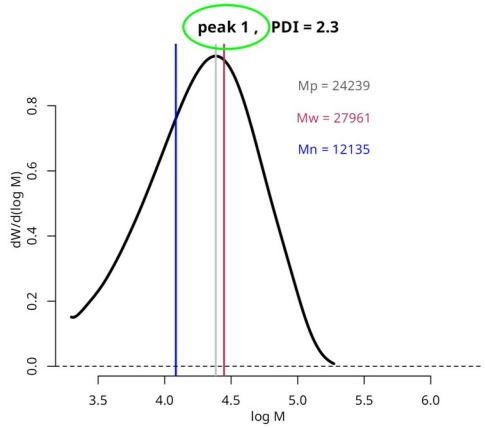
◦ And the second type of plots is the molecular weight distribution. The number of these plots is equal to the number of peaks in the chromatogram.

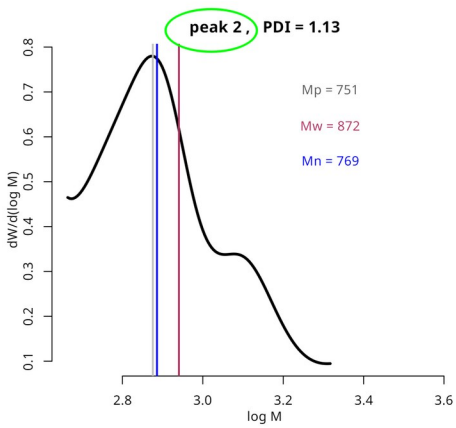


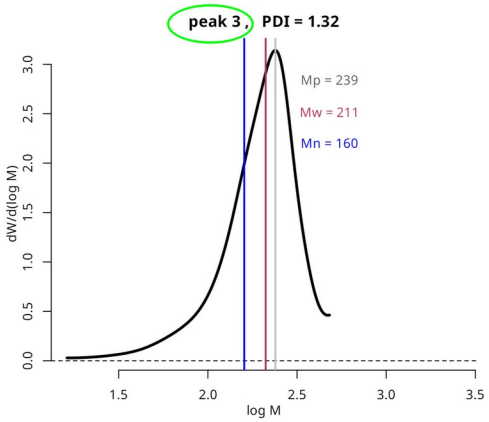






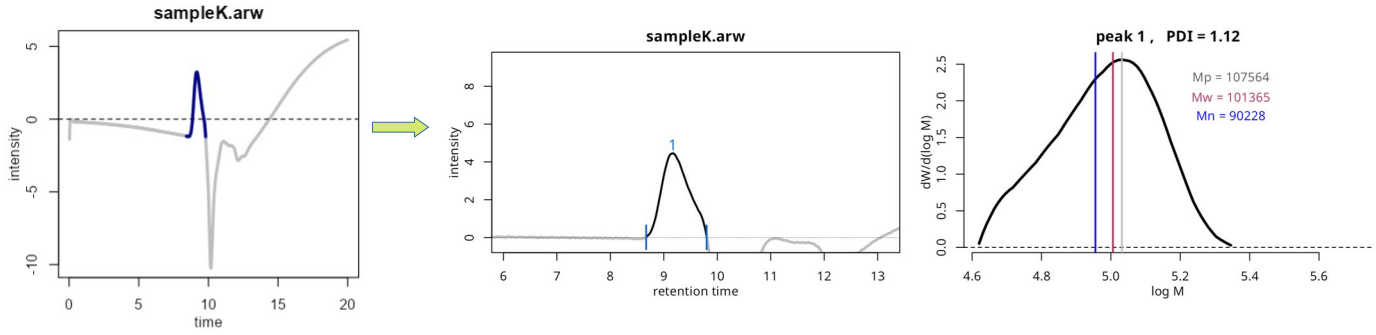






Base line correction

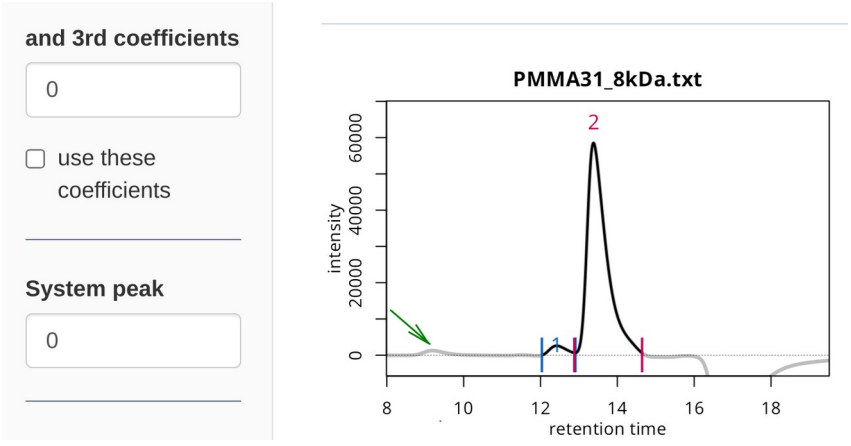
GPC chromatograms may have baseline drift or, shift, other problems [3]. Therefore, the application performs baseline correction so that the peaks start at 0 intensity.



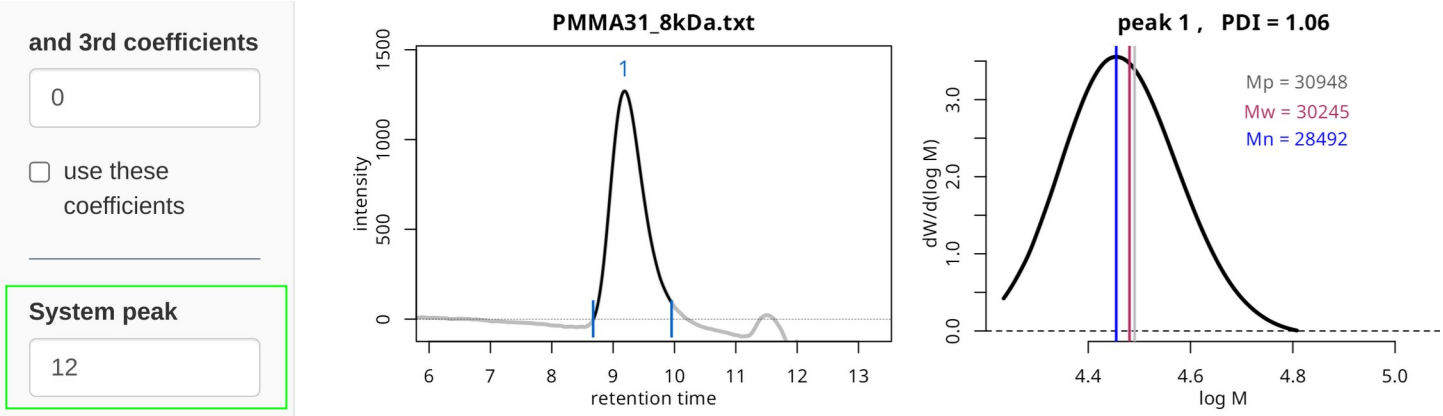
Baseline anomalies are more often present at the beginning of the chromatogram, and the application does not look for peaks in this part of the graph. Therefore, if the analyte peak starts at the beginning, it will not be detected.

System peak

System peaks may be present in GPC chromatograms [4]. If the peak of interest is too small compared to the system peak, the application may not detect it correctly.



Thus if the wrong peak is shown on the plots, you may try to enter the *system peak time*.

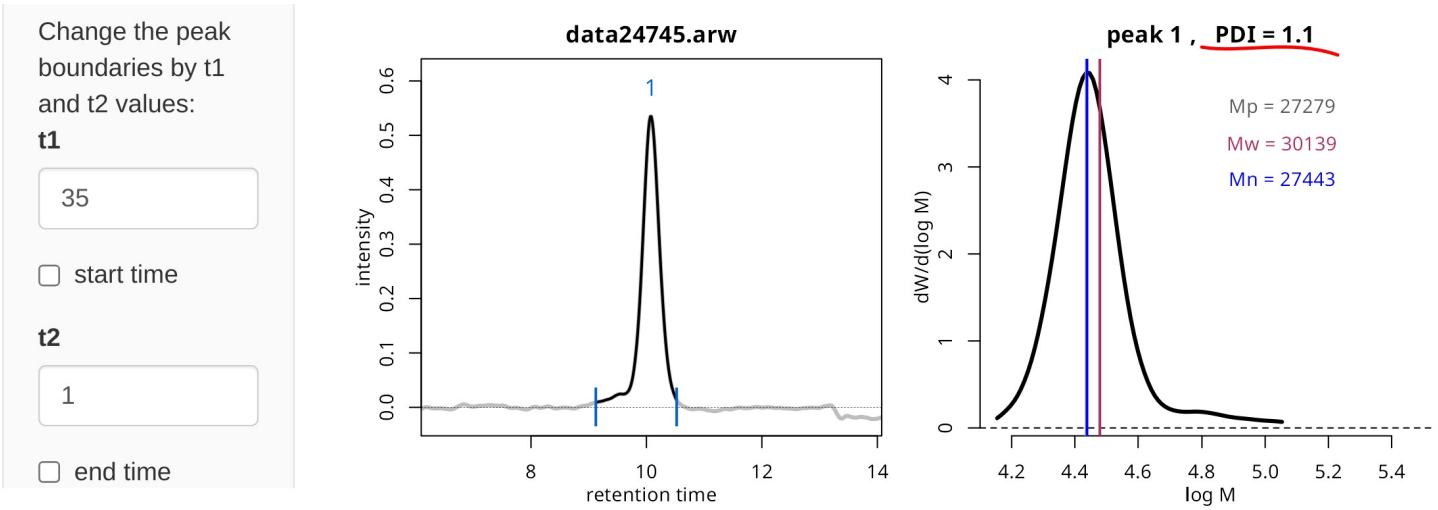


For this sample, the Mp should be 31800 Da and the PDI should be about 1.03, which is close to what was calculated by the app if we cut off the system peak and the peak next to it (possibly caused by some impurities in the sample).

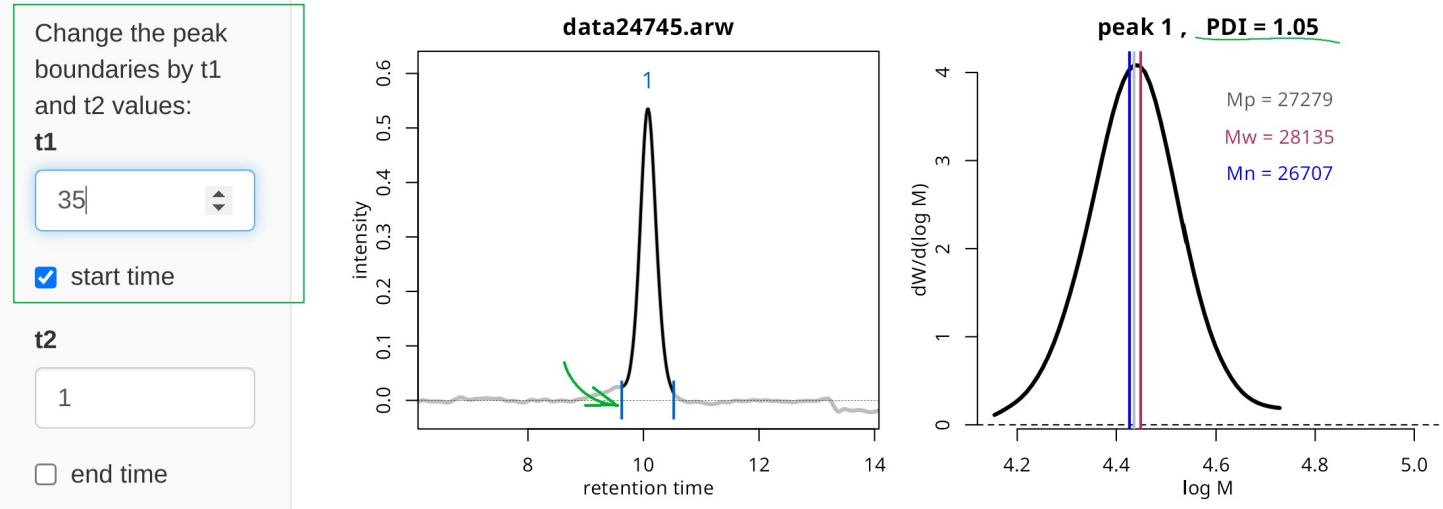
Note that you cannot set the system peak to less than 20% of the maximum retention time.

Peak start and end

The start and the end of the peak (short vertical lines on the chromatograms) can be shifted by *t1* and *t2* values (percentage of the peak width).

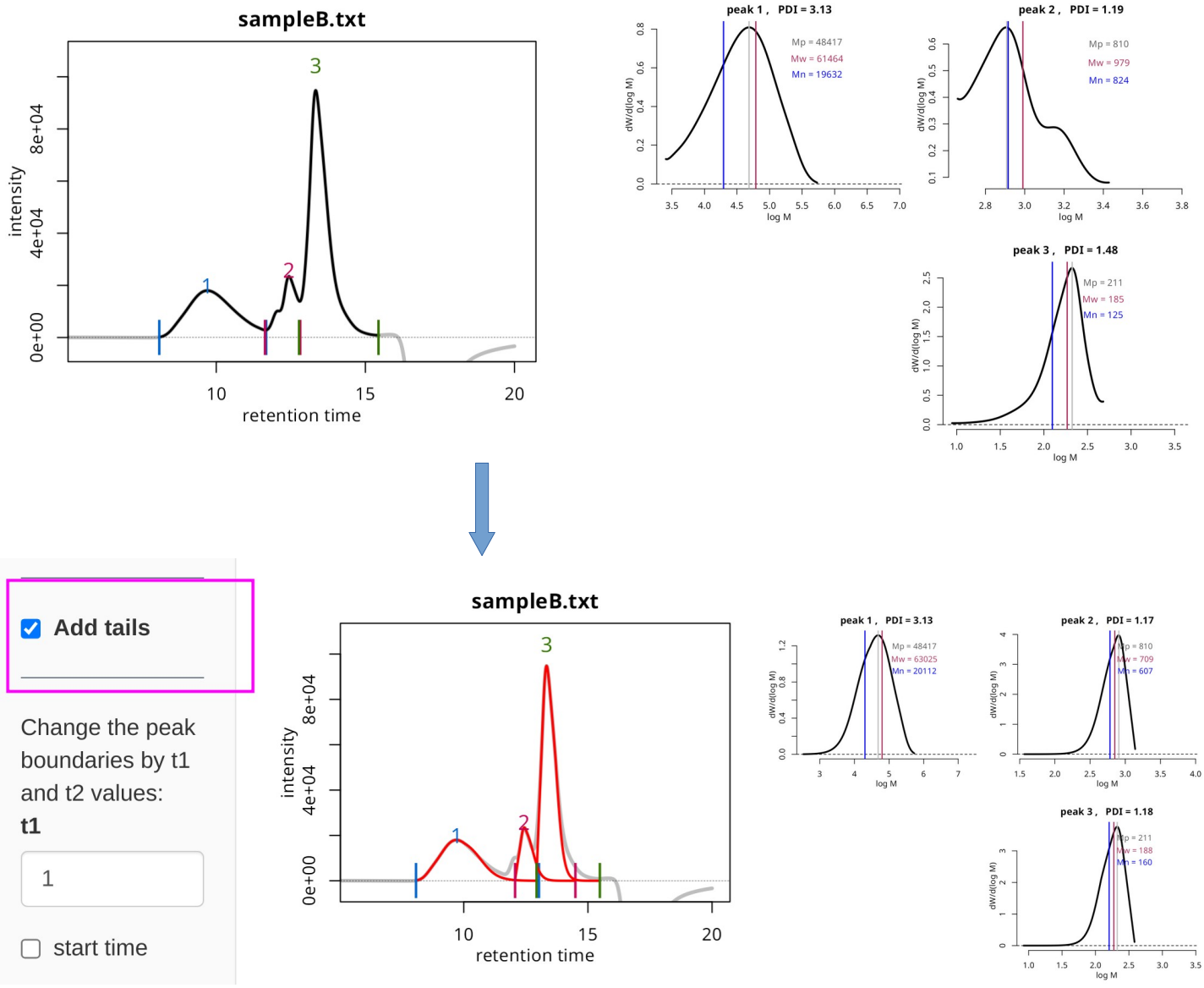


On the sidebar enter the value **t1** (or/and **t2**) and check the boxes below to apply the changes.



Adding tails

If peaks are merged you can add tails: check **Add tails** box on the sidebar.



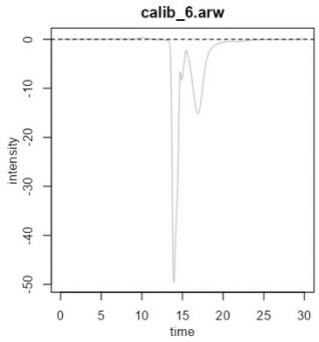
The application will try to fit the sigmoid equation given in [5] and [6] and add the missing part of the chromatogram. If a large portion of the GPC curve is missing, the results may be inaccurate. Note that this app does not make any peak broadening corrections for molecular weight averages and does not add Gaussian curve.

Smoothing

If the chromatogram is noisy, the application will apply smoothing before finding the peaks. However, in some cases the noise may be detected as peaks.

No peaks

If no peaks are detected, only the raw data will be plotted.



This application was created using R version **4.4.2** and R packages:

DT **0.33** purrr **1.0.2** dplyr **1.1.4.9000** stringr **1.5.1** tidyr **1.3.1** shiny **1.9.1**

---

[1] S. Mori and H. G. Barth, *Size Exclusion Chromatography*, 1st ed. Springer Berlin, Heidelberg, 1999.

[2] R. Freitag and L. G. Berruex, “Separation and Purification of Biochemicals,” in *Encyclopedia of Physical Science and Technology*, 3rd ed., Elsevier, 2003, pp. 651–673.

[3] S. Kromidas, “Quantification in LC and GC: a practical guide to good chromatographic data,” 2009.

[4] S. Levin and E. Grushka, “System peaks in liquid chromatography: their origin, formation, and importance,” *Anal. Chem.*, vol. 58, no. 8, pp. 1602–1607, 1986.

[5] K. Heusser, R. Heusser, J. Jordan, V. Urechie, A. Diedrich, and J. Tank, “Baroreflex Curve Fitting Using a WYSIWYG Boltzmann Sigmoidal Equation.,” *Front. Neurosci.*, vol. 15, p. 697582, 2021.

[6] H. Motulsky and A. Christopoulos, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*. in Mathematics. Oxford University Press, 2004.