

## Automatic mass calibration

The automatic mass calibration of unknown spectra is intended to be a method that enables calculation of unit mass resolution sticks and automatically finds peaks from the spectra. The accuracy is not as good as with known calibration substances, but the method is robust and fully automated. The performance of each calibration steps is listed in table 1S. Step 0 is the starting point of the calibration, and the instrumental parameter  $b_0$  has been assumed to be 0, and  $a_0$  is calculated from equation 1S.

$$a_0 = \sqrt{\text{median}(\Delta t^2)} \quad 1S$$

Where  $\Delta t$  is the difference between adjacent peaks in time-of-flight space. This method assumes that the majority of peaks are separated by 1 Th.

Step 1 assumes that each peak is at its integer mass (only masses <300 Th have been used) and that step 0 produced a calibration with errors smaller than 0.5 Th for all masses <300 Th. The mass axis is re-calibrated, this time including the parameter  $b_0$ . Step 2 minimizes the baseline (using masses <800 Th). A positive ambient spectrum with a series of pyridines as test peaks (80.0495, 94.0651, 108.0808, 122.0964, 136.1121, 150.1277, 164.1434, 178.1590) was used for testing. In manual mode the pyridine peaks were selected, and a traditional mass calibration was performed. It can be seen that already Step 1 is sufficient for unit mass resolution sticks (accuracy 666 ppm, that translates to an error in 100 Th of 0.067 Th) and the final step 2 improves the calibration of unknown sample even further. For the peak at 100 Th the final calibration error would be 0.007 Th, that is only 2 times higher than with the manual calibration.

Table 1S. Mass accuracy of each calibration step. For clarity also the error at 100 Th and assumption made at each step are reported.

	Accuracy ppm	Error at 100Th Th	Assumption
Step 0	2600.3	0.260	Majority of peaks 1 Th apart
Step 1	665.9	0.067	Peaks at integer mass (<300 Th)
Step 2	68.1	0.007	Baseline minimization (<800 Th)
Manual	20.2	0.002	Known calibration peaks

The method is sensitive to the initial starting point and it is crucial that the spectrum contains many adjacent peaks that are required in the step 0. If organic compounds have been measured, the mass defect of hydrocarbons can become a problem, the mass defect can be high enough that the peak is at the mid-point between two integer masses (e.g.  $C_{30}H_{60}$  has a mass of 420.4695 Da). This can affect the mass calibration if only part of the spectrum is used, but if the whole spectrum is utilized the lower masses dominate the calibration. However, this is case specific and in our case, where most of the signal is below 500Th, the mass defect of organics is not a problem.

This is demonstrated in figure 1S, where panel A depicts the calibrated spectrum with highlighted baseline (0.4 Th wide area between peaks), in panel B that same spectrum with same mass calibration parameters has been plotted, but now the mass axes has been shifted by mass defect of CH fragment (factor 13/13.0078). In panel A can be seen that the masses above 250 Th starting to have so high mass defect that the defined baseline is covering partly the peaks. However, the panel B demonstrates that the mass calibration has not been effected by mass defect of larger organic masses; in fact the manual mass calibration has the same behavior. It has to be mentioned than the usable range that can be calibrated this way depends on the amount of signal and also the instrument.

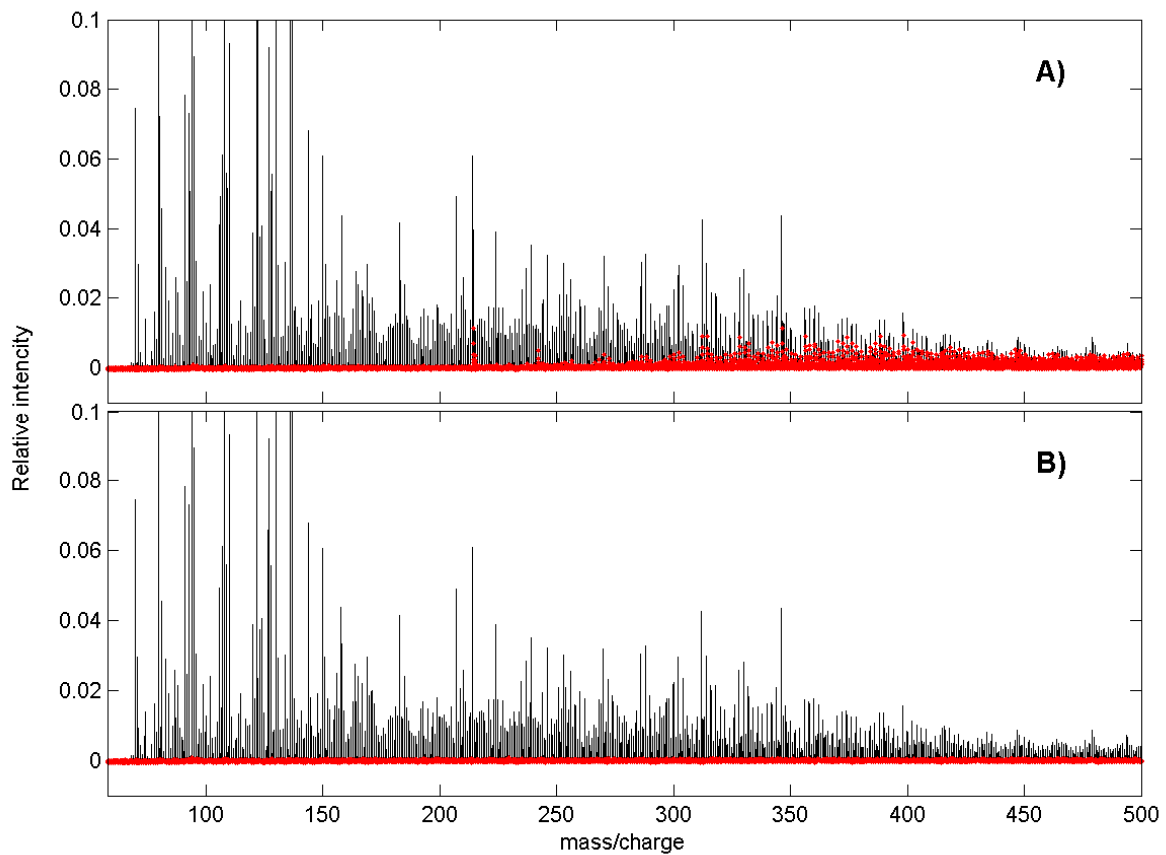


Figure 1S. Result of mass calibration with baseline minimization on positive ambient spectrum (Panel A) and the same mass calibration with shifted mass axes to compensate the mass defect of hydrogen (Panel B). Red dots are 0.4Th wide area between peaks and representing the baseline.