Supporting information for "Asn$_3$, a reliable, robust and universal lock mass for improved accuracy in LC-MS and LC-MS/MS"

An Staes[a,b], Jonathan Vandenbussche[a,b], Hans Demol[a,b], Marc Goethals[a,b], Şule Yilmaz[a,b], Niels Hulstaert[a,b], Sven Degroeve[a,b], Pieter Kelchtermans[a,b], Lennart Martens[a,b] and Kris Gevaert[a,b,*]

[a]Department of Medical Protein Research, VIB, B-9000 Ghent, Belgium

[b]Department of Biochemistry, Ghent University, B-9000 Ghent, Belgium

*Corresponding author: Prof. Dr. Kris Gevaert, A. Baertsoenkaai 3, B-9000 Ghent, Belgium, e-mail: kris.gevaert@vib-ugent.be, fax: +32-92649496

Table of contents:

Figure S-1: Public data derived from the ProteomExchange repository. Bar plots of the percentage of scans in which the lock mass is found (blue) or absent (red) in each LC-MS/MS run. All data shown here were derived from data analyzed by LTQ-Orbitrap instruments. Runs belonging to the same project and hence coming from the same lab are boxed together. For each raw data file, a PCM peak was used as lock mass.
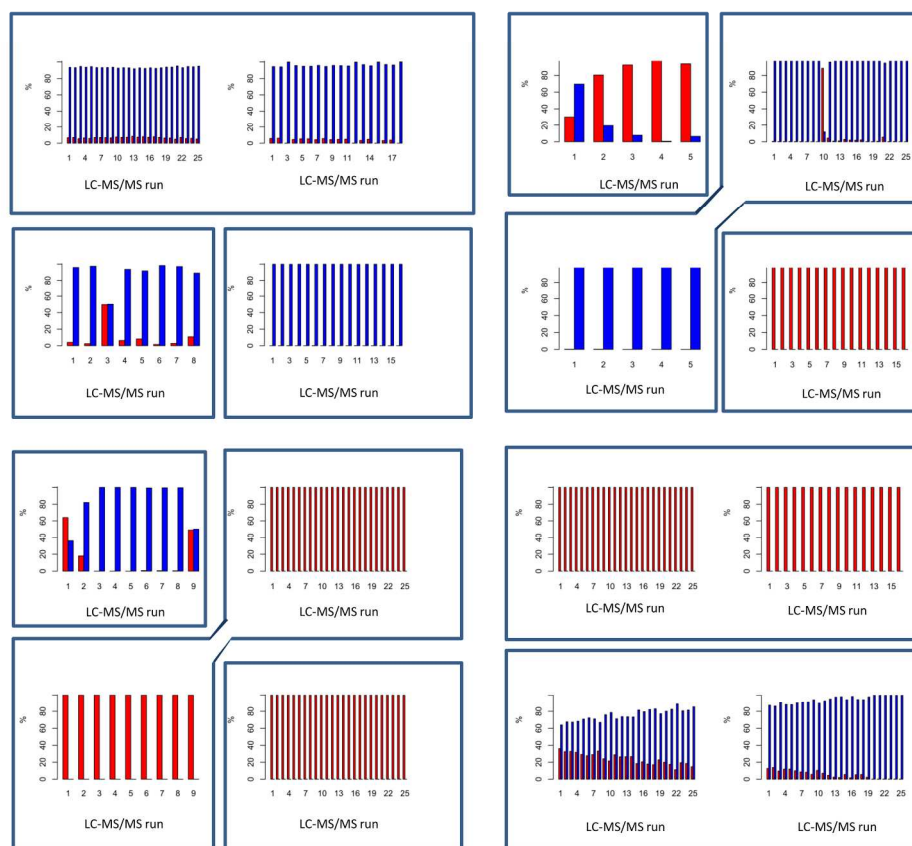
Figure S-2: Extracted ion chromatograms of the lock mass for all LMA0 runs. Panel A shows the overlayed extracted ion chromatograms (XIC) of the lock mass for 20 runs where Asn$_3$ was used as a lock mass (top graph) and for 20 runs were PCM was used as a lock mass (bottom graph). Panel B shows the same data, only zoomed in for the region where peptides actually elute. For PCM, the signal is lost to a high degree in the dense peptide region, as evident from the lower stability of the signal of PCM. For Asn$_3$, and as expected, a higher stability throughout the run is observed.
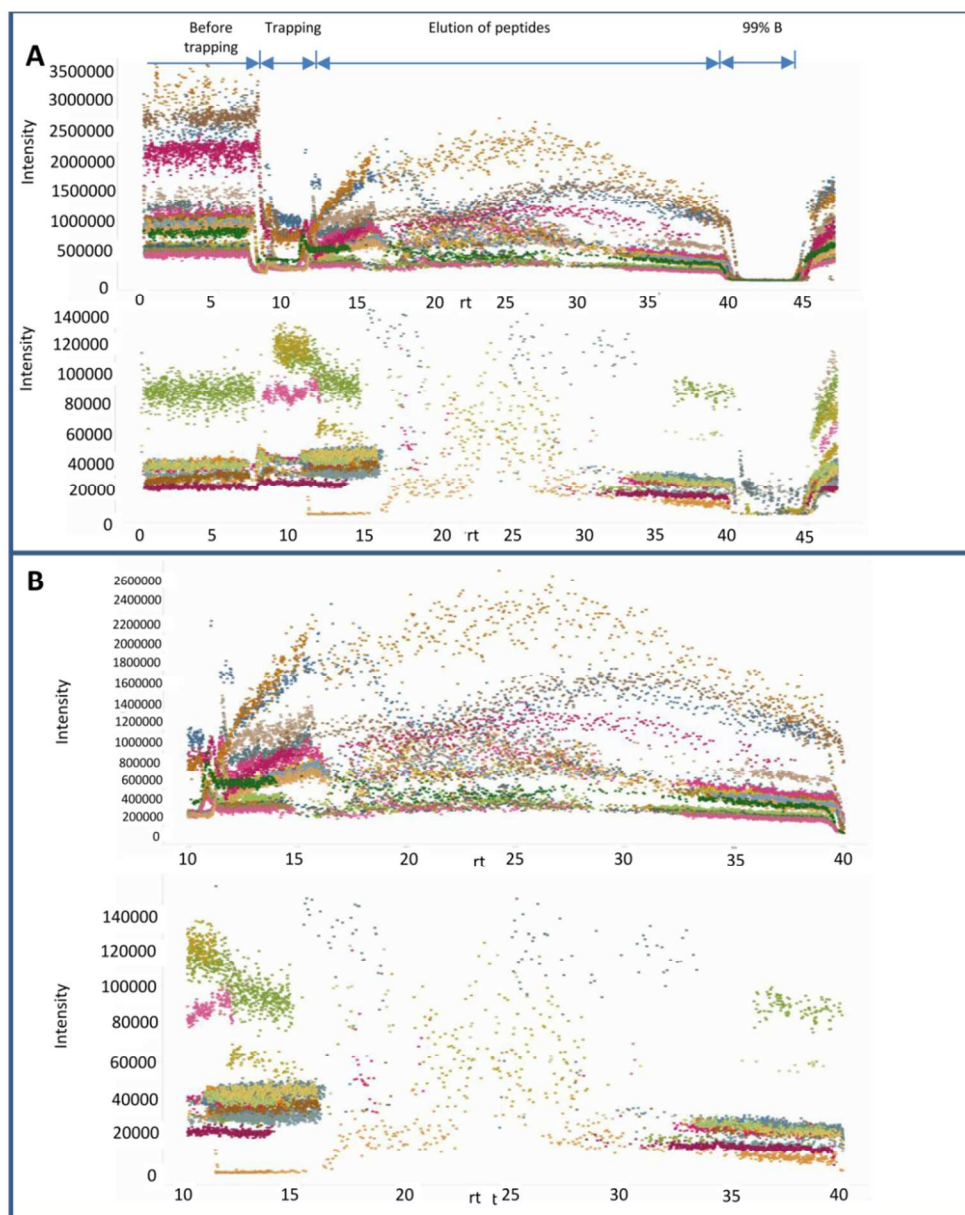
Figure S-3: Asn$_3$ as lock mass has very limited influence on LC peak broadening and on peptide ionization. Panel A displays the extracted ion chromatograms of the same MS peak identified in an analysis with solvents containing 5 nmol Asn$_3$/ml or no Asn$_3$. The retention time of the peak and its full width at half maximum (FWHM) are shown. Panel B shows the number of identified peptides when analyzing 5 µl of the same sample with different amounts of Asn$_3$ dosed in the LC solvents. Panel C shows the number of identified peptides when analyzing 1 µl of the same sample as in panel B, representing a low abundant sample, with different amounts of Asn$_3$ dosed in the LC solvents.
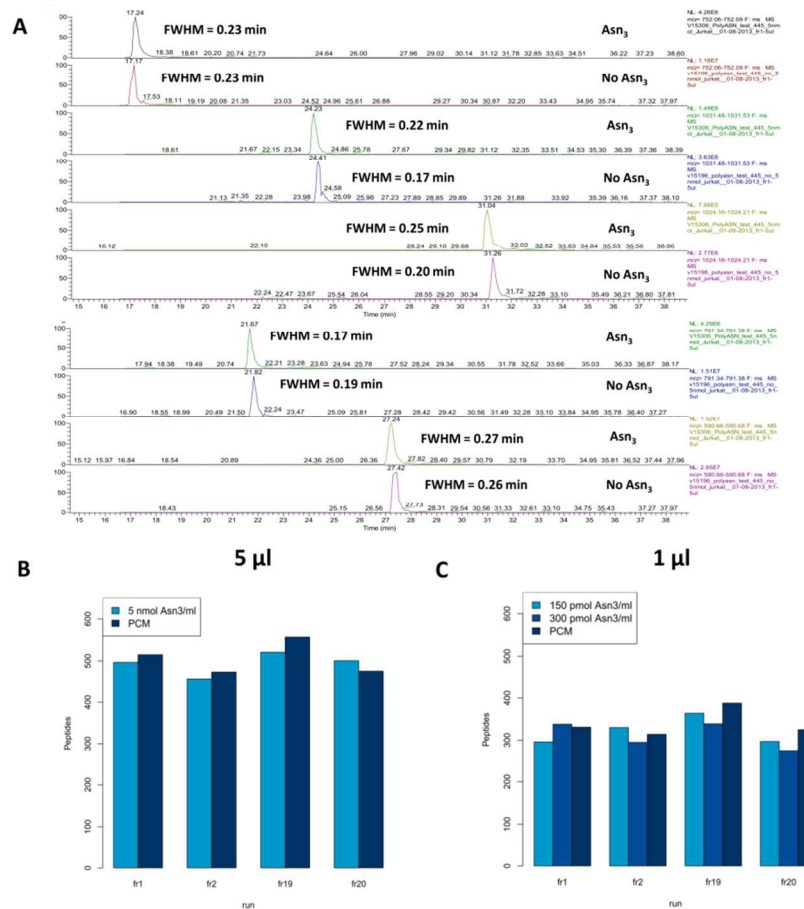
Figure S-4: Recalibrated mass error distributions in MaxQuant. Plots of all 20 samples run with PCM as lock mass or Asn$_3$ as lock mass, in which the difference in Area Under the Curve (AUC) is calculated between quantile 35 and 65 of the Asn$_3$ dataset for each dataset. A positive value means there is a higher density of accurate data in the Asn$_3$ dataset compared to the PCM dataset.
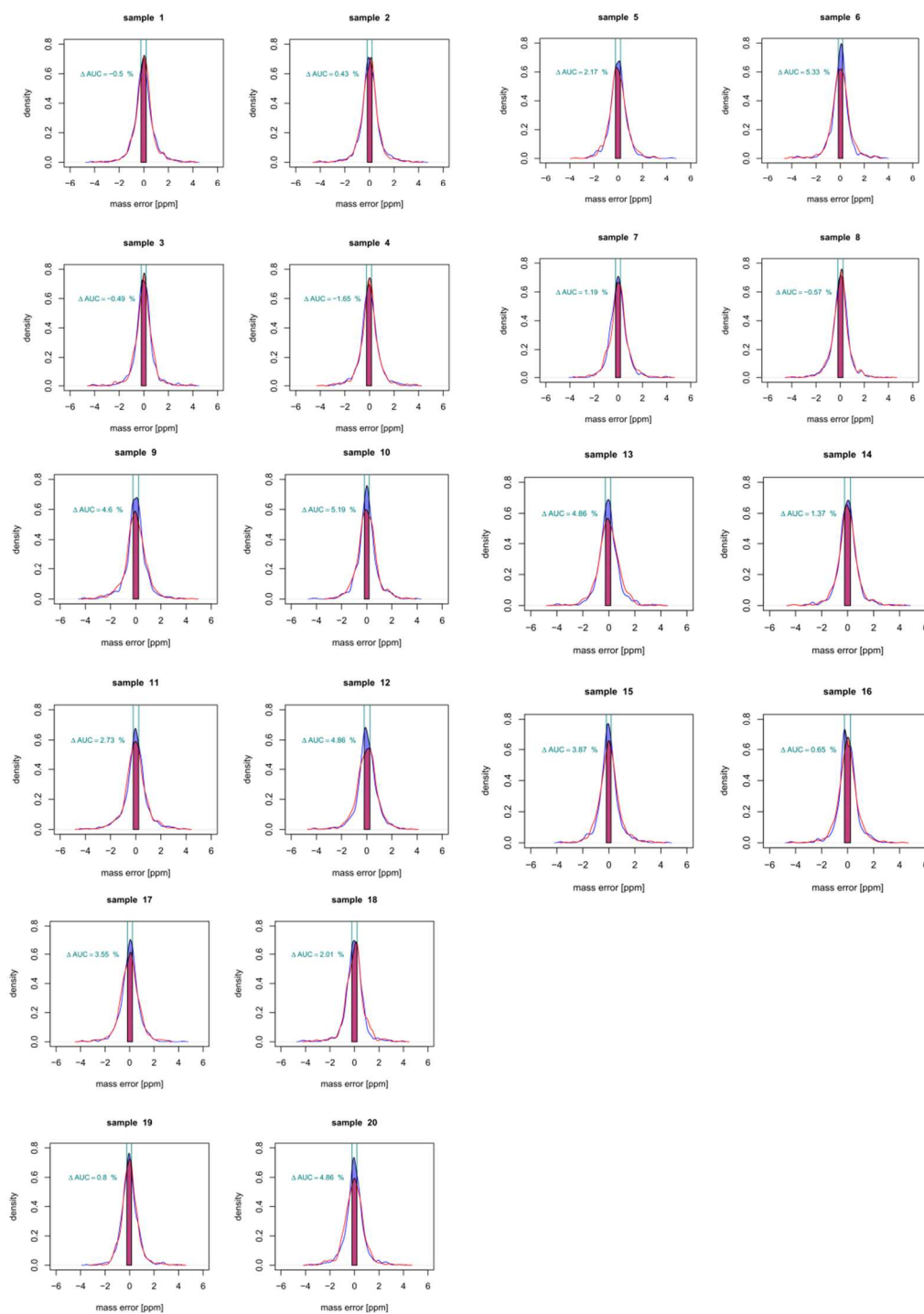
Table S-1: p-values for the two sided t-test and F-test for each sample run with Asn_3 as lock mass and a PCM peak as lock mass. With only a few exceptions (indicated in italics), the p-values are below 0.05 for the F-test, indicating that the variances of the mass measurement errors are significantly different between the different lock mass compounds. The median mass measurement errors however, are typically not significantly different as shown by the t-test results.

Sample	p-value t-test	p-value F-test
fr1	*0.267*	*0.935*
fr2	0.018	0.007
fr3	0.020	*0.247*
fr4	0.000	0.001
fr5	0.023	0.000
fr6	0.000	0.000
fr7	0.023	0.000
fr8	0.000	0.003
fr9	*0.907*	0.000
fr10	*0.288*	0.000
fr11	0.000	0.000
fr12	0.000	0.000
fr13	*0.176*	0.000
fr14	0.001	0.000
fr15	*0.721*	0.000
fr16	0.000	0.000
fr17	*0.513*	0.001
fr18	*0.719*	0.032
fr19	*0.055*	*0.383*
fr20	*0.836*	*0.104*