**Automated evaluation of LC-HRMS full scan data for the rapid characterization of surface water samples**

Martin Krauss1, Christine Hug1, Robert Bloch1, Tobias schulze1, Werner Brack1

1 Department Effect-Directed Analysis, Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, 04318 Leipzig, Germany, martin.krauss@ufz.de

With the increased availability of high resolution–high accuracy mass spectrometers such as QToF and Orbitrap instruments, a nontargeted analysis of environmental samples by LC-HRMS has gained popularity. Still, the data evaluation is a major bottleneck in nontarget screening, as typically several thousands of peaks are obtained from peak picking, for which manual processing and evaluation is tedious.

We present an approach which allows for a rapid characterisation of surface water samples based on “peak attributes”, which can be detected by automatic software routines with a good reliability. The individual processing steps include (i) an automated peak picking using the MZmine 2 software (Pluskal et al., 2010), (ii) blank removal using an R script, and (iii) detection of isotopes, adducts, and homologue series using the R package “nontarget” (Loos, 2012). This workflow allows obtaining the total number of peaks, number of peaks with negative mass defects, 13C, 34S, 15N, 37Cl, and 81Br isotope peaks, and number of homologue series within a processing time of < 3 hours for a set of 30 samples. The approach was applied to a set of 31 samples from rivers and streams of different size and a different portion of wastewater from the Saale and Mulde catchments. These were solid-phase extracted using a cartridge containing four different sorbents and analysed using LC-HRMS using an LTQ Orbitrap XL instrument (Thermo) in ESI+ and ESI- mode.

The total number of peaks was positively correlated with the fraction of wastewater (Spearman rank correlation R = 0.69 for ESI+ and 0.76 for ESI-, respectively) and the number of peaks showing a 13C, 34S, 15N, 37Cl, and 81Br isotope pattern, which shows that the LC-HRMS amendable compound inventory is in general dominated by compounds derived from wastewater. Surfactants are an important contributor to this inventory, particularly for sites impacted by untreated wastewater, as indicated by a large number of homologue series. Beyond this general trend, the evaluation revealed several sites with an outstandingly large number of high-intensity peaks or a large number of peaks showing 34S or 37Cl isotopes. It turned out that at these sites specific inputs from industrial wastewater or former superfund sites occur, resulting in specific compounds groups not typically analysed within monitoring programs. Finally, some of these compounds could be identified by manual evaluation, which benefitted from the readily available information on isotope and adduct peaks and homologue series. Thus, the proposed approach is suitable to rapidly characterize surface water samples and allows for a prioritization of sites or compound groups for further in-depth studies.

Loos, M. (2012): Nontarget: Detecting, combining and filtering isotope, adduct and homologue series relations in high-resolution mass spectrometry (HRMS) data. http://www.eawag.ch/forschung/uchem/software/R\_package\_start

Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešic, M. (2010): MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics 11*, 395.