Student Research – SR 2024.004

Semi-automated process to build a suspect list with water contaminants and their transformation products

Peli Angouraki – Chemistry for Environment & Health, Amsterdam Institute for Life and Environment (A-LIFE), Vrije Universiteit Amsterdam

Supervised by
Nienke Meekel – Chemical Water Quality and Health, KWR Water Research Institute
Frederic Béen – Chemical Water Quality and Health, KWR Water Research Institute

Chemical products are widely used in modern society. Examples are pharmaceuticals in healthcare and pesticides in agriculture. These chemicals might enter the environment via run-off or the wastewater treatment plant. Municipal wastewater is collected mainly from residential areas but involves industrial wastewater as well, containing a plethora of chemical compounds. The treatment of municipal wastewater is focused on the removal of organic material and microorganisms, while chemical substances are not fully eliminated. On top of that, the initial substances might undergo partial degradation to form so-called transformation products, about which very little is known.
To obtain a better understanding of the underlying hazards of these chemicals and their transformation products that are released into the environment, it is important to develop methods to detect and identify them. Liquid chromatography (LC) coupled with high-resolution mass spectrometry (HRMS) is an analytical technique that is often used for the detection of these chemicals and their transformation products. The resulting measurement data can be screened using a list of specific substances that are expected to be present in the sample, so-called suspect screening. Suspect screening can be applied to search for transformation products of a set of chemical substances, with the use of prediction models that predict their chemical structure based on possible transformation reactions.
During this internship project (May 2023 – March 2024), a (semi-)automated workflow was developed for the prediction of transformation products. The input is composed of a given list of chemical substances that are expected to enter the wastewater treatment plant (e.g., emission registry). The output consists of a greatly extended suspect list compiling the known chemicals, and their reported and predicted transformation products, that are expected to be found in wastewater effluent. These compounds are then prioritized based on their predicted affinity with water, and their amenability with LC-HRMS, to flag chemicals that are unlikely to be present in water and/or cannot be detected using LC-HRMS. The workflow was built and tested with an initial dataset containing > 5,000 chemical entities, of which 509 relevant and unique organic chemical structures remained after filtering. At the end of the process, two suspect lists were created, one for each mass spectrometry ionization mode, with ±1,400 unique chemical structures (parents and their transformation products) expected to be detectable with reversed-phase chromatography coupled to HRMS. Extra care was put in the validation of the chemical substances after each step of the workflow, to avoid the presence of irrelevant or inconsistent chemical structures in the final suspect lists.
This work can be the base for the creation of other extensive lists of chemicals for suspect screening workflows, with access to different transformation product prediction models, and using other prioritization criteria, such as biodegradability and toxicity.

Keywords: transformation products, wastewater effluent, suspect screening, *in silico* tools