FP7 SOLUTIONS project - Fact Sheet 003

Name SOLUTIONS Tool or Service

Non-target screening and structure elucidation workflow

Description

1. Objective

You have reached this Fact Sheet because you are interested in identifying unknown substances with non-target screening techniques.

Mass spectrometry (MS) is often used to analyse and determine the inventory of substances from rivers and other sampling sites. While target screening, the traditional approach, requires *a priori* knowledge of the chemicals to be detected, suspect and non-target screening cover more substances and are becoming increasingly popular. In suspect screening, chemicals such as all REACH chemicals are suspected to be present in an environmental sample and can be searched for explicitly. In non-target screening, no prior information is available.

2. Methodology

The non-target screening and structure elucidation workflow developed in SOLUTIONS enables a comprehensive chemical assessment of monitoring samples, especially water samples. Specifically, this workflow uses acquired chemical data (liquid chromatography-mass spectrometry data) as input and performs the data pre-processing and mass spectral interpretation steps involved in target, suspect and non-target screening. This workflow is designed to gather all information available from the analysed sample and combine this with the structure elucidation efforts to identify relevant emerging contaminants. The structure elucidation software includes meta-information specific for the environmental context, such as the number of references, patents, data sources, as well as compound filters and user-defined scores for candidate ranking including presence in suspect lists and information about the elemental composition or substructure occurrence. These user-defined scoring terms also allow the inclusion of toxicity information for candidate structures. Experimental information such as retention time and isotopic labelling can also be included. Together, these methods were used to identify non-target chemicals within the case studies on the rivers Rhine and Danube, and are applicable to many applications, e.g. screening for RBSP, surveillance monitoring, effect directed analysis.

The non-target screening workflow is shown in Figure 1. It was compiled to be as a flexible a workflow as possible, such that each user can tailor the workflow to their own needs. Thus, the focus was on a flexible combination of several ‘blocks‘ of a workflow that can be pieced together as needed, while each block can also be used alone and also accept alternative inputs. The focus has been on developing an open source workflow that uses external resources that are also open source or openly accessible wherever possible. The workflow is available in the programming language R and also comes with two graphical user interfaces. Additional functions in the workflow allow for the interpretation and plotting of the results, to assist in the selection of non-target substances for the structure elucidation efforts using the second part of the workflow, MetFrag.

The structure elucidation approach is shown in Figure 2. The MetFrag workflow starts by retrieving candidate structures, which are then fragmented using a bond dissociation approach. These fragments are then compared with the measured MS/MS spectrum to determine which candidate best matches the measured data. Additional criteria are also considered, including the number of references, data sources and/or patents for a substance (a high number of literature references or the listing in many patents may imply the substance is of high use and thus more likely to be found in the environment). The use of retention time information is also possible, as well as element and substructure selection where available, as well as ways of ‘flagging’ structures potentially of interest. Furthermore, ‘suspect screening’ approaches were implemented to find

candidates of interest in different lists of chemicals. User-defined scores were added to allow the inclusion of toxicity and other information. The information from hydrogen-deuterium exchange experiments was incorporated with three additional scores.

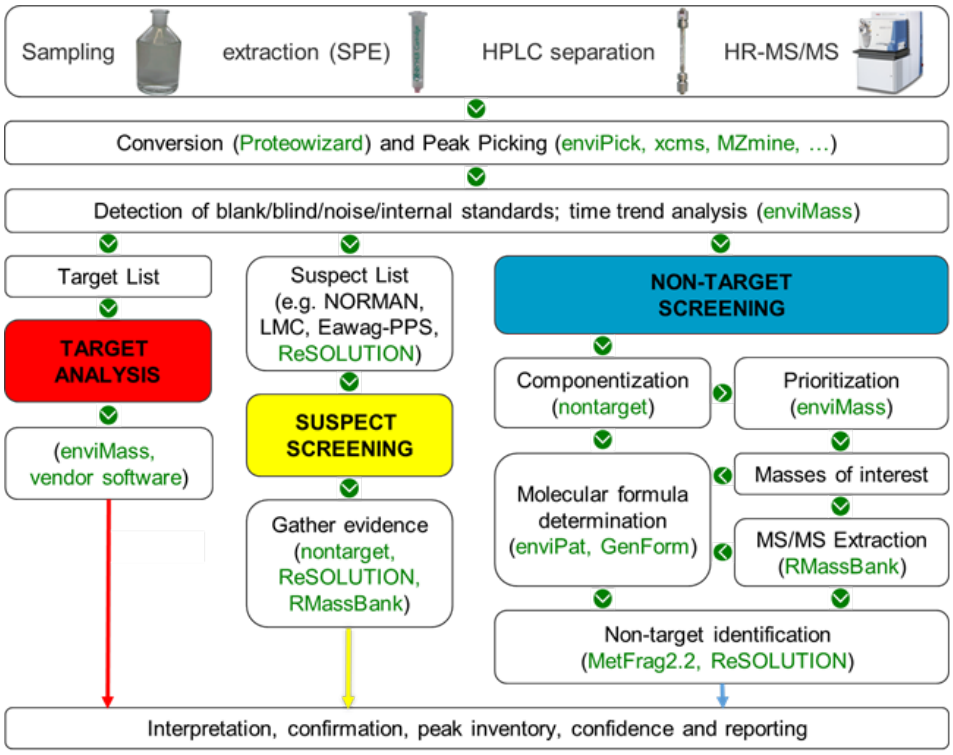

Figure 1. The non-target screening workflow, showing the interconnection between the various parts. The workflow parts are listed for every step (websites given below).

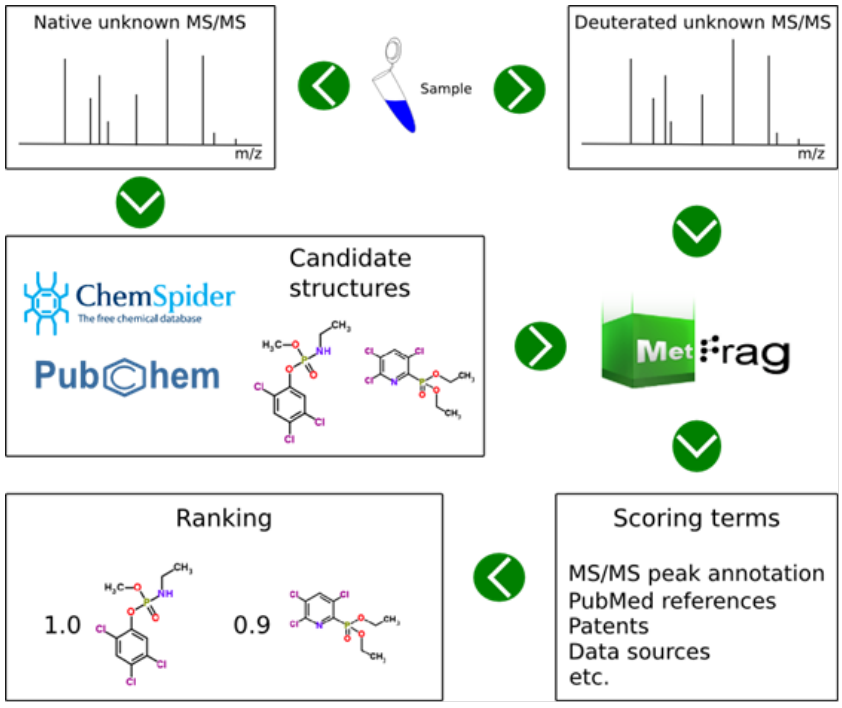

Figure 2. Schematic of the structure elucidation approach used in MetFrag2.2.

3. Application

The changes to MetFrag were evaluated on a dataset of 473 merged high resolution tandem mass spectra (HR-MS/MS). Using HR-MS/MS information only, MetFrag2.2 had 30 Top 1 ranks, while including reference and retention information improved this to 420 and 336 Top 1 ranks using the

two largest compound databases as a source of candidate molecules. The optimal parameters and weights were verified using three additional datasets (see Ruttkies *et al.* 2016 [Ref01]). MetFrag2.2 is available at http://c-ruttkies.github.io/MetFrag/.

The combination of the workflow and structure elucidation parts yields a flexible, high-performing workflow for non-target screening in environmental samples ([Ref02], [Ref03]) that has been applied in the SOLUTIONS case studies FS042, FS040. In many cases, additional information is available from the experimental context to add to small molecule identification, which is especially useful where the mass spectrum alone is not sufficient for candidate selection from a large number of candidates. The combination of the non-target screening workflow with the new features in MetFrag2.2 greatly improves the results from comprehensive environmental analysis. This method helps determining contaminants of emerging concern (and thus of interest) in environmental samples in a complementary matter to FS001 as part of FS051.

References

1. Ruttkies, C. E.L. Schymanski, S. Wolf, J. Hollender and S. Neumann, 2016. MetFrag Relaunched: Incorporating strategies beyond in silico fragmentation, J. Cheminform. 8:3; https://doi.org/10.1186/s13321-016-0115-9
2. Schymanski, E.L., H.P. Singer, Ph. Longrée, M. Loos, M. Ruff, M.A. Stravs, C. Ripollés Vidal and J. Hollender, 2014. Strategies to characterize polar organic contamination in wastewater: Exploring the capability of high resolution mass spectrometry. Environ. Sci. Technol. 48(3): 1811-1819; https://doi.org/10.1021/es4044374
3. Schymanski, E.L., H.P. Singer, J. Slobodnik, I.M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, N.S. Thomaidis, A. Bletsou, C. Zwiener, M. Ibáñez, T. Portolés, R. de Boer, M.J. Reid, M. Onghena, U. Kunkel, W. Schulz, A. Guillon, N. Noyon, G. Leroy, Ph. Bados, S. Bogialli, D. Stipaničev, P. Rostkowski and J. Hollender, (2015. Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Anal. Bioanal. Chem. 407: 6237-6255; https://doi.org/10.1007/s00216-015-8681-7

Websites:

enviMass https://github.com/blosloos/enviMass [30/09/2015]
enviPat http://www.envipat.eawag.ch/index.php [30/01/2014]
enviPick http://cran.r-project.org/web/packages/enviPick/index.html [30/09/2015]
GenForm http://sourceforge.net/projects/genform/ [25/01/2016]
MetFrag http://c-ruttkies.github.io/MetFrag/ [25/01/2016]
MZmine2 http://mzmine.github.io/ [25/01/2016]
Nontarget http://cran.r-project.org/web/packages/nontarget/index.html [30/09/2015]
NORMAN Suspect List Exchange http://www.norman-network.com/?q=node/236 [30/09/2015]
Proteowizard http://proteowizard.sourceforge.net/ [25/01/2016]
RMassBank http://bioconductor.org/packages/release/bioc/html/RMassBank.html [25/01/2016]
Xcms https://bioconductor.org/packages/release/bioc/html/xcms.html [25/01/2016]

Keywords

Non-target screening, mass spectrometry, structure elucidation, chemical assessment, identification

Related topics

Analytical strategies FS051
Protocols for target analysis of emerging contaminants in water, sediment and biota FS001
Standard operational procedures (SOPs) for individual organic compounds FS078

Standard operational procedures (SOPs) for organic compounds classes FS079

Syntheses of reference standards for SOLUTIONS FS005

Suspect screening FS052

Contact information

E. Schymanski (emma.schymanski@uni.lu), J. Hollender (juliane.hollender@eawag.ch)
Eawag, Dübendorf, Switzerland

C. Ruttkies (cruttkie@ipb-halle.de), S. Neumann (Steffen.Neumann@ipb-halle.de)
Leibniz Institute of Plant Biochemistry (IPB), Halle, Germany