Screening emerging chemicals in human matrices to support biomonitoring and environmental health studies: methods, challenges and promises

Pourchet M[1], Cariou R[1], Antignac JP[1], Le Bizec B[1]

[1]LABERCA, Oniris, INRA, Université Bretagne-Loire, F-44307, Nantes, France, laberca@oniris-nantes.fr

Introduction

Emerging chemicals clearly appears as a major current concern for the scientific community, societal actors, and public authorities. However, this issue is covering different aspects from both conceptual and methodological points of view. Basically, "emerging chemicals" may be considered as compounds recently appeared in the environment, for instance newly developed substitutes of banned or currently under regulation substances. They can also be considered as "chemicals of emerging concern", *i.e.* compounds present for a while in the environment-food-human continuum but for which the concern has increased only recently[1]. Such new concerns can arise due to analytical methods sensitivity improvements, allowing the detection at trace/ultra-trace level of formerly not detected chemicals. In addition, new application fields developed by the chemical industry for a known chemical can open up a new route of exposure. Alongside, recent toxicological facts can also change the perspective for risk assessment on a given chemical.

A (probably wide) number of these emerging substances present in the current Human chemical exposome are not yet included in existing Human biomonitoring programs, mainly due to the absence of analytical methods available to determine the considered chemicals or their metabolites in human specimens. For this reason, efficient analytical techniques able to analyse these compounds in the broadest way are mandatory. Moreover, these techniques have to be flexible in order to characterise the evolution of the human exposome over time. Last generation of chromatography coupled to mass spectrometry instrumentation today appears as the gold standard for that purpose, due to high performances it offers both in term of selectivity and sensitivity. On the basis of such instrumentations, three different screening approaches have been developed, namely "targeted", "suspect" (or "semi-targeted") or "untargeted".

Targeted approach usually refers to quantitative determination of relatively limited number of known compounds without any doubt on their identification and using a specific signal acquisition mode. Conversely, suspect screening approach commonly refers to the analysis of a wider range of "known unknowns"[2], using, at least in a first stage, full scan signal acquisition mode. In that case appropriate matching databases are used to annotate, and ideally unambiguously identify, as many markers as possible by comparing experimental data with reference MS data, previously stored for a set of *a priori* known compounds. The last possible screening mode is the untargeted one, where "unknown unknowns" are expected to be detected also by using full scan signal acquisition mode coupled to complementary structural elucidation work.

In the present work, we have investigated the capabilities of Liquid Chromatography coupled to High Resolution Mass Spectrometry (LC-HRMS) to develop suspect and untargeted screening approaches aiming to characterise human matrices and potentially find new exposure markers related to emerging chemicals.

Materials and methods

Chromatography

LC is generally more frequently used than Gas phase Chromatography (GC) for suspect and/or untargeted screening, since it allows a broader coverage of chemical diversity without dedicated sample treatment. GC is also usually introducing a derivatisation step to facilitate the detection but that may complicate the identification of unknown

compounds in the case of untargeted screening. However, GC based instrumentation basically represents a relevant and complementary tool to address non-polar and volatile substances for which the sensitivity is often lower with LC based approaches. Conventional reversed phase (RP) systems (e.g. C_{18} stationary phases) are still the most commonly employed in LC. Hydrophilic Interaction Liquid Chromatography (HILIC) represents a complementary alternative of interest to separate hydrophilic to highly hydrophilic compounds. The nature of the mobile phase as well as the elution gradient parameters are additional factors influencing the chromatographic separation. In contrast to targeted approaches focusing on a limited number of known compounds for which the elution system may be precisely optimised, suspect and untargeted screening approaches are usually based on less selective and more generic chromatographic conditions to cover the most extended range of possible markers. In the present work, an UltiMate 3000 UHPLC pumping system coupled to an Orbitrap Q-Exactive mass spectrometer fitted with a heated ESI source (HESI, Thermo Fisher Scientific, San José, CA, USA) was used. Instrument control and data processing were carried out by Chromelon Xpress and Xcalibur softwares (Thermo Fisher Scientific). The chromatographic separation was achieved using RP chromatography on a Hypersil Gold C_{18} column (100x2.1 mm, 1.9 µm particles, ThermoFisher Scientic) kept at 45 °C. Mobile phase consisted of 10 mM ammonium acetate in water (A) and acetonitrile (B). The gradient began with (A/B) 80:20 (v/v) for 2 min, then ramped linearly to 20:80 (v/v) over 5.5 min, a second linear ramp to 0:100 over 6.5 min, this ratio is maintained for 6 min and return to 80:20 (v/v) over 1 min. Flow rate and sample injection volume were set at 0.4 mL/min and 5 µL respectively.

Mass spectrometry

Following the chromatographic separation step, a selection of appropriate ionisation and signal acquisition modes has to be operated. For LC-HRMS based screening, Electrospray Ionisation (ESI) in both positive and negative modes is most commonly used due to its high versatility and common availability. On GC-MS based systems, electron ionisation (EI) is by far the most commonly employed ionisation mode, and the richest MS databases available contain EI generated mass spectra. As far as the signal acquisition is concerned, suspect and untargeted screening require a non-selective full scan acquisition mode in the scope of detecting a wider set of compounds without pre-selection. High resolution mass analysers are then preferred to obtain the required sensitivity and selectivity[3]. The most widely used mass filters in that context are Time-of-Flight (ToF or Q-ToF) and Orbitrap devices.

In the present study, both positive and negative ionisation mode were applied with HESI parameters as follow: sheath gas flow, 50 arbitrary units (AU); auxiliary gas flow, 5 AU; capillary temperature, 350 °C; source heater temperature, 350°C; spray voltage, 3 kV and 2.5 kV in positive and negative mode respectively; S-lens radio frequency, 50 AU. HRMS data were acquired in full scan mode over the m/z range 100-1 000 and 1 000-2 000 at a resolving power of 140,000 full width half maximum (FWHM) at m/z 200. The automatic gain control (AGC Target) was set at high dynamic range (5×10^5) and the maximum injection time was set to 500 ms.

Data processing and compound identification

Suspect screening data processing basically uses a reference library to compare and to match each detected signal in the analysed samples to a list of already known and well characterised compounds. Even if some external and web accessible resources may be used for that purpose, this reference library is usually in-house elaborated and research project dependent. Finally, special attention has to be paid to the confidence level associated to this identification process, the maximal level being reached when experimental data are successfully confronted to data obtained from a pure reference standard. This annotation and identification issue still remains the main bottleneck in this field. For untargeted screening, data processing is even more challenging and requires particular bioinformatics and/or statistical tools and skills together with high-level expertise in mass spectrometry (Fig. 1). When chemicals of

interest are *a priori* unknown, several approaches may be used to extract relevant information from the whole generated dataset. A first approach may consist in comparing the untargeted chemical profiles obtained for two sub-population groups, for instance known – or presumed – to display low and high exposure levels respectively. Associated to multivariate statistics (e.g. PCA, PLS…) this comparative data processing may allow evidencing possible markers related to this higher exposure. A second approach for untargeted screening includes generating and comparing global chemical profiles obtained from an identical (or similar) sub-population group characterised at two different time periods, *i.e.* two samples collected at a recent *versus* older time point. The application of appropriate multivariate statistics may then allow highlighting trends from these data that may correspond to possible markers of exposure related to emerging chemicals. However, these two approaches are facing the limitations of confounding factors that may contribute to the discrimination between the two considered sample batches apart from the expected difference of exposure. A third approach consists to exploit a particular chemical signature associated to a given substance group as a filter to extract them from the whole dataset. One relevant illustrating example is a focus on halogenated substances, which display significant mass defects. Thanks to the chlorine and bromine atoms characteristics, isotopic patterns can be easily detected in untargeted spectrometric data[4]. However, the development and implementation of dedicated bioinformatics tools is necessary to achieve this data processing that is not manually feasible. In our research, this task is achieved with the help of HaloSeeker v1.0, a user-friendly software application specifically developed for screening halogenated chemicals from untargeted high resolution mass spectrometry data.

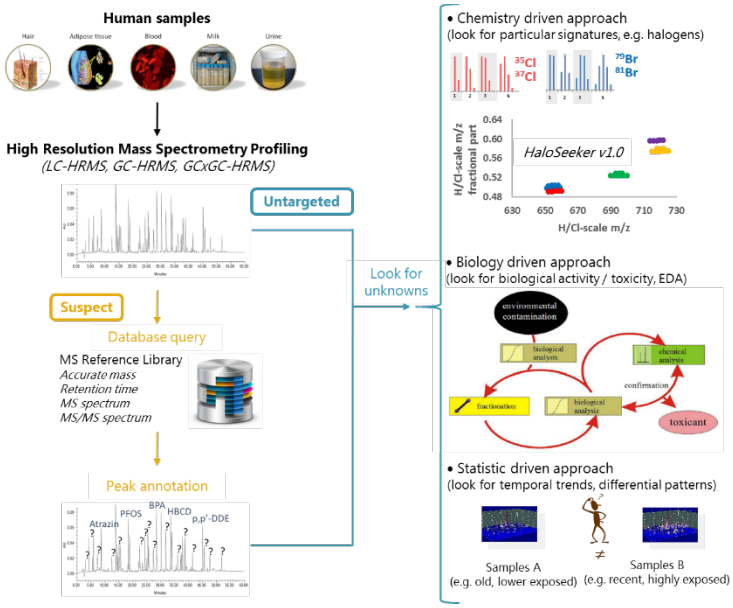

Figure 1: Bioinformatics and statistical tools for suspect and untargeted screening data processing.

Results and discussion

In the present work we aim to investigate the capabilities of LC-HRMS based suspect and untargeted screening from various human matrices including storage/bioaccumulation (adipose tissue, meconium), circulating (blood) or excretion (urine, milk) compartments. All these supporting human samples used for method development and first exploratory application were originated from various research projects conducted within our laboratory at national scale between 2006 and 2016. One critical issue directly influencing expected/obtained results from these approaches is the procedure applied for sample preparation. It should be basically as limited as possible, in order to prevent as much loss of information as possible. However, the injection of crude biological matrices without any pre-treatment does not appear conceivable at this time since both chromatographic and mass spectrometric systems cannot accommodate such complexity. For adipose tissue, both sulphuric acid digestion and Gel Permeation Chromatography (GPC) have been tested and compared, showing a complementarity of both strategies. In particular,

GPC appears more adapted for compounds that are degraded in acidic conditions (e.g. some organochlorine pesticides such as β-HCH, aldrin, heptachlor, some novel halogenated flame retardants…). For other matrices, liquid/liquid partitioning between aqueous (H_2O/ACN) and organic (Hexane) phases followed by supplementary solid phase extraction (C_{18} reverse and SiOH normal phases) was first applied, then compared to a less selective strategy inspired from a typical Bligh and Dyer partitioning. The influence of initial sample amount (*i.e.* µL/mg *versus* mL/g range) in term of resulting useful information/noise ratio was also evaluated, demonstrating globally a better suitability of the latter. Moreover, ion suppression induced by the applied sample preparation strategy (matrix effect) was evaluated thanks to a direct introduction of standard compounds in parallel with the injection of biological sample extracts through the LC system[5]. Regarding the data analysis process, a first step consisted in characterising recoveries of known reference substances added in each analysed sample, namely hexabromocyclododecane (HBCD) and tetrabromobisphenol A (TBBPA) as models for two classes of emerging compounds (brominated flame retardants and halogenated phenols respectively), to evaluate the applied sample preparation efficiency and relevance. In a second step, some particular signals corresponding to known chemicals were manually researched in the generated profiles. In a last and more systematic step, the HaloSeeker v1.0 application was applied to screen halogenated compounds presence. At this stage (Fig. 2) the obtained results confirmed (i) the capability of at least some of the applied sample preparation strategies to finally generate informative untargeted chemical profiles, (ii) the compatibility of such profiles with suspect screening process through a peak annotation based on appropriate reference library, and (iii) the suitability of the HaloSeeker v1.0 application to identify halogenated compounds from these data sets on a more systematic and semi-automated way. Now these approaches are still facing some limitations that include (i) the direct impact of the applied sample preparation procedure on the obtained results while no unique and definitive reference protocol may be proposed yet, (ii) the huge resources needed to go ahead with the identification of possible relevant markers then the urgent need of a more extended and QA/QC consolidated reference library, which is under creation, for a validated annotation based on MS and MS/MS data, and (iii) the lower exposure levels typically encountered in human matrices compared to those observed in environmental and food samples that appear clearly more challenging in term of required performances.

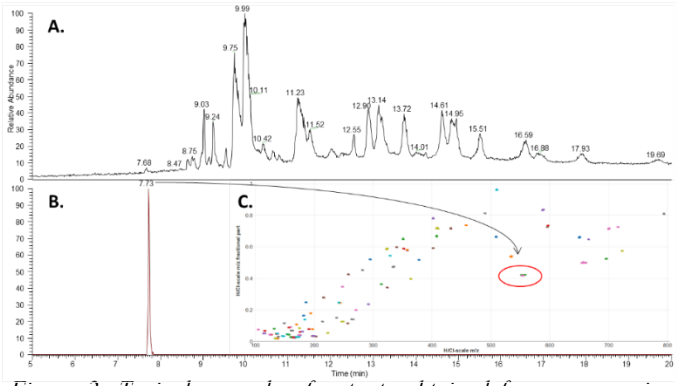

Figure 2: Typical example of outputs obtained from a meconium sample specimen: A. LC-HRMS total ion chromatogram B. ^{13}C-TBBPA (rt=7.73 min, m/z 554.7859) specific extracted ion chromatogram and C. mass defect based map sorted out by HaloSeeker v1.0. application pointing out halogenated features.

Acknowledgements

The authors want to express their acknowledgments to HBM4EU project (EU H2020 SC1-2016-RTD call, Cofund project nb. 733032) and "Région Pays de la Loire, France" for their financial support.

References

1. Sauve S, Desrosiers M. (2014); *Chem Cent J.* 8(1): 15
2. Schymanski EL, Williams AJ. (2017); *Environ. Sci. & Technol.* 51(10): 5357-5359
3. Kaufmann A. (2012); *Anal. Bioanal. Chem.* 403(5): 1233-1249

4. Cariou R, Omer E, Léon A, et al. (2016); Anal. Chim. Acta 936: 130-138
5. Antignac JP, Wasch K, Monteau F, et al. (2005); Anal. Chim. Acta 529: 129-136