Environmental Science & Technology

Viewpoint

pubs.acs.org/est

Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence

Emma L. Schymanski,*[,†] Junho Jeon,[†] Rebekka Gulde,[†,‡] Kathrin Fenner,[†,‡] Matthias Ruff,[†] Heinz P. Singer,[†] and Juliane Hollender*[,†,‡]

[†]Eawag: Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland

[‡]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092, Zurich, Switzerland

The increased availability of high resolution mass spectrometry (HR-MS) in chemical analysis has dramatically improved the detection and identification of compounds in environmental (and other) samples. This has opened up new research opportunities in environmental sciences, demonstrated by over 200 research papers per year, increasing strongly (source: SCOPUS keywords "high resolution mass spectromet", subject "envi"). The elucidation of small molecules such as emerging pollutants and their transformation products using HR-MS-based suspect and nontarget analysis is gaining in relevance, also in other fields (e.g., metabolomics, drug discovery, forensics). However, confidence in these HR-MS-based identifications varies between studies and substances, since it is not always possible or even meaningful to synthesize each substance or confirm them via complementary methods (e.g., nuclear magnetic resonance). These varying levels of confidence are very difficult to communicate to readers concisely and accurately.

In Figure 1 we propose a level system, which arose from intense discussions within our department, to ease the communication of identification confidence and form the basis of further discussions on this topic. This level system is not intended to replace guidance documents (e.g., EU Guideline 2002/657/EG), but specifically covers the new possibilities in HR-MS-based analysis.

Our discussion started with the levels published by the Metabolomics Standards Initiative (MSI),[1] as we experienced many cases that fitted "in between" their proposed levels. While Jeon et al.[2] first refined these levels, these were tailored to the specific investigation. The levels in Figure 1 reconcile differences in the two proposals, contain additional levels pertinent to screening methods and are clarified in the text below.

LEVELS 1 AND 2: STRUCTURE IDENTIFICATION

Here the experimental information strongly suggests one exact molecular structure, but the supporting evidence differs.

Level 1. ***Confirmed structure*** represents the ideal situation, where the proposed structure has been confirmed via appropriate measurement of a reference standard with MS, MS/MS and retention time matching. If possible, an orthogonal method should also be used.

Level 2. ***Probable structure*** indicates that it was possible to propose an exact structure using different evidence. For **Level 2a: Library** this involves matching literature or library spectrum data where the spectrum-structure match is unambiguous. Care is needed when comparing spectra recorded with different acquisition parameters (e.g., resolution, collision energy, ionization, MS level) to ensure the validity of the match and decision criteria should be clearly presented. Desirable additional evidence such as retention behavior would require, for example, a retention index for both the measured and matched spectrum, established for GC-MS but not yet sufficiently for LC-MS-based techniques. **Level 2b: Diagnostic** represents the case where no other structure fits the experimental information, but no standard or literature information is available for confirmation. Evidence can include diagnostic MS/MS fragments and/or ionization behavior, parent compound information and the experimental context. A good example is the hydroxylation at the tert-butyl group of irgarol[2] (see Figure 1). Although the breakdown into (a) and (b) is useful for research purposes, for practical reporting a level 2 classification may often suffice as "probable structure" indicates the confidence in the candidate.

LEVELS 3 TO 5: SUBSTANCE CLASS, FORMULA OR "MASS OF INTEREST"

Level 3. ***Tentative candidate(s)*** describes a "grey zone", where evidence exists for possible structure(s), but insufficient information for one exact structure only (e.g., positional isomers). Although there are many shades of gray (i.e., certainty) possible for many different situations, for all intents and purposes

Received: January 14, 2014
Accepted: January 17, 2014
Published: January 29, 2014

   dx.doi.org/10.1021/es5002105 | Environ. Sci. Technol. 2014, 48, 2097−2098

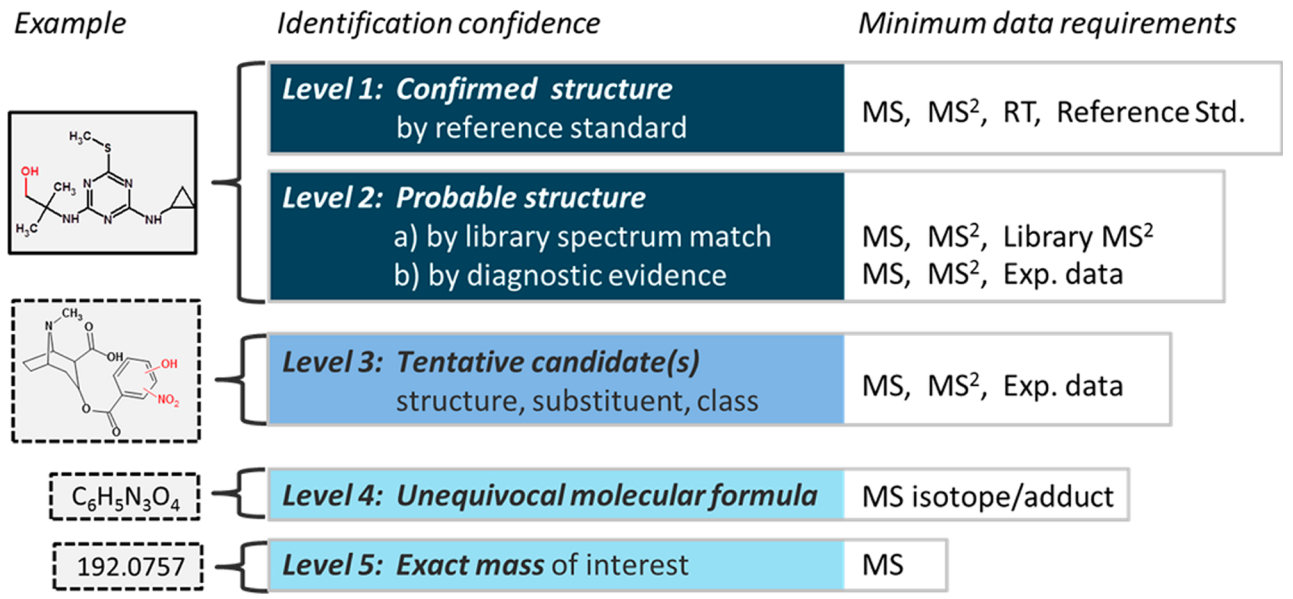

Figure 1. Proposed identification confidence levels in high resolution mass spectrometric analysis. Note: MS^2 is intended to also represent any form of MS fragmentation (e.g., MS^e, MS^n).

the exact structure remains speculative at level 3 (i.e., tentative in environmental, putative in metabolomics terms). Two examples are PCV TP251 from Prasse et al.,[3] where the position of the hydroxylation was not unequivocal, and TP-P3 from Bijlsma et al.,[4] where the position of the substituents could not be clarified (see Figure 1, substituents marked red). Further examples are top-ranked structures from in silico fragmentation of candidates from compound database searches, and/or suspects selected using additional information such as high predicted likeliness of being a transformation product, the number of references or retention behavior (e.g., Jeon et al.,[2] Hug et al.[5]). Although sublevels could be defined, cases are often so study-specific that sublevels would "drown" in details and reduce the generic applicability. Rather, sublevels should be defined on a per-study basis where evidence supporting different proposed structures is clearly presented, especially where their identity is central to the conclusions of the study.

Level 4. ***Unequivocal molecular formula*** is possible when a formula can be unambiguously assigned using the spectral information (e.g., adduct, isotope, and/or fragment information), but insufficient evidence exists to propose possible structures. The MS/MS could be uninformative, contain interferences or not even exist. However, the formula provides some information and is worth presenting as it can be traced in future studies.

Level 5. ***Exact mass*** (m/z) can be measured in a sample and be of specific interest for the investigation, but lack information to assign even a formula. Screening and nontarget methods allow the tracing of these masses in other investigations, but the level 5 indicates that no unequivocal information about the structure or formula exists. It is even possible to record the MS/MS of a level 5 mass and save it as an "unknown" spectrum in a database. This level should only apply to a few masses of specific interest, since it would be counterproductive to label all masses in a sample as level 5. Blank measurements should be used to ensure the substance does not arise from sample preparation or measurement.

To summarize, our aim with this viewpoint is to stimulate discussion across various research groups and organisations to establish an agreement on how to best communicate identification confidence for the exchange of identified, tentative, and unknown compounds via literature and databases for the benefit of all investigations.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: emma.schymanski@eawag.ch.

*E-mail: juliane.hollender@eawag.ch.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank our co-workers from the Department of Environmental Chemistry at Eawag who participated in the discussion. E.S. was funded by the EU Marie Curie Fellowship 299734.

REFERENCES

(1) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 211−221.

(2) Jeon, J.; Kurth, D.; Hollender, J. Biotransformation pathways of biocides and pharmaceuticals in freshwater crustaceans based on structure elucidation of metabolites using high resolution mass spectrometry. *Chem. Res. Toxicol.* **2013**, *26*, 313−324.

(3) Prasse, C.; Wagner, M.; Schulz, R.; Ternes, T. A. Biotransformation of the antiviral drugs Acyclovir and Penciclovir in activated sludge treatment. *Environ. Sci. Technol.* **2011**, *45* (7), 2761−2769.

(4) Bijlsma, L.; Boix, C.; Niessen, W. M. A.; Ibanez, M.; Sancho, J. V.; Hernandez, F. Investigation of degradation products of cocaine and benzoylecgonine in the aquatic environment. *Sci. Total Environ.* **2013**, *443*, 200−208.

(5) Hug, C.; Ulrich, N.; Schulze, T.; Brack, W.; Krauss, M. Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening. *Environ. Pollut.* **2014**, *184*, 25−32.