Improving target and suspect screening high resolution mass spectrometry workflows in environmental analysis by ion mobility separation

Alberto Celma [a], Juan V. Sancho [a], Emma L. Schymanski [b], David Fabregat-Safont [a], María Ibáñez [a], Jeff Goshawk [c], Gitte Barknowitz [c], Félix Hernández [a], Lubertus Bijlsma [a,*]

[a] Environmental and Public Health Analytical Chemistry, Research Institute for Pesticides and Water, University Jaume I, Avda. Sos Baynat s/n, E-12071 Castellón, Spain.

[b] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 6, avenue du Swing, L-4367 Belvaux, Luxembourg.

[c] Waters Corporation, Stamford Avenue, Altrincham Road, Wilmslow, Cheshire, SK9 4AX, U.K.

*Corresponding author: bijlsma@uji.es Tel.: +34 964 387382 Fax: +34 964 387368, ORCID: 0000-0001-7005-8775

0. Abstract

Currently, the most powerful approach to monitor organic micropollutants (OMPs) in environmental samples is the combination of target, suspect and non-target screening strategies using high resolution mass spectrometry (HRMS). However, the high complexity of sample matrices as well as the huge number of OMPs potentially present in samples at low concentrations pose an analytical challenge. Ion mobility separation (IMS) combined with HRMS instruments (IMS-HRMS) introduces an additional analytical dimension, providing extra information which facilitates the identification of OMPs. The collision cross section (CCS) value provided by IMS is unaffected by the matrix or chromatographic separation. Consequently, the creation of CCS databases and the inclusion of ion mobility within identification criteria are of high interest for an enhanced and robust screening strategy. In this work, a CCS library for IMS-HRMS, which is online and freely available, was developed for 556 OMPs in both positive and negative ionization modes using electrospray ionization. The inclusion of ion mobility data in widely adopted confidence levels for identification in environmental reporting is discussed. Illustrative examples of OMPs found in environmental samples are presented to highlight the potential of IMS-HRMS and to demonstrate the additional value of CCS data in various screening strategies.

Keywords: ion mobility separation; identification levels; collision cross section library; wide-scope screening; environment; high-resolution mass spectrometry

1. Introduction

High resolution mass spectrometry (HRMS) offers a powerful and suitable alternative to former targeted screening methods using low resolution mass spectrometry [1–5]. The high mass accuracy and resolution, together with the extensive variety of available acquisition modes for a wide mass-to-charge range (*m/z* 50-1000), make HRMS the technique of choice for wide-scope screening of thousands of organic micropollutants (OMPs) and their transformation products in aquatic matrices such as surface water or wastewater [6–9].

Hybrid HRMS mass analyzers, such as quadrupole – time-of-flight (QTOF), offer the possibility of sequentially acquiring information about the ionized molecule and fragment ions which vastly increases the identification potential of the screening strategy without significantly compromising the sensitivity of the analysis. However, when data independent acquisition (DIA) modes are used, fragmentation occurs not only for the compound of interest but for other co-eluting compounds and, therefore, fragments of multiple precursor ions can contribute to the fragmentation spectrum [6]. Particularly in complex matrices, interferences may be present due to fragment ions from precursors other than the one of interest. As a result, the possibility of misidentification increases. The large amount of data generated, the extensive databases used as well as the untargeted acquisition mode require meticulous strategies for the identification of compounds in the results obtained. The use of retention time and mass accuracy tolerance alone during screening analyses can lead to a notable number of false-positive findings [10,11]. To address this, different identification levels have been proposed in the scientific literature that depend on the information obtained by HRMS analysis [12–18]. The 5-level classification, from the most confident scenario (level 1, confirmed structure by reference standard) to the most uncertain scenario (level 5, exact mass of interest) proposed by Schymanski *et al*. [13] is currently widely used in the environmental literature. While discussions are ongoing for a revised set of identification levels, especially in the metabolomics community, these have not yet achieved community consensus.

The coupling of ion mobility separation (IMS) to HRMS instruments (IMS-HRMS) has promising applications for both targeted and untargeted screening. Briefly, IMS separates ions depending on their size, shape and charge in a gas phase, usually nitrogen (N_2) or helium (He), in the presence of an electric field [19]. Owing to their different mobility through the drift cell, IMS enables, in theory, the separation of isobaric or isomeric compounds that could not be previously resolved using liquid chromatography (LC) and/or HRMS [6,19–21]. The time needed by an ion to travel through the mobility separation device, the drift time (DT), is used for the determination of the collision cross section (CCS) of this particular ion based on the measurement of calibrating standards with already established CCS values for travelling wave IMS (TWIMS) or trapped IMS (TIMS) instruments, or based on the application of Mason-Schamp equation for drift tube IMS (DTIMS) instruments [22]. While measured DT is not comparable between different instruments [19], CCS is an instrument independent value that allows the comparison of CCS libraries with the actual measurement of a candidate in a sample even between different commercially available IMS-HRMS instruments [23]. In light of this, some publications dealing with the creation or use of CCS libraries for hundreds of compounds of different families have been published [24–29]. However, only very few studies have considered the inclusion of ion mobility data into the identification criteria [6,30–36]. Nuñez et al.[30] present an automated scoring engine for the processing of IMS-HRMS data by comparing empirical mass spectrometric and ion mobility data with *in silico* libraries. However, neither the chromatographic separation nor mass fragmentation was considered, which may increase the occurrence of false positives. The study conducted by Monge et al.[31] proposes a scoring system for the identification of metabolites in untargeted metabolomics as an update for previously reported confidence levels of Sumner et al.[18] through the combination of chromatography, mass spectrometry, ion mobility separation and nuclear magnetic resonance. However, these publications did not establish the minimum requirements for compound identification.

The aim of this work was: *i)* to develop an extensive database of CCS values for hundreds of OMPs in both positive and negative ionization mode, *ii)* to incorporate ion mobility information into a

87 widely community-adopted confidence levels for non-target and suspect screening strategies, and *iii)*
88 to demonstrate the improved utility of IMS-HRMS in screening of OMP in environmental samples via
89 illustrative examples gathered in different research projects. The information provided in this work
90 will be of interest in the near future, as it is expected that ion mobility will be incorporated as a
91 complementary criterion for reliable identification in different areas of analytical research.

92

2. Materials and Methods

2.1 Chemicals and materials

A total of 556 reference standards comprising illicit drugs, hormones, mycotoxins, new psychoactive substances, pesticides and pharmaceuticals were injected for the development of a CCS library and the subsequent application of the library to screening analyses. **Table S1** of the **Supporting Information** shows the complete set of compounds used in the study with their SMILES (simplified molecular-input line-entry system) representation, structure and measured CCS data. The database is also available on the NORMAN Suspect List Exchange website [37], the *Zenodo* online repository [38] and the CCS values have been integrated into PubChem [39]. JChem for Office (version 19.9.0.467) in Excel (from ChemAxon, www.chemaxon.com) was used for chemical parameters and structure calculation [40].

2.2 Instrumentation

Analyses were performed with a Waters Acquity I-Class UPLC system (Waters, Milford, MA, USA) connected to a VION IMS-QTOF mass spectrometer, using an electrospray ionization (ESI) interface operating in both positive and negative ionization mode.

The chromatographic column used was a CORTECS® C18 2.1 x 100 mm, 2.7 µm fused core column (Waters) at a flow rate of 300 µL min^{-1}. Gradient elution was performed using H_2O (A) and MeOH (B) as mobile phases, both with 0.01% formic acid. The initial percentage of B was 10%, which was immediately linearly increased to 90% over 14 min, followed by a 2 min isocratic period, then returned to initial conditions (at 16.1 min) with a 2 min equilibration of the column. The total run time was 18 min. The injection volume was 5 µL.

A capillary voltage of 0.8 kV and cone voltage of 40 V were used. The desolvation temperature was set to 550 °C, and the source temperature to 120 °C. Nitrogen was used as the drying gas and

nebulizing gas. The cone gas flow was 250 L h^{-1} and desolvation gas flow of 1000 L h^{-1}. The column temperature was set to 40 °C and the sample temperature to 10 °C. MS data were acquired using the VION in HDMSe mode, over the range *m/z* 50-1000, with N_2 as the drift gas, an IMS wave velocity of 250 m s^{-1} and wave height ramp of 20-50 V. Leucine enkephalin (*m/z* 556.27658 and *m/z* 554.26202) was used for mass correction in positive and negative ionization modes, respectively. Two independent scans with different collision energies were acquired during the run: a collision energy of 6 eV for low energy (LE) and a ramp of 28-56 eV for high energy (HE). A scan time of 0.3 s was set in both LE and HE functions. Nitrogen (≥ 99.999%) was used as collision-induced dissociation (CID) gas. All data were examined using an in-house built accurate mass screening workflow within the UNIFI platform (version 1.8.2) from Waters Corporation.

2.3 Collision cross section library

The whole set of reference standards was divided into different mixtures of up to 20 compounds depending on substances classes, based on previous knowledge about chromatographic separation and avoiding the presence of isobaric and isomeric compounds in the same mixture.

To obtain an accurate CCS value for each compound, the following workflow was used. Prior to the standard injection, the instrument was calibrated both for *m/z* measurements and CCS calculation following the manufacturer instructions. Then a 'system suitability test' (SST) containing 9 compounds was injected ten times to check the accuracy of the instrument measurements. **Table S2** shows the compounds included in the SST mix together with their molecular formula, SMILES and expected *m/z* and CCS value. Expected CCS values were provided by manufacturer: data measurements were performed in triplicate at three different pressures of N_2 with a minimum of eight different voltage gradients (RSD were typically < 0.3%) using a modified Synapt G2-Si (linear drift tube in place of the standard Travelling Wave cell). Next, in this study, reference standard mixtures at 1, 10 and 100 µg L^{-1} were injected in triplicates. After every mixture sequence (i.e. all injections of the three

concentration levels per mix), an SST injection was performed for a temporal evolution and the continuous control of the stability of the measurement. At the end of the sequence, the SST was run again (n=10). For the data to be considered acceptable, mass accuracy and CCS error (percentage deviation from the expected value) for the start, end and interspersed SST injections had to be within an acceptable tolerance (5 ppm in mass accuracy and 2% deviation in CCS). **Figure S1** shows the temporal evolution of mass and CCS accuracy across a representative injection run of standards during the CCS library building with interspersed SST in positive ionization mode. As expected, the empirical CCS deviation was below ± 2% deviation (mostly < 1%) ensuring a good robustness of CCS measurement.

The actual value of CCS for a compound was established by averaging the 9 values obtained at the three concentrations tested. In the cases where no signal was observed in the lower concentration level, the CCS value was established by averaging the data for the other concentration levels.

3. Results and discussion

3.1 Collision cross section library

A library containing CCS information of a total of 970 different adducts corresponding to 556 compounds (209 pesticides, 170 pharmaceuticals, 128 illicit drugs and new psychoactive substances, and 49 hormones and mycotoxins) was built to enhance target workflows with IMS. The library contains 472 protonated adducts ($[M+H]^+$), 248 sodium adducts ($[M+Na]^+$), 26 water loss in-source fragments ($[M+H-H_2O]^+$), 9 ammonia loss in-source fragments ($[M+H-NH_3]^+$), 162 deprotonated adducts ($[M-H]^-$), 25 chlorinated adducts ($[M+Cl]^-$) and 31 formate adducts ($[M+HCOO]^-$). The complete library is available in **Table S1** of the **Supporting Information** and also publicly available on the NORMAN Suspect List Exchange website [37], *Zenodo* online repository [38] and on PubChem [39].

As previously mentioned, the CCS for each adduct was obtained as an average value of the replicates injected at 3 different concentration levels. In general, the relative standard deviation (RSD) observed between replicates was 0.1-0.3%., and no trend was observed in the CCS measurement precision depending on the concentration of the reference standard. As an example of the main trend, **Figure S2** shows the RSD in the measurement of CCS value for a set of 46 pesticides. The robustness of CCS measurements across injections supports the use of ion mobility as a powerful and promising tool for improved identification of candidates.

In general, the CCS value of a certain adduct is strongly related to the molecular mass, such that different adducts of the same molecule generally result in different CCS values due to the difference (mainly in size) of the ion incorporated in or removed from the structure [41] (**Figure S3**). However, the non-perfect linear correlation between CCS and molecular mass (see **Section S1**) highlights that CCS values are also affected by other molecular parameters, such as the chemical backbone, ionization site or how the molecule can rearrange its structure to stabilize the electric charge. This is particularly the case of X-ray agents *ioversol, iopromide, iomeprol* and *iopamidol* that with a molecular mass of approximately 800 Da yield an unexpected low CCS value due to the intrinsic

characteristics of the chemical backbone and substituents (**Figure S4**). Furthermore, it is noteworthy that among the complete set of 556 reference standards analyzed, only protomers were observed for the quinolone antibiotics *sarafloxacin* (I: 187.09 Å^2 and II: 202.00 Å^2), *ciprofloxacin* (I: 175.38 Å^2 and II: 188.89 Å^2) and *norfloxacin* (I: 171.88 Å^2 and II: 187.60 Å^2). In these particular cases, protonation on the cyclic ketone or the piperazine moiety [21] (**Figure S5**) resulted in different conformational changes, being distinct enough to be resolved by IMS. Consequently, these protomers could be qualitatively identified in real samples using IMS-HRMS without the need to consider abundances within the identification strategy.

A detailed and comprehensive discussion concerning the general trends observed for CCS values as well as these particular cases can be found in the **Supporting Information** (**Section S1** and **Figures S3-S6**).

3.2 Identification levels for IMS-HRMS screening strategies

Having well-defined criteria accepted by the scientific community for the identification of candidates in screening strategies is pivotal for an accurate dissemination of results and comparison with other studies. For that purpose, Schymanski et al. [13] proposed a 5-level criteria for the identification of small molecules using chromatographic separation coupled to HRMS. This classification also included cases in which the solely available information was the molecular formula or exact mass (*level 4 -unequivocal molecular formula-* and *level 5 -exact mass of interest-*, respectively). At these levels, insufficient information is available to propose tentative candidates. However, the data available for *level 3 -tentative candidate(s*)- allows the proposition of more than one chemical structure (for example, positional isomers). Candidate structures elucidated by *in silico* fragmentation tools are usually most appropriately classified as level 3 features. *Level 2 -probable structure*- is related to candidates that could unambiguously be assigned to a certain chemical structure based on the scientific literature, mass spectral libraries or diagnostic evidence. Finally, *level*

206 *1 -confirmed structure-* represents the ideal situation where chromatographic and mass
207 spectrometric evidence are confirmed with a reference standard. These criteria have been widely
208 adopted by environmental researchers [3,42–44]. Even though the fragmentation information gathered
209 with HRMS instruments often determines the potential for identification of candidates, the utilization
210 of additional orthogonal methods is recommended [13,45]. In this sense, the incorporation of IMS-HRMS
211 in screening strategies permits to gain even more confidence in the identification and adds an extra
212 dimension to further improve screening analyses [46]. The inclusion of IMS may also help to discriminate
213 between isomeric level 3 candidates and move one of them up to level 2. In this work,
214 recommendations are given how to apply the 5-level criteria from Schymanski et al. [13] for users of
215 state-of-the-art IMS-HRMS instruments. The analytical experience gathered during CCS library
216 building has been taken into account in proposing these criteria. The classification is intended to
217 enhance these widely adopted criteria and suggest how to apply them to IMS-HRMS measurements,
218 as well as to contribute to the community discussion on how to incorporate multiple lines of evidence
219 into identification confidence schemes.

220 **Figure 1** shows the different levels of confidence proposed in this work for the identification
221 of a compound using LC-IMS-HRMS based on chromatographic, ion mobility and mass spectrometric
222 parameters. Typically, the accuracy of empirical data for mass spectrometric measurements is
223 established at a maximum deviation of 5 ppm (or 2 mDa) from the theoretical *m/z,* as well as
224 compliance with the expected isotopic pattern [45]. However, as most HRMS instruments can provide
225 higher levels of accuracy, the threshold for deviation in mass spectrometric measurements could
226 nowadays be adjusted to 3 ppm. The criterion for retention time is less harmonized among the
227 scientific community, and it is surely more debatable. In this work, a maximum retention time
228 deviation of ± 0.1 min from that of the standard is proposed in agreement with SANTE 2017 guideline
229 [45], implying that both sample and reference standard are run under the same chromatographic
230 conditions. However, the SANTE guideline is applied for food analysis and not environmental analysis.
231 As such, the maximum deviation is an indicative value, and should be adapted depending on the

particular conditions of the analysis. The results obtained and the examples presented in this study may open the dialogue to develop more applicable criteria for environmental studies, where matrix effects can potentially lead to high deviations. In the case of CCS, there are no regulatory guidelines yet and, therefore, there is still no agreement on which is the maximum threshold permitted for CCS deviation. Based on the experience gathered during the development of the CCS library included in this study, together with the background knowledge acquired during screening campaigns using IMS-HRMS, we propose a maximum deviation of 2% for CCS values. Depending on the availability of reference standards, in addition to the accuracy of the acquired empirical data, the level classification previously proposed by Schymanski et al. [13] is updated for IMS-HRMS users as follows:

- ***Level 5 –exact mass of interest-*** represents the level where least information about the candidate is available. However, the exact mass together with its specific CCS value is considered relevant for the study and worth being monitored in future campaigns.
- ***Level 4 –unequivocal molecular formula-*** encompasses the cases where a molecular formula can be assigned. MS, RT and CCS information alone, without fragmentation information, is commonly not enough to propose a potential structure and, therefore, RT and CCS data measured typically do not provide sufficient additional information for identification.
- ***Level 3 –tentative candidate(s)-*** comprises the cases where different chemical structures are compatible with the empirical RT, CCS and MS data but not enough information is available to distinguish which one is the most likely. In these cases, empirical information about the chromatography, ion mobility and mass spectrometry behavior of the candidates could be compared with *predicted* parameters. The predictions about the value for RT, CCS or mass fragmentation can give extra confidence to the proposed tentative candidates [47–54], or help prioritize potential candidates [46]. Despite the additional value of such tools, the predicted values should be considered as an orientation. Hence, rejecting candidate structures solely because of a disagreement between empirical and predicted values is not recommended. The utilization of

retention time indexing strategies (RTI) to compare the empirical data with online available databases can also provide extra confidence in the tentative identification of candidates [11,55].

- ***Level 2 –probable structure-*** indicates that an exact structure could be proposed based on experimental evidence. This level can be divided into two sub-levels. ***Level 2a – probable structure by library match,*** comprises those cases when the structure of the compound is proposed based on the agreement between experimental data and literature or available libraries for both HRMS and CCS. The high robustness of CCS measurement between different instruments permits the utilization of home-made or third-party CCS libraries to compare with experimental data, reaching a high level of confidence in the identification. ***Level 2b – probable structure by diagnostic evidence-*** makes use of the available data to unambiguously propose a structure in the case that no other candidate fits the empirical evidence. The slender difference between level 2b and level 3 is the fact that in level 2b *only one structure* satisfies the experimental evidence (and all other candidates can be eliminated), while in level 3 there is not enough evidence to distinguish between more than one candidate structures. Level 2b identifications are generally quite rare and often require experimental context (*e.g.* transformation experiments where the parent is known). For both level 2a and 2b, a reference standard is required for a final confirmation of the structure to achieve the highest confidence (Level 1).
- **Level 1 – confirmed structure-** is the ideal situation, where the empirical data fully agrees with that of a reference standard in terms of MS, fragmentation, retention time and CCS. This is the case where the highest confidence in the identification is obtained with HRMS. For a proper level 1 identification, all orthogonal techniques (MS, fragmentation, RT and CCS) should be in accordance with that of the reference standard. However, the comparison of reference standard information to empirical data from samples can result in different sublevels of identification confidence. Hence, the combined adoption of an Identification Points (IP) scoring system to

address this often challenging task is proposed in agreement with the Commission Decision 2002/657/EC [56] and recently reported IP proposals [42,57]. Briefly:

Empirical MS information matches the reference standard	1 IP
Empirical RT information matches the reference standard	1 IP
Empirical CCS information matches the reference standard	1.5 IP
Two or more matching HRMS fragments	2.5 IP
Minimum IP for Level 1 identification with CCS	***5 IP***

Although the ideal situation should yield a maximum of 6 IP (1 for MS, 1 for RT, 2.5 for HRMS fragmentation and 1.5 for CCS), a minimum value of 5 IP should be considered sufficient for the confirmation of the identity. While some studies have proposed different criteria for the identification of compounds [13,42,45,56,57], very few consider the likely case in which any of the parameters measured (retention time, CCS or mass spectrometric data) fails to meet the requirements. In such cases, establishing the level of confidence of the identification is not a straightforward decision and usually further investigation is required to accurately report the detection. Mass spectrometric data can be affected by several factors and, therefore, when the mass accuracy is barely higher than the established threshold different actions can be followed. The immediate verification should be the instrument performance by checking the mass accuracy with a set of reference standards injected alongside the sample injection run as quality controls. In addition, spectral interferences can affect the mass accuracy, which can be improved by a reinjection of the sample with enhanced resolution (which is often not available for many instruments). Also important is the dependence of mass error on the signal intensity. The lower the number of ions measured, the higher the mass error; therefore, low abundant fragments often show higher mass errors [58]. The same applies for high intensity ions, which can distort mass accuracy because of detector saturation.

On another point, either a RT error slightly higher than 0.1 min or a ΔCCS faintly greater than 2% would require the fortification of the original sample with the candidate compound and/or modification in the chromatographic conditions to fully confirm its identity. However, in our own experience the chance of having deviations greater than 2% in the CCS is low because of the

304 robustness of the CCS measurements. So, not all the parameter deviations should be weighted
305 uniformly, since retention time is more prone to be shifted by sample matrix [11]. Consequently, a
306 variation in RT slightly questions the identification of a candidate that perfectly matches the reference
307 standard for HRMS data and CCS. On the contrary, a CCS deviation higher than 2% strongly questions
308 the identification. In this sense, the minimum requirement for identity confirmation as Level 1 is
309 established at 5 IP, which already considers the possibility of deviations in RT but needs an agreement
310 of CCS. For those particular cases when empirical data do not completely fit the reference standard,
311 reporting the candidate at the corresponding level with a reduced score (< 5 IP) and accompanied with
312 a clarification on the parameter failing in the requirements is proposed in order to comprehensively
313 report the data (e.g. highlighted with an asterisk as Level *x**). Obviously, the fact that one parameter
314 (commonly RT and mass error) is slightly out of tolerance (typically 0.1 min and 5 ppm, respectively)
315 would reduce the confidence, but might not be as crucial as other important parameters, such as CCS
316 deviation or the presence of fragment ions in agreement with experimental data or spectral libraries.
317

318 **3.3 Application to environmental water samples**

319 The application of efficient strategies for the wide-scope screening of OMPs in environmental
320 samples has become essential. While strategies involving HRMS may lead to misidentifications in some
321 cases [11,35,59–63], IMS-HRMS instruments provide an extra identification parameter that improves the
322 performance and helps to reduce the number of false positives/negatives [10]. In this section, we
323 highlight different identification scenarios using the developed CCS library to show the potential of
324 IMS-HRMS in environmental analysis. It summarizes some of the experience gathered through the
325 utilization of IMS-HRMS in different research studies.

326 **Figure 2** shows the confirmation at level 1 of *4-acetamidoantipyrin* in surface water from a
327 nature reserve in Spain after pre-concentrating the sample using solid phase extraction. Despite being
328 a protected area, the sampling location was contaminated through the introduction of the effluent

stream of an urban wastewater treatment plant. The presence of this human metabolite of *metamizole* can be attributed to an inefficient removal of this pharmaceutical metabolite during wastewater treatment. The entry in the CCS library for the reference standard of *4-acetamidoantipyrin* showed a retention time of 3.01 min with a CCS value of 154.06 $Å^2$ for the protonated molecule, and HE fragment ions with *m/z* 228.1132 and *m/z* 104.0495. The candidate observed in the surface water sample eluted at 3.09 min (+ 0.08 min of deviation) and both the protonated molecule and the HE fragments were observed at their *m/z* (mass error <5 ppm). In addition, the experimental CCS for the candidate was 154.08 $Å^2$, which only deviates by + 0.01 % from the standard. In the light of the full agreement of all these measurements and using the criteria previously proposed, the identification of this candidate as *4-acetamidoantipyrin* was confirmed as level 1 with 6 IP (MS + RT + >2 HRMS fragments + CCS).

As stated above, in environmental samples, the matrix composition can strongly influence compound retention and, therefore, the RT for most of the analytes [11]. This fact may lead to a notable increase in the number of misidentifications because of significant RT deviation between standard and sample. Nevertheless, the excellent reproducibility observed for CCS values, and the fact that this parameter is not affected by matrix composition, provides extra identification power, which is especially useful for compounds partially out of the confirmation criteria. As an illustrative example, **Figure 3** shows the detection of *thiabendazole*, a fungicide used to control fungal diseases in fruits and vegetables, in the mouth of a Spanish river in the Mediterranean basin identified at level 1* (i.e. RT deviation beyond limits). The RT for *thiabendazole* reference standard was 3.27 min with a CCS value of 137.44 $Å^2$. However, the RT in the sample was 3.51 min, and seemed notably affected by matrix composition, with a deviation of + 0.24 min. The RT difference between standard and sample is far in excess of the typical criterion established for confirmation (± 0.1 min) (**Figure 1**) not earning, in consequence, the 1 IP for RT agreement. On the contrary, ion mobility was not affected by the matrix and resulted in a CCS value of 137.27 $Å^2$, which only deviated -0.12 % from the standard. In addition, the protonated molecule and three fragments were observed with mass error below 5 ppm. Under

these conditions, the identity of this compound as *thiabendazole* could be confirmed at level 1* with 5 IP (MS + >2 HRMS fragments + CCS). This example illustrates that RT affected by matrix composition may hamper the confirmation process in wide-scope screening, while the application of CCS provides the extra value needed for confirmation. In cases in which the RT notably deviates from the standard, some guidelines recommend to spike the sample with the candidate standard to confirm the identity of the compound [45]. However, the additional confidence gathered with the CCS measurement in a single-injection reduces time and costs of spiking and re-injecting the sample, as two separate pieces of evidence already exist (MS + >2 HRMS fragments + CCS). This is of special interest in environmental screening strategies where ion mobility can be included as an additional criterion for reliable identification in forthcoming guidelines in different fields of analytical research.

Moreover, the robustness of CCS measurements allows this parameter to be used also as an extra point of confidence when the reference standard is not available. Prediction tools can offer an estimation of the CCS value that can easily be compared to the measured value of the tentatively identified compound [49,51,54]. This is the case of the tentative identification of *valifenalate* in spinach samples reported by Bijlsma et al [49], who found a potential positive with an experimental CCS of 196.97 $Å^2$, although no reference standard was available for confirmation. By means of a predictive model developed using Artificial Neural Networks, the authors were able to predict CCS values for small molecules. The predicted CCS for *valifenalate* was 194.34 $Å^2$ which deviated only 1.4% from the experimental value, resulting in higher confidence in the tentative identification. Similarly, in the present work, a suspect screening of pesticides in surface water revealed a potential positive of *tricyclazole*, commonly used for the control of *Magnaporthe grisea* fungi during rice blast. The candidate peak ($[M+H]^+$; *m/z* 190.04354) showed a RT of 5.74 min with the fragment ions *m/z* 163.03251, *m/z* 136.02158, *m/z* 109.01057 and *m/z* 92.04961, and a measured CCS of 133.93 $Å^2$ (**Figure 4**). HRMS information contained in the free online-available mass spectral database *Mass Bank of North* America [64] included four fragment ions for *tricyclazole* (*m/z* 163.0333 - $C_8H_7N_2S^+$, *m/z* 136.0220 – $C_7H_6NS^+$, *m/z* 109.0106 – $C_6H_5S^+$, and *m/z* 92.0496 – $C_6H_6N^+$), which fully agreed with our

experimental data. Additionally, the CCS prediction model developed by Bijlsma et al. [49] predicted a CCS value of 136.2626 $Å^2$, with a deviation of +1.74% from the experimental measurement. Although the reference standard should be acquired for the full confirmation of the identity, the agreement of all these parameters gave high confidence to the tentative identification of *tricyclazole* in the surface water sample, at level 2a. At a later stage, reference standard was purchased and it allowed the identification of *tricyclazole* at level 1 since fully agreement between empirical and reference standard data was achieved (reference standard data: RT 5.78 min, CCS 132.98 $Å^2$, $[M+H]^+$ *m/z* 190.04354 – $C_9H_8N_3S^+$, and fragments *m/z* 163.03245 - $C_8H_7N_2S^+$, *m/z* 136.02155 – $C_7H_6NS^+$, *m/z* 109.01065 – $C_6H_5S^+$, and *m/z* 92.04948 – $C_6H_6N^+$).

It is worth emphasizing at this point that the proposed levels of confidence in the identification and the discussion of the examples are both based on the knowledge gathered by the authors through the use of IMS-HRMS in several studies. The expertise of the mass spectrometrist should be the rationale behind the application of the levels of confidence for IMS-HRMS analyses. The results from the screening should be deeply reviewed by experienced researchers and data critically discussed if there is a deviation on the criteria (such as mass spectrometric accuracy or RT deviation), avoiding immediate exclusion of potential positives by an automated application of strict criteria. Although the use of mass spectrometric databases and/or predictive models give more confidence into the results, the experience of the analyst is crucial in the elucidation of compounds through the utilization of common mass fragmentation rules [65]. Additionally, the sample origin and its characteristics can be determinant when considering potential candidate structures for the empirical features, and this knowledge can only come up from a human being and not (yet) from an automated processing software.

3.4 Strengths and limitations of IMS-HRMS

The use of IMS-HRMS for wide-scope screening of OMPs in environmental analyses is a powerful instrument for an enhanced analytical performance. One of the major benefits of ion

406 mobility, which is usually insufficiently acknowledged in the scientific literature, is the simplification
407 of mass spectral interpretation. In addition to separating chromatographically co-eluting ions, ion
408 mobility also filters both LE and HE spectra, removing ions that do not belong to the candidate of
409 interest [6,33]. This includes the removal of other co-eluting compounds that could be producing HE
410 fragments as well as the reduction of matrix-endogenous interferences, thereby decreasing the
411 number of peaks in a spectrum to be interpreted and thus also the risk of false fragment library
412 matching. As an illustrative example, **Figure 5** shows the comparison of LE and HE spectra of
413 *benzoylecgonine*, the main metabolite resulting from cocaine use, of a reference standard (**Figure 5a**)
414 and a positive finding in a wastewater sample with the drift time aligned (**Figure 5b**) and non-drift
415 time aligned spectra (**Figure 5c**). When no ion mobility separation is applied (**Figure 5c**), the spectrum
416 is much more populated with ions that do not originate from *benzoylecgonine* than in the drift time
417 aligned spectrum (**Figure 5b**), with a quality comparable to the reference standard spectrum. The fact
418 that IMS-HRMS provides 'clean' spectra, because of matrix interferences and co-eluting ions
419 separation, strongly facilitates the spectral interpretation and identification process in wide-scope
420 screening strategies, especially in comparison with non-ion mobility HRMS instruments [6].

421 Despite the benefits of IMS-HRMS, some limitations should also be mentioned. The IMS-HRMS
422 instrument used in this study, VION IMS-QTOF mass spectrometer from Waters, has the mobility
423 separation cell located between the ionization source and the mass analyzer. Therefore, ions
424 constantly produced in the ionization source need to be packed in small groups of ions every 14 ms in
425 order to separate them by their mobility. To this aim, a trap is located before the separation cell.
426 Unfortunately, the release process of the trapped ions seemed to cause additional fragmentation in
427 the LE function for labile (de)protonated molecules. As an example, **Figure 6** highlights the LE
428 fragmentation for the new psychoactive substance *2,5-dimethoxy-4-ethylphenethylamine* (*2C-E*). A
429 routine revision of HRMS data in screening analyses is often performed making use of the
430 aforementioned advantages of IMS-HRMS, and therefore, revising drift time aligned MS data. That
431 would be the case of spectra shown in **Figure 6a**, which apparently is a proper spectrum for a potential

positive of *2C-E* with a protonated adduct *m/z* 210.14883 in the LE function and significant fragments in the HE function. However, the non-drift time aligned MS spectra (**Figure 6b**), shows that the most abundant ion does not really correspond to the protonated adduct (*m/z* 210.14886, green shadowed) but to the ammonia loss fragment (*m/z* 193.12222, blue shadowed) followed by other LE fragments such as *m/z* 178.09871 and *m/z* 163.07529. Further investigation revealed that all these ions showed different ion mobility (different DT) (**Figure 6c**), which confirms that they were produced at some stage before the mobility separator device. The extra fragmentation observed was confirmed to be a 'pre-mobility' fragmentation behavior but not an enhanced 'in-source' fragmentation since the fragmentation did not occur when working in conventional MS^E mode (i.e. with no mobility separation) (**Figure S7**). This 'pre-mobility' fragmentation produced a ten-fold decrease in the intensity of the protonated adduct of *2C-E*, which may hamper the discovery of this compound in a real-sample scenario. Therefore, this particular 'pre-mobility' fragmentation may have negative consequences in environmental analysis where most of detections and subsequent identifications are based on the presence of the protonated molecules. The reduced intensity of the protonated adduct of the molecule can favor false negative identifications, especially for low abundant and very labile compounds such as some psychoactive substances in wastewater samples. It is noteworthy that this particular example was observed using a VION IMS-QTOF instrument and, therefore, cannot be directly extrapolated to other IMS instrument. However, the nature of IMS separation and the building of mobility devices make it feasible that other manufacturer instruments may suffer from a similar 'pre-mobility' phenomenon.

In summary, although the above-mentioned limitations have been observed, IMS-HRMS has strong potential for wide-scope screening of OMPs and notably facilitates screening strategies in highly complex matrices. The much cleaner drift time aligned MS spectra enhances the identification process, and the excellent robustness of CCS measurements in different matrices enables CCS prediction tools to help in tentative identification of candidates when the reference standard is not available. This enhances the confirmation rate if the reference is eventually acquired for confirmation.

458 Furthermore, freely and/or commercially available CCS libraries, both measured and computational,
459 can be used to facilitate target/suspect screening, due to the stability and extra identification power
460 provided by ion mobility when RT shifts are likely to occur.

461 In this paper, we provided a publicly available dataset of 970 CCS values, illustrated the
462 potential of IMS-HRMS and suggested IMS-based scoring criteria to enhance commonly applied
463 identification reporting levels in environmental analyses. The work was supported by real examples
464 taking into account the additional value of ion mobility and demonstrated an improved screening
465 strategy for OMPs in environmental samples based on state-of-the-art IMS-HRMS technologies.

466

4. Acknowledgements

The authors appreciate the kind collaboration of our colleagues Tania Portolés, Elena Pitarch, Eduardo Beltrán and Leticia Lacalle who kindly contributed to the CCS library building process, as well as Evan Bolton and Jian (Jeff) Zhang (NCBI, NLM, NIH) for integrating these CCS values in PubChem. A. Celma acknowledges the Spanish Ministry of Economy and Competiveness for his predoctoral grant (BES-2016-076914). D. Fabregat-Safont acknowledges Spanish Ministry of Education, Culture and Sports in Spain for his predoctoral grant (FPU15/02033). L. Bijlsma acknowledges the financial support by the Swedish Research Council (FORMAS) for his post-doctoral grant in "Effect-directed analysis as a tool towards a non-toxic environment - identification of mixture effects and toxicity drivers in water (DANTE)" project (2018-02256). ELS is supported by an ATTRACT fellowship from the Luxembourg National Research Fund (FNR, Grant A18/BM/12341006). Authors acknowledge the financial support from Universitat Jaume I (UJI-B2018-19 and UJI-B2018-55). F. Hernández acknowledges the financial support of Spanish Ministry of Science, Innovation and Universities (RTI2018-097417-B-100) and of Generalitat Valenciana (Research Group of Excellence Prometeo 2019/040).

5. Supporting information

This document, available online, includes 2 tables, 7 figures and a short section to have supportive visual information on the written text. Table S1: contain the complete database for 970 (de)protonated adducts, sodiated adducts, ammonia loss, water loss, chlorine adducts and formate adducts and can be consulted online at https://www.norman-network.com/nds/SLE/ (List S61); https://doi.org/10.5281/zenodo.3549476; https://pubchem.ncbi.nlm.nih.gov/source/23819#data=Annotations. Table S2: Information on the compounds used for the "system suitability test". Figure S1: Temporal evolution of mass accuracy and CCS accuracy, Figure S2: General trend observed in the ion mobility measurement during CCS library building process, Figure S3: CCS values of different adducts versus the neutral mass of the molecule,

Figure S4: Chemicals structures of X-ray agents ioversol, iopromide, iomeprol and iopamidol, Figure S5: Different protomers separated by IMS-HRMS, Figure S6: Absolute variation of CCS value observed in molecules showing more than one ionic species, Figure S7: HRMS spectra for 2C-E. Section S1: Main trends and particularities observed during CCS library development.

6. References

(1) Ibáñez, M.; Sancho, J. V.; Bijlsma, L.; van Nuijs, A. L. N.; Covaci, A.; Hernández, F. Comprehensive Analytical Strategies Based on High-Resolution Time-of-Flight Mass Spectrometry to Identify New Psychoactive Substances. *TrAC - Trends Anal. Chem.* **2014**, *57*, 107–117. https://doi.org/10.1016/j.trac.2014.02.009.

(2) Leendert, V.; Van Langenhove, H.; Demeestere, K. Trends in Liquid Chromatography Coupled to High-Resolution Mass Spectrometry for Multi-Residue Analysis of Organic Micropollutants in Aquatic Environments. *TrAC - Trends Anal. Chem.* **2015**, *67*, 192–208. https://doi.org/10.1016/j.trac.2015.01.010.

(3) Bletsou, A. A.; Jeon, J.; Hollender, J.; Archontaki, E.; Thomaidis, N. S. Targeted and Non-Targeted Liquid Chromatography-Mass Spectrometric Workflows for Identification of Transformation Products of Emerging Pollutants in the Aquatic Environment. *TrAC - Trends Anal. Chem.* **2015**, *66*, 32–44. https://doi.org/10.1016/j.trac.2014.11.009.

(4) López, A.; Yusà, V.; Millet, M.; Coscollà, C. Retrospective Screening of Pesticide Metabolites in Ambient Air Using Liquid Chromatography Coupled to High-Resolution Mass Spectrometry. *Talanta* **2016**, *150*, 27–36. https://doi.org/10.1016/j.talanta.2015.11.068.

(5) Knolhoff, A. M.; Croley, T. R. Non-Targeted Screening Approaches for Contaminants and Adulterants in Food Using Liquid Chromatography Hyphenated to High Resolution Mass Spectrometry. *J. Chromatogr. A* **2016**, *1428*, 86–96. https://doi.org/10.1016/j.chroma.2015.08.059.

(6) Regueiro, J.; Negreira, N.; Berntssen, M. H. G. Ion-Mobility-Derived Collision Cross Section as an Additional Identification Point for Multiresidue Screening of Pesticides in Fish Feed. *Anal. Chem.* **2016**, *88* (22), 11169–11177. https://doi.org/10.1021/acs.analchem.6b03381.

(7) Hernández, F.; Bakker, J.; Bijlsma, L.; de Boer, J.; Botero-Coy, A. M.; Bruinen de Bruin, Y.;

Fischer, S.; Hollender, J.; Kasprzyk-Hordern, B.; Lamoree, M.; López, F. J.; Laak, T. L. te.; van Leerdam, J. A.; Sancho, J. V.; Schymanski, E. L.; de Voogt, P.; Hogendoorn, E. A. The Role of Analytical Chemistry in Exposure Science: Focus on the Aquatic Environment. *Chemosphere* **2019**, *222*, 564–583. https://doi.org/10.1016/j.chemosphere.2019.01.118.

(8) Hernández, F.; Ibáñez, M.; Botero-Coy, A. M.; Bade, R.; Bustos-López, M. C.; Rincón, J.; Moncayo, A.; Bijlsma, L. LC-QTOF MS Screening of More than 1,000 Licit and Illicit Drugs and Their Metabolites in Wastewater and Surface Waters from the Area of Bogotá, Colombia. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6405–6416. https://doi.org/10.1007/s00216-015-8796-x.

(9) Hernández, F.; Sancho, J. V; Ibáñez, M.; Abad, E.; Portolés, T.; Mattioli, L. Current Use of High-Resolution Mass Spectrometry in the Environmental Sciences. *Anal. Bioanal. Chem.* **2012**, *403* (5), 1251–1264. https://doi.org/10.1007/s00216-012-5844-7.

(10) Regueiro, J.; Negreira, N.; Hannisdal, R.; Berntssen, M. H. G. Targeted Approach for Qualitative Screening of Pesticides in Salmon Feed by Liquid Chromatography Coupled to Traveling-Wave Ion Mobility/Quadrupole Time-of-Flight Mass Spectrometry. *Food Control* **2017**, *78*, 116–125. https://doi.org/10.1016/j.foodcont.2017.02.053.

(11) Celma, A.; Bijlsma, L.; López, F. J.; Sancho, J. V. Development of a Retention Time Interpolation Scale (RTi) for Liquid Chromatography Coupled to Mass Spectrometry in Both Positive and Negative Ionization Modes. *J. Chromatogr. A* **2018**, *1568*, 101–107. https://doi.org/10.1016/j.chroma.2018.07.030.

(12) Nácher-Mestre, J.; Ibáñez, M.; Serrano, R.; Boix, C.; Bijlsma, L.; Lunestad, B. T.; Hannisdal, R.; Alm, M.; Hernández, F.; Berntssen, M. H. G. Investigation of Pharmaceuticals in Processed Animal By-Products by Liquid Chromatography Coupled to High-Resolution Mass Spectrometry. *Chemosphere* **2016**, *154*, 231–239. https://doi.org/10.1016/j.chemosphere.2016.03.091.

(13) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098. https://doi.org/10.1021/es5002105.

(14) Rochat, B. Proposed Confidence Scale and ID Score in the Identification of Known-Unknown Compounds Using High Resolution MS Data. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 709–723. https://doi.org/10.1007/s13361-016-1556-0.

(15) Herrera-Lopez, S.; Hernando, M. D.; García-Calvo, E.; Fernández-Alba, A. R.; Ulaszewska, M. M. Simultaneous Screening of Targeted and Nontargeted Contaminants Using an LC-QTOF-MS System and Automated MS/MS Library Searching. *J. Mass Spectrom.* **2014**, *49* (9), 878–893. https://doi.org/10.1002/jms.3428.

(16) Rivier, L. Criteria for the Identification of Compounds by Liquid Chromatography-Mass Spectrometry and Liquid Chromatography-Multiple Mass Spectrometry in Forensic Toxicology and Doping Analysis. *Anal. Chim. Acta* **2003**, *492* (1–2), 69–82. https://doi.org/10.1016/S0003-2670(03)00889-4.

(17) Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* **2018**, *8* (2), 1–23. https://doi.org/10.3390/metabo8020031.

(18) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; 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. https://doi.org/10.1007/s11306-007-0082-2.

(19) Lee, J. W. Basics of Ion Mobility Mass Spectrometry. *Mass Spectrom. Lett.* **2017**, *8* (4), 79–89.

https://doi.org/10.5478/MSL.2017.8.4.79.

(20) Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2008**, *80* (24), 9689–9699. https://doi.org/10.1021/ac8016295.

(21) D'Atri, V.; Causon, T.; Hernandez-Alba, O.; Mutabazi, A.; Veuthey, J. L.; Cianferani, S.; Guillarme, D. Adding a New Separation Dimension to MS and LC–MS: What Is the Utility of Ion Mobility Spectrometry? *J. Sep. Sci.* **2018**, *41* (1), 20–67. https://doi.org/10.1002/jssc.201700919.

(22) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrom. Rev.* **2019**, 1–30. https://doi.org/10.1002/mas.21585.

(23) Hinnenkamp, V.; Klein, J.; Meckelmann, S. W.; Balsaa, P.; Schmidt, T. C.; Schmitz, O. J. Comparison of CCS Values Determined by Traveling Wave Ion Mobility Mass Spectrometry and Drift Tube Ion Mobility Mass Spectrometry. *Anal. Chem.* **2018**, *90* (20), 12042–12050. https://doi.org/10.1021/acs.analchem.8b02711.

(24) Hernández-Mesa, M.; Le Bizec, B.; Monteau, F.; García-Campaña, A. M.; Dervilly-Pinel, G. Collision Cross Section (CCS) Database: An Additional Measure to Characterize Steroids. *Anal. Chem.* **2018**, *90* (7), 4616–4625. https://doi.org/10.1021/acs.analchem.7b05117.

(25) Picache, J.; McLean, J. S50 | CCSCOMPEND | The Unified Collision Cross Section (CCS) Compendium (Version NORMAN-SLE-S50.0.1.0) [Data Set]. *Zenodo* **2019**. https://doi.org/10.5281/zenodo.2658163.

(26) Picache, J. A.; Rose, B. S.; Balinski, A.; Leaptrot, K. L.; Sherrod, S. D.; May, J. C.; McLean, J. A.

592 Collision Cross Section Compendium to Annotate and Predict Multi-Omic Compound Identities.
593 *Chem. Sci.* **2019**, *10* (4), 983–993. https://doi.org/10.1039/c8sc04396e.

594 (27) Metz, T. O.; Baker, E. S.; Schymanski, E. L.; Renslow, R. S.; Thomas, D. G.; Causon, T. K.; Webb,
595 I. K.; Hann, S.; Smith, R. D.; Teeguarden, J. G. Integrating Ion Mobility Spectrometry into Mass
596 Spectrometry-Based Exposome Measurements: What Can It Add and How Far Can It Go?
597 *Bioanalysis* **2017**, *9* (1), 81–98. https://doi.org/10.4155/bio-2016-0244.

598 (28) Zheng, X.; Aly, N. A.; Zhou, Y.; Dupuis, K. T.; Bilbao, A.; Paurus, V. L.; Orton, D. J.; Wilson, R.;
599 Payne, S. H.; Smith, R. D.; Baker, E. S. A Structural Examination and Collision Cross Section
600 Database for over 500 Metabolites and Xenobiotics Using Drift Tube Ion Mobility
601 Spectrometry. *Chem. Sci.* **2017**, *8* (11), 7724–7736. https://doi.org/10.1039/c7sc03464d.

602 (29) Colby, S. M.; Thomas, D. G.; Nunez, J. R.; Baxter, D. J.; Glaesemann, K. R.; Brown, J. M.; Pirrung,
603 M. A.; Govind, N.; Teeguarden, J. G.; Metz, T. O.; Renslow, R. S. ISiCLE: A Quantum Chemistry
604 Pipeline for Establishing in Silico Collision Cross Section Libraries. *Anal. Chem.* **2019**, *91* (7),
605 4346–4356. https://doi.org/10.1021/acs.analchem.8b04567.

606 (30) Nuñez, J. R.; Colby, S. M.; Thomas, D. G.; Tfaily, M. M.; Tolic, N.; Ulrich, E. M.; Sobus, J. R.; Metz,
607 T. O.; Teeguarden, J. G.; Renslow, R. S. Evaluation of In Silico Multifeature Libraries for Providing
608 Evidence for the Presence of Small Molecules in Synthetic Blinded Samples. *J. Chem. Inf. Model.*
609 **2019**, *59* (9), 4052–4060. https://doi.org/10.1021/acs.jcim.9b00444.

610 (31) Monge, M. E.; Dodds, J. N.; Baker, E. S.; Edison, A. S.; Fernández, F. M. Challenges in Identifying
611 the Dark Molecules of Life. *Annu. Rev. Anal. Chem.* **2019**, *12* (1), 177–199.
612 https://doi.org/10.1146/annurev-anchem-061318-114959.

613 (32) Lian, R.; Zhang, F.; Zhang, Y.; Wu, Z.; Ye, H.; Ni, C.; Lv, X.; Guo, Y. Ion Mobility Derived Collision
614 Cross Section as an Additional Measure to Support the Rapid Analysis of Abused Drugs and
615 Toxic Compounds Using Electrospray Ion Mobility Time-of-Flight Mass Spectrometry. *Anal.*

Methods **2018**, *10* (7), 749–756. https://doi.org/10.1039/C7AY02808C.

(33) Goscinny, S.; McCullagh, M.; Far, J.; De Pauw, E.; Eppe, G. Towards the Use of Ion Mobility Mass Spectrometry-Derived Collision Cross Section as a Screening Approach for Unambiguous Identification of Targeted Pesticides in Food. *Rapid Commun. Mass Spectrom.* **2019**, No. December 2018, 1–15. https://doi.org/10.1002/rcm.8395.

(34) Goscinny, S.; Joly, L.; De Pauw, E.; Hanot, V.; Eppe, G. Travelling-Wave Ion Mobility Time-of-Flight Mass Spectrometry as an Alternative Strategy for Screening of Multi-Class Pesticides in Fruits and Vegetables. *J. Chromatogr. A* **2015**, *1405*, 85–93. https://doi.org/10.1016/j.chroma.2015.05.057.

(35) Bauer, A.; Kuballa, J.; Rohn, S.; Jantzen, E.; Luetjohann, J. Evaluation and Validation of an Ion Mobility Quadrupole Time-of-Flight Mass Spectrometry Pesticide Screening Approach. *J. Sep. Sci.* **2018**, *41* (10), 2178–2187. https://doi.org/10.1002/jssc.201701059.

(36) Tejada-Casado, C.; Hernández-Mesa, M.; Monteau, F.; Lara, F. J.; Olmo-Iruela, M. del; García-Campaña, A. M.; Le Bizec, B.; Dervilly-Pinel, G. Collision Cross Section (CCS) as a Complementary Parameter to Characterize Human and Veterinary Drugs. *Anal. Chim. Acta* **2018**, *1043*, 52–63. https://doi.org/10.1016/j.aca.2018.09.065.

(37) NORMAN network. NORMAN Suspect List Exchange https://www.norman-network.com/nds/SLE/ (accessed Jul 27, 2020).

(38) Celma, A.; Fabregat-Safont, D.; Ibáñez, M.; Bijlsma, L.; Hernández, F.; Sancho, J. V. S61 | UJICCSLIB | Collision Cross Section (CCS) Library from UJI (Version S61.0.1.0) [Data Set]. *Zenodo* **2019**. https://doi.org/10.5281/zenodo.3549477.

(39) National Center for Biotechnology Information; National Library of Medicine; National Institutes of Health. Annotations from NORMAN Suspect List Exchange https://pubchem.ncbi.nlm.nih.gov/source/23819#data=Annotations (accessed Jul 27, 2020).

(40) ChemAxon. JChem for Office (V19.9.0.467). Budapest (Hungary) 2019.

(41) National Center for Biotechnology Information; National Library of Medicine; National Institutes of Health. Carbamazepine. PubChem Compound. https://pubchem.ncbi.nlm.nih.gov/compound/2554#section=Information-Sources (accessed Jul 27, 2020).

(42) Schymanski, E. L.; Singer, H. P.; Slobodnik, J.; Ipolyi, I. M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; Thomaidis, N. S.; Bletsou, A.; Zwiener, C.; Ibáñez, M.; Portolés, T.; De Boer, R.; Reid, M. J.; Onghena, M.; Kunkel, U.; Schulz, W.; Guillon, A.; Noyon, N.; Leroy, G.; Bados, P.; Bogialli, S.; Stipaničev, D.; Rostkowski, P.; Hollender, J. Non-Target Screening with High-Resolution Mass Spectrometry: Critical Review Using a Collaborative Trial on Water Analysis. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6237–6255. https://doi.org/10.1007/s00216-015-8681-7.

(43) Chiaia-Hernandez, A. C.; Keller, A.; Wächter, D.; Steinlin, C.; Camenzuli, L.; Hollender, J.; Krauss, M. Long-Term Persistence of Pesticides and TPs in Archived Agricultural Soil Samples and Comparison with Pesticide Application. *Environ. Sci. Technol.* **2017**, *51* (18), 10642–10651. https://doi.org/10.1021/acs.est.7b02529.

(44) Alygizakis, N. A.; Samanipour, S.; Hollender, J.; Ibáñez, M.; Kaserzon, S.; Kokkali, V.; van Leerdam, J. A.; Mueller, J. F.; Pijnappels, M.; Reid, M. J.; Schymanski, E. L.; Slobodnik, J.; Thomaidis, N. S.; Thomas, K. V. Exploring the Potential of a Global Emerging Contaminant Early Warning Network through the Use of Retrospective Suspect Screening with High-Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2018**, *52* (9), 5135–5144. https://doi.org/10.1021/acs.est.8b00365.

(45) European Commission. Directorate General for Health and Food Safety. *Guidance Document on Analytical Quality Control and Method Validation Procedures for Pesticide Residue and*

Analysis in Food and Feed. SANTE/11813/2017; 2017; p 42.

(46) Bijlsma, L.; Berntssen, M. H. G.; Merel, S. A Refined Nontarget Workflow for the Investigation of Metabolites through the Prioritization by in Silico Prediction Tools. *Anal. Chem.* **2019**, *91* (9), 6321–6328. https://doi.org/10.1021/acs.analchem.9b01218.

(47) Bade, R.; Bijlsma, L.; Miller, T. H.; Barron, L. P.; Sancho, J. V.; Hernández, F. Suspect Screening of Large Numbers of Emerging Contaminants in Environmental Waters Using Artificial Neural Networks for Chromatographic Retention Time Prediction and High Resolution Mass Spectrometry Data Analysis. *Sci. Total Environ.* **2015**, *538*, 934–941. https://doi.org/10.1016/j.scitotenv.2015.08.078.

(48) Bade, R.; Bijlsma, L.; Sancho, J. V.; Hernández, F. Critical Evaluation of a Simple Retention Time Predictor Based on LogKow as a Complementary Tool in the Identification of Emerging Contaminants in Water. *Talanta* **2015**, *139*, 143–149. https://doi.org/10.1016/j.talanta.2015.02.055.

(49) Bijlsma, L.; Bade, R.; Celma, A.; Mullin, L.; Cleland, G.; Stead, S.; Hernandez, F.; Sancho, J. V. J. V. Prediction of Collision Cross-Section Values for Small Molecules: Application to Pesticide Residue Analysis. *Anal. Chem.* **2017**, *89* (12), 6583–6589. https://doi.org/10.1021/acs.analchem.7b00741.

(50) Kaufmann, A.; Butcher, P.; Maden, K.; Walker, S.; Widmer, M. Practical Application of in Silico Fragmentation Based Residue Screening with Ion Mobility High-Resolution Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (13), 1147–1157. https://doi.org/10.1002/rcm.7890.

(51) Mollerup, C. B.; Mardal, M.; Dalsgaard, P. W.; Linnet, K.; Barron, L. P. Prediction of Collision Cross Section and Retention Time for Broad Scope Screening in Gradient Reversed-Phase Liquid Chromatography-Ion Mobility-High Resolution Accurate Mass Spectrometry. *J. Chromatogr. A* **2018**, *1542*, 82–88. https://doi.org/10.1016/j.chroma.2018.02.025.

(52) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. MetFrag Relaunched: Incorporating Strategies beyond in Silico Fragmentation. *J. Cheminform.* **2016**, *8* (1), 1–16. https://doi.org/10.1186/s13321-016-0115-9.

(53) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. In Silico Fragmentation for Computer Assisted Identification of Metabolite Mass Spectra. *BMC Bioinformatics* **2010**, *11*. https://doi.org/10.1186/1471-2105-11-148.

(54) Zhou, Z.; Shen, X.; Tu, J.; Zhu, Z. J. Large-Scale Prediction of Collision Cross-Section Values for Metabolites in Ion Mobility-Mass Spectrometry. *Anal. Chem.* **2016**, *88* (22), 11084–11091. https://doi.org/10.1021/acs.analchem.6b03091.

(55) Aalizadeh, R.; Nika, M.-C.; Thomaidis, N. S. Development and Application of Retention Time Prediction Models in the Suspect and Non-Target Screening of Emerging Contaminants. *J. Hazard. Mater.* **2019**, *363* (September 2018), 277–285. https://doi.org/10.1016/j.jhazmat.2018.09.047.

(56) European Commission. *Commission Decision of 12 August 2002 Implementing Council Directive 96/23/EC Concerning the Performance of Analytical Methods and the Interpretation of Results (2002/657/EC)*; 2002; Vol. L221, pp 8–36.

(57) Gago-Ferrero, P.; Bletsou, A. A.; Damalas, D. E.; Aalizadeh, R.; Alygizakis, N. A.; Singer, H. P.; Hollender, J.; Thomaidis, N. S. Wide-Scope Target Screening of >2000 Emerging Contaminants in Wastewater Samples with UPLC-Q-ToF-HRMS/MS and Smart Evaluation of Its Performance through the Validation of 195 Selected Representative Analytes. *J. Hazard. Mater.* **2019**, *387*, 121712. https://doi.org/10.1016/j.jhazmat.2019.121712.

(58) Blom, K. F. Estimating the Precision of Exact Mass Measurements on an Orthogonal Time-of-Flight Mass Spectrometer. *Anal. Chem.* **2001**, *73* (3), 715–719. https://doi.org/10.1021/ac001064v.

712 (59) Vergeynst, L.; Van Langenhove, H.; Joos, P.; Demeestere, K. Suspect Screening and Target
713 Quantification of Multi-Class Pharmaceuticals in Surface Water Based on Large-Volume
714 Injection Liquid Chromatography and Time-of-Flight Mass Spectrometry. *Anal. Bioanal. Chem.*
715 **2014**, *406* (11), 2533–2547. https://doi.org/10.1007/s00216-014-7672-4.

716 (60) Bader, T.; Schulz, W.; Kümmerer, K.; Winzenbacher, R. General Strategies to Increase the
717 Repeatability in Non-Target Screening by Liquid Chromatography-High Resolution Mass
718 Spectrometry. *Anal. Chim. Acta* **2016**, *935*, 173–186.
719 https://doi.org/10.1016/j.aca.2016.06.030.

720 (61) Delatour, T.; Savoy, M. C.; Tarres, A.; Bessaire, T.; Mottier, P.; Desmarchelier, A. Low False
721 Response Rates in Screening a Hundred Veterinary Drug Residues in Foodstuffs by LC-MS/MS
722 with Analyte-Specific Correction of the Matrix Effect. *Food Control* **2018**, *94* (July), 353–360.
723 https://doi.org/10.1016/j.foodcont.2018.07.014.

724 (62) Vergeynst, L.; Van Langenhove, H.; Demeestere, K. Balancing the False Negative and Positive
725 Rates in Suspect Screening with High-Resolution Orbitrap Mass Spectrometry Using
726 Multivariate Statistics. *Anal. Chem.* **2015**, *87* (4), 2170–2177.
727 https://doi.org/10.1021/ac503426k.

728 (63) Schollée, J. E.; Schymanski, E. L.; Stravs, M. A.; Gulde, R.; Thomaidis, N. S.; Hollender, J.
729 Similarity of High-Resolution Tandem Mass Spectrometry Spectra of Structurally Related
730 Micropollutants and Transformation Products. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (12), 2692–
731 2704. https://doi.org/10.1007/s13361-017-1797-6.

732 (64) MassBank of North America. MassBank of North America (MoNA) - Tricyclazole
733 https://mona.fiehnlab.ucdavis.edu/spectra/display/CCMSLIB00001058767 (accessed Oct 18,
734 2019).

735 (65) Bijlsma, L.; Boix, C.; Niessen, W. M. A.; Ibáñez, M.; Sancho, J. V.; Hernández, F. Investigation of

736 Degradation Products of Cocaine and Benzoylecgonine in the Aquatic Environment. *Sci. Total*

737 *Environ.* **2013**, *443*, 200–208. https://doi.org/10.1016/j.scitotenv.2012.11.006.

738

739

740

741 **Figure captions**

742

743 **Figure 1**. Different confidence levels established in the identification of a compound applying
744 ion mobility high resolution mass spectrometry target, suspect and non-target screening workflows
745 based on the levels provided by Schymanski et al. [13]. *MS* refers to accurate mass of the precursor ion,
746 MS^n to accurate mass of the fragment ions, *RT* is the retention time, *RTI* refers to retention time
747 indexing systems, *CCS* means Collision Cross-Section, and the sub index *Pred.* indicates that the value
748 is in accordance with predictive models applied.

749

750 **Figure 2**. Identification at Level 1 of 4-acetamidoantipyrin in a surface water sample. *(a)*
751 Structure, RT and CCS comparison of experimental and standard data, *(b)* Extracted ion
752 chromatograms for $[M+H]^+$ ion (*m/z* 246.1240) and two characteristic fragments (*m/z* 228.1132 and
753 *m/z* 104.0495) and *(c)* Drift time aligned MS data along with the empirical mass error of the
754 corresponding fragment ions observed.

755

756 **Figure 3.** Identification as level 1* of the fungicide thiabendazole in a Spanish River mouth
757 including structure and CCS comparison of experimental and expected data (right top panel),
758 extracted ion chromatograms for $[M+H]^+$ ion (*m/z* 202.0433) and 3 representative fragments (*m/z*
759 175.0326, *m/z* 131.0604, *m/z* 92.0495) (left panel) and drift time aligned MS data with the empirical
760 mass error of the fragment ions observed (right-bottom panel).

761

762 **Figure 4.** Identification at Level 2a of tricyclazole in a surface water. *(a)* Structure and CCS
763 comparison of experimental and predicted data, *(b)* Extracted ion chromatogram for $[M+H]^+$ ion (*m/z*
764 190.0435) of tricyclazole and *(c)* Drift time aligned MS data along with the empirical mass error of the
765 fragment ions observed.

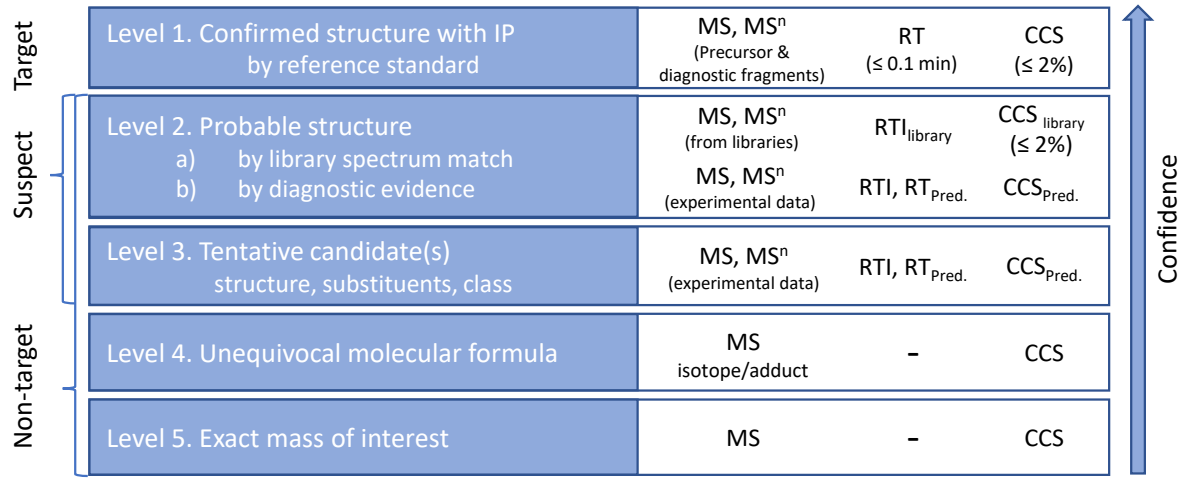
766

767 **Figure 5.** Comparison of HRMS spectra for benzoylecgonine in analytical reference standard
768 solution (a), drift time aligned data of positive finding in wastewater sample (b) and non-drift time
769 aligned data of the same positive finding in wastewater (c).

770

771 **Figure 6.** 'Pre-mobility' fragmentation of 2C-E resulting in LE fragments with different drift
772 time (blue-shadowed points) *(c)* which are omitted in the drift-time aligned data for protonated
773 adduct (green-shadowed peak) *(a)* but present in the non-drift time aligned data (blue-shadowed
774 peaks) *(b).*

775



776

777 **Figure 1.**

778

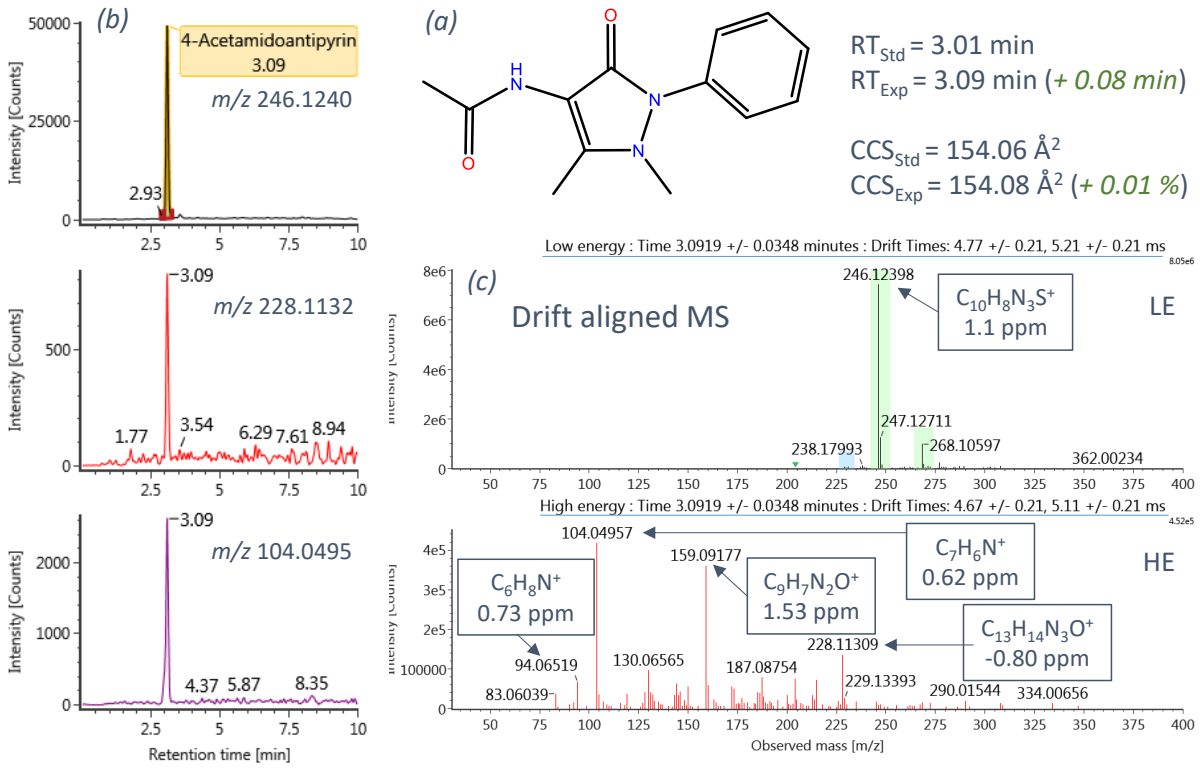
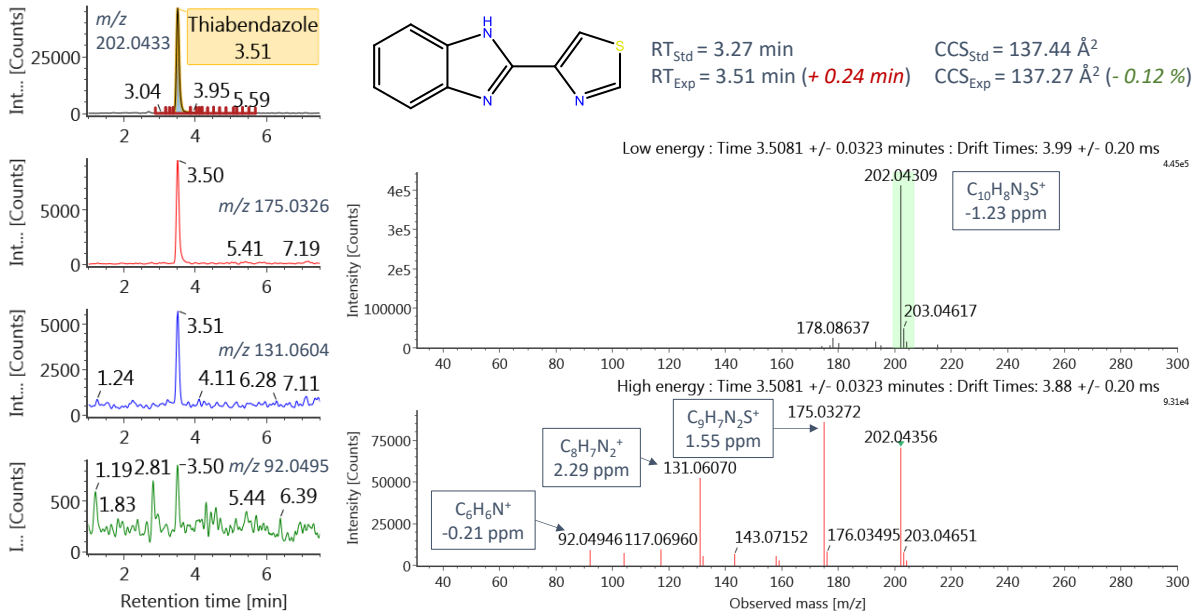

(b)
4-Acetamidoantipyrin
3.09
m/z 246.1240
2.93
m/z 228.1132
3.09
1.77
3.54
6.29
7.61
8.94
m/z 104.0495
3.09
4.37
5.87
8.35
Intensity [Counts]
Retention time [min]
(a)
RT_Std = 3.01 min
RT_Exp = 3.09 min (+ 0.08 min)
CCS_Std = 154.06 Å²
CCS_Exp = 154.08 Å² (+ 0.01 %)
(c)
Drift aligned MS
Low energy : Time 3.0919 +/- 0.0348 minutes : Drift Times: 4.77 +/- 0.21, 5.21 +/- 0.21 ms
246.12398
C10H8N3S+
1.1 ppm
LE
247.12711
238.17993
268.10597
362.00234
High energy : Time 3.0919 +/- 0.0348 minutes : Drift Times: 4.67 +/- 0.21, 5.11 +/- 0.21 ms
104.04957
C7H6N+
0.62 ppm
HE
159.09177
C9H7N2O+
1.53 ppm
C6H8N+
0.73 ppm
228.11309
C13H14N3O+
-0.80 ppm
94.06519
130.06565
187.08754
229.13393
83.06039
290.01544
334.00656
Observed mass [m/z]

Figure 2.

783

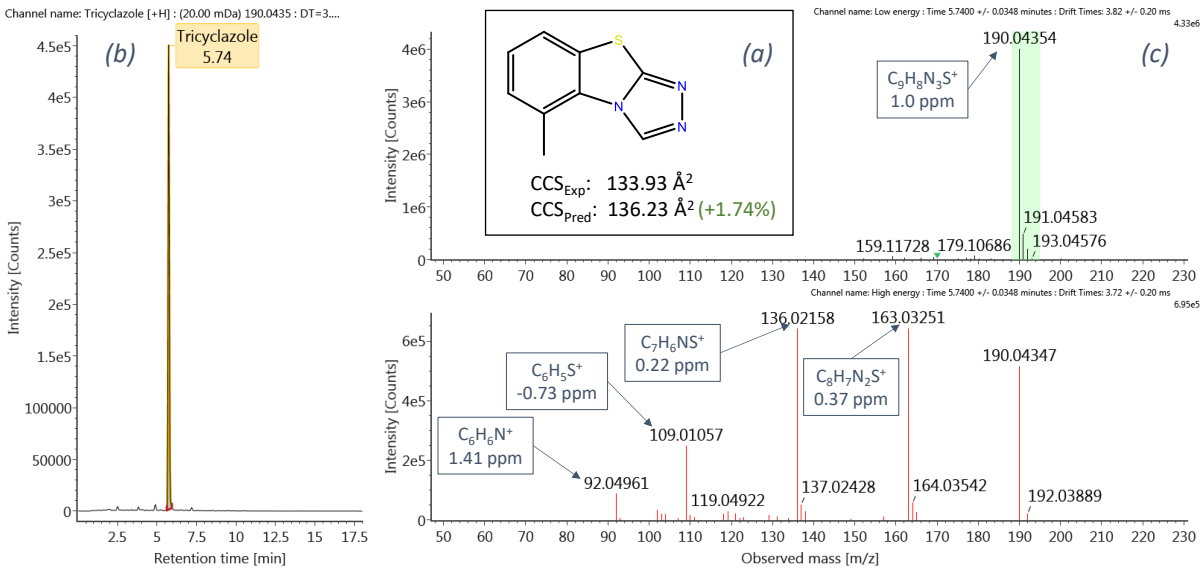


784

785 **Figure 3.**

786

787



788

789 **Figure 4.**

790

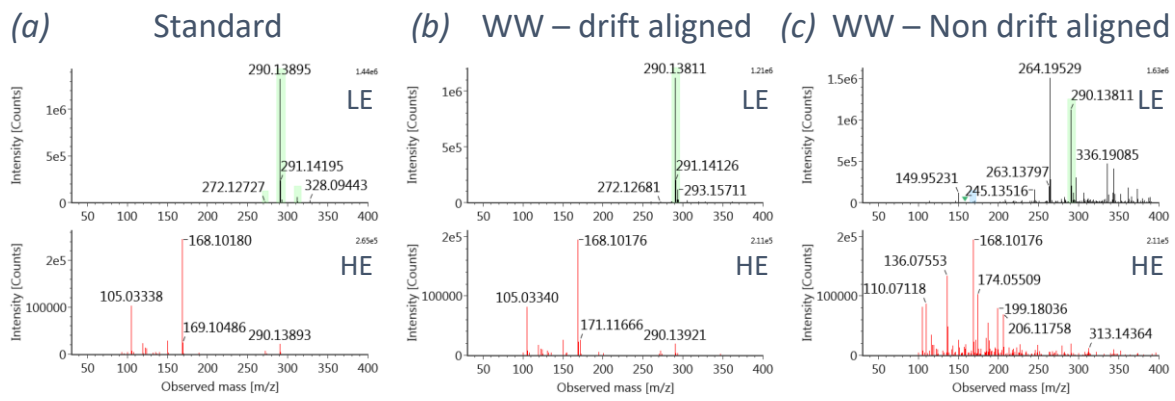

(a) Standard
(b) WW – drift aligned
(c) WW – Non drift aligned
LE
HE
Intensity [Counts]
Observed mass [m/z]
290.13895
291.14195
272.12727
328.09443
168.10180
105.03338
169.10486
290.13893
290.13811
291.14126
272.12681
293.15711
168.10176
105.03340
171.11666
290.13921
264.19529
290.13811
336.19085
149.95231
263.13797
245.13516
168.10176
136.07553
174.05509
110.07118
199.18036
206.11758
313.14364

Figure 5.

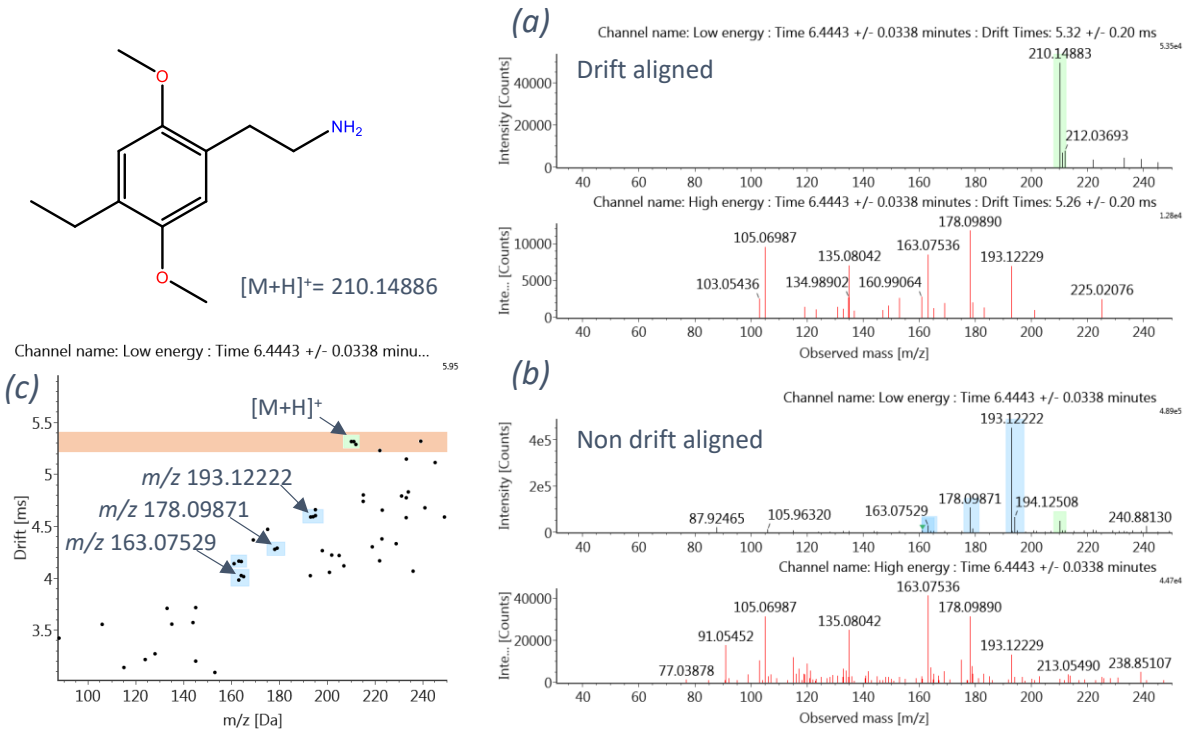

(a)
Channel name: Low energy : Time 6.4443 +/- 0.0338 minutes : Drift Times: 5.32 +/- 0.20 ms
Drift aligned
Intensity [Counts]
210.14883
212.03693
Channel name: High energy : Time 6.4443 +/- 0.0338 minutes : Drift Times: 5.26 +/- 0.20 ms
178.09890
105.06987
163.07536
135.08042
193.12229
103.05436
134.98902
160.99064
225.02076
Observed mass [m/z]
(b)
Channel name: Low energy : Time 6.4443 +/- 0.0338 minutes
Non drift aligned
193.12222
87.92465
105.96320
163.07529
178.09871
194.12508
240.88130
Channel name: High energy : Time 6.4443 +/- 0.0338 minutes
163.07536
105.06987
178.09890
135.08042
91.05452
193.12229
77.03878
213.05490
238.85107
Observed mass [m/z]
[M+H]+= 210.14886
Channel name: Low energy : Time 6.4443 +/- 0.0338 minu...
(c)
[M+H]+
m/z 193.12222
m/z 178.09871
m/z 163.07529
Drift [ms]
m/z [Da]

Figure 6.