Article

Reference library for suspect screening of environmental toxicants using ion mobility spectrometry-mass spectrometry

Check for updates

Devin Teri[1], Noor A. Aly[1], James N. Dodds [2], Jian Zhang[3], Paul A. Thiessen [3], Evan E. Bolton[3], Kara M. Joseph[2], Antony J. Williams [4], Emma L. Schymanski [5], Ivan Rusyn [1] ✉ & Erin S. Baker [2]✉

Due to the potential health risks related to chemical exposure, rapidly assessing xenobiotic molecules in the environment and those already in the body is imperative. Targeted analytical methods coupling either gas or liquid chromatography with mass spectrometry (GC-MS or LC-MS) are commonly utilized in current exposure assessments. While these methods are accepted as the gold standard for exposure analyses, they often require multiple sample preparation steps and analysis times > 30 min. These limitations have resulted in an evolving interest in using ion mobility spectrometry and MS (IMS-MS), either with or without chromatography, to improve throughput and annotation confidence. To increase IMS-MS information availability for exposure studies, here we utilized drift tube IMS-MS to evaluate 4685 xenobiotic chemical standards from the Environmental Protection Agency Toxicity Forecaster (ToxCast) programme, including pesticides, industrial chemicals, pharmaceuticals, consumer products, and per- and polyfluoroalkyl substances. Collision cross section (CCS) and m/z values were detected for 2144 unique chemicals with high confidence and reproducibility ($\leq$1% error intra-laboratory and $\leq$2% inter-laboratory), resulting in values for 4004 $[M + H]^+$, $[M+Na]^+$, $[M-H]^-$ and $[M]^{\bullet+}$ ion types. This multidimensional database therefore supports suspect screening for a wider range of environmental contaminants, faster exposure response times, and assessments of xenobiotic-disease connections.

Ion mobility spectrometry coupled with mass spectrometry (IMS-MS) is an analytical technique that has stimulated great interest in both analytical chemistry and exposure science due to its ability to resolve isomeric species, distinguish halogenated chemicals and perform a variety of studies including targeted, suspect screening, and non-targeted analyses[1–7]. IMS separates ions by their shape and charge state in the gas phase through the use of an electric field, buffer gas and sometimes gas flow, depending on the IMS type utilized (e.g., drift tube, travelling wave, trapped IMS, etc.)[8,9]. In addition to its millisecond analysis speed, most IMS methods also allow for the calculation of collision cross section (CCS) values for each ion detected[10,11]. CCS values correspond to the collision area between a charged analyte and its neutral collision partner or the buffer gas molecules in the IMS cell. In various studies, CCS values have proven to be highly reproducible (often less than 2% of each other) between laboratories using the same instrumentation, even in complex samples[12]. Additionally, CCS values for the same molecule evaluated with the different IMS methods are often within 2% of each other[8]. Therefore, CCS values are regarded as highly informative data which provide a way to increase confidence in the identification of individual molecules in non-targeted analytical studies of complex substances and mixtures[13,14]. Moreover, in IMS-MS studies, ions separated by IMS can be evaluated for their m/z value, and if fragmentation is applied, 3-dimensional characterization is possible (CCS, MS and MS/MS)[15,16].

Historically, IMS-MS has been used to study peptides, proteins, carbohydrates and inorganic molecules[17]. More recently, studies of environmental and industrial chemicals have provided additional CCS information to enable analyses of complex environmental exposures and separate halogenated chemicals from those containing mainly hydrogens to lessen sample preparation needs[6,8,18–22]. These studies have demonstrated the use of CCS libraries in suspect screening of complex substances[5,23–26] and in

[1]Department of Veterinary Physiology and Pharmacology, Texas A&M University, College Station, TX, USA. [2]Department of Chemistry, University of North Carolina, Chapel Hill, NC, USA. [3]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA. [4]Center for Computational Toxicology and Exposure, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA. [5]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, Luxembourg. ✉e-mail: irusyn@tamu.edu; erinmsb@unc.edu

assessing environmental samples[18,27]. The opportunities for further advancing chemical exposure analyses and determining potential adverse effects on human health are therefore rapidly evolving with analytical chemistry advances such as IMS-MS[28–31]. However, improved linkages between exposure measurements and potential hazards of chemicals, such as high-throughput toxicity data[32], are still lacking due to the absence of analytical methods that can be used for rapid analyses of complex samples and mixtures[33–35]. Additionally, refinements to characterization of in vitro to in vivo extrapolations (IVIVE) of individual chemicals[36,37] and mixtures[38] would therefore benefit from improved analytical methods able to assess diverse chemicals in the environment and their metabolites in biological systems. Fortunately, government agencies have encouraged collaborative analyses of large-scale chemical libraries and mixtures by providing access to their compound libraries[39], enabling opportunities for developing large-scale reference datasets for suspect screening using rapid analytical methods such as IMS-MS.

In this study, we follow up on recent studies that have aimed to establish CCS libraries for xenobiotic chemicals[8,18,21,22] by utilizing drift tube IMS (DTIMS) and nitrogen buffer gas (denoted as $^{DT}CCS_{N2}$) to evaluate chemicals of concern to human health which have also been tested in hundreds of in vitro assays by the ToxCast/Tox21 programmes[32,40]. For these analyses, we utilized an electrospray ionization (ESI) source in both positive and negative mode and an atmospheric pressure chemical ionization (APCI) source in positive mode for more comprehensive molecular coverage of the diverse chemicals. The database created from this study contains CCS values for 2144 unique chemicals and will be of immediate interest for the chemical analyses of complex samples and environmental mixtures[41].

Methods
Sample preparation

The 4685 xenobiotic chemical standards analysed in this study (Fig. 1 and Tables S1–S2 within the Supplementary Data file) were obtained from Evotec (Branford, CT; order #13169), a contractor of the U.S. Environmental Protection Agency's (US EPA) ToxCast library of chemicals[42]. Chemical selection from a ToxCast library was performed by the U.S. EPA (Dr. Ann M. Richard) based on chemical availability and the purity of the individual compounds at the time of assembly and all standards were supplied at 0.4 mM in pure dimethyl sulfoxide (DMSO). The process for selecting and registering chemicals into the ToxCast library including the description of chemical procurement from commercial suppliers, cataloguing the chemicals, storage and chemical quality control is detailed at refs. 43,44. After selection, all chemicals were placed into sealed 384-well plates, shipped, and then stored frozen (−80 °C). The certificate of analysis obtained with the chemical shipment stated all compounds were >90% purity as determined by the supplier. Before experiments, the stock plates were thawed at room temperature, 5 μL of each chemical was then transferred to individual 9 mm amber screw top vials with fused inserts (Ibis Scientific, #4400-FIV2W), and each standard was diluted to 10 μM with 195 μL of 50:50 methanol/water. The high purity (≥99.9%) water, methanol, and acetonitrile used in this study were obtained from Sigma-Aldrich (St. Louis, MO).
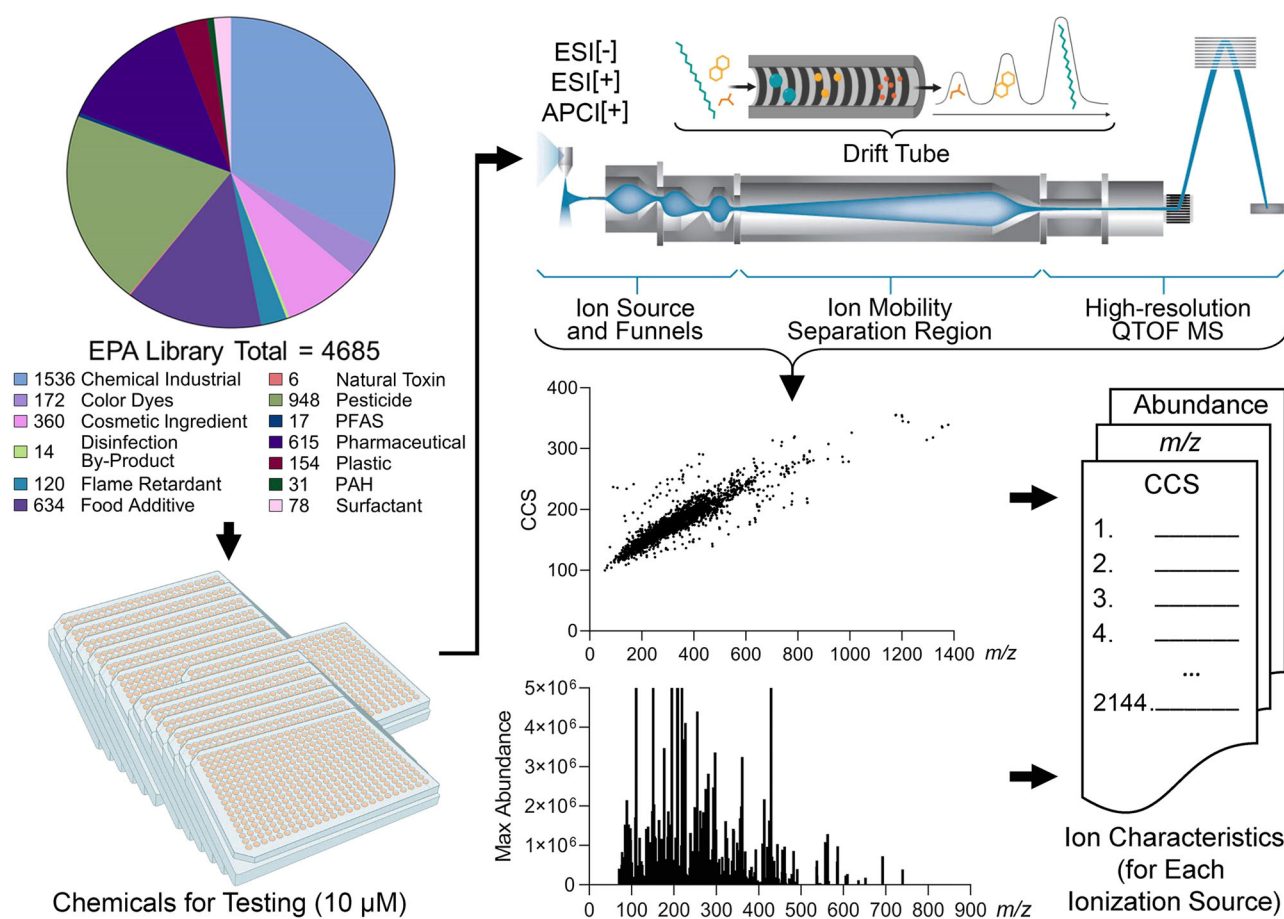
Fig. 1 | Project workflow. In this study, we first classified the 4685 chemical standards into 13 classes. The 4685 standards were then diluted to 10 μM and analysed with IMS-MS using ESI[+], ESI[−] and APCI[+]. Feature identification and collision cross section calculations were performed using Agilent IM-MS Browser and Agilent Mass Profiler. CCS values for 2144 unique chemicals were obtained in the study and 4004 ions from [M + H]⁺, [M+Na]⁺, [M-H]⁻ and [M]•⁺ ion types.
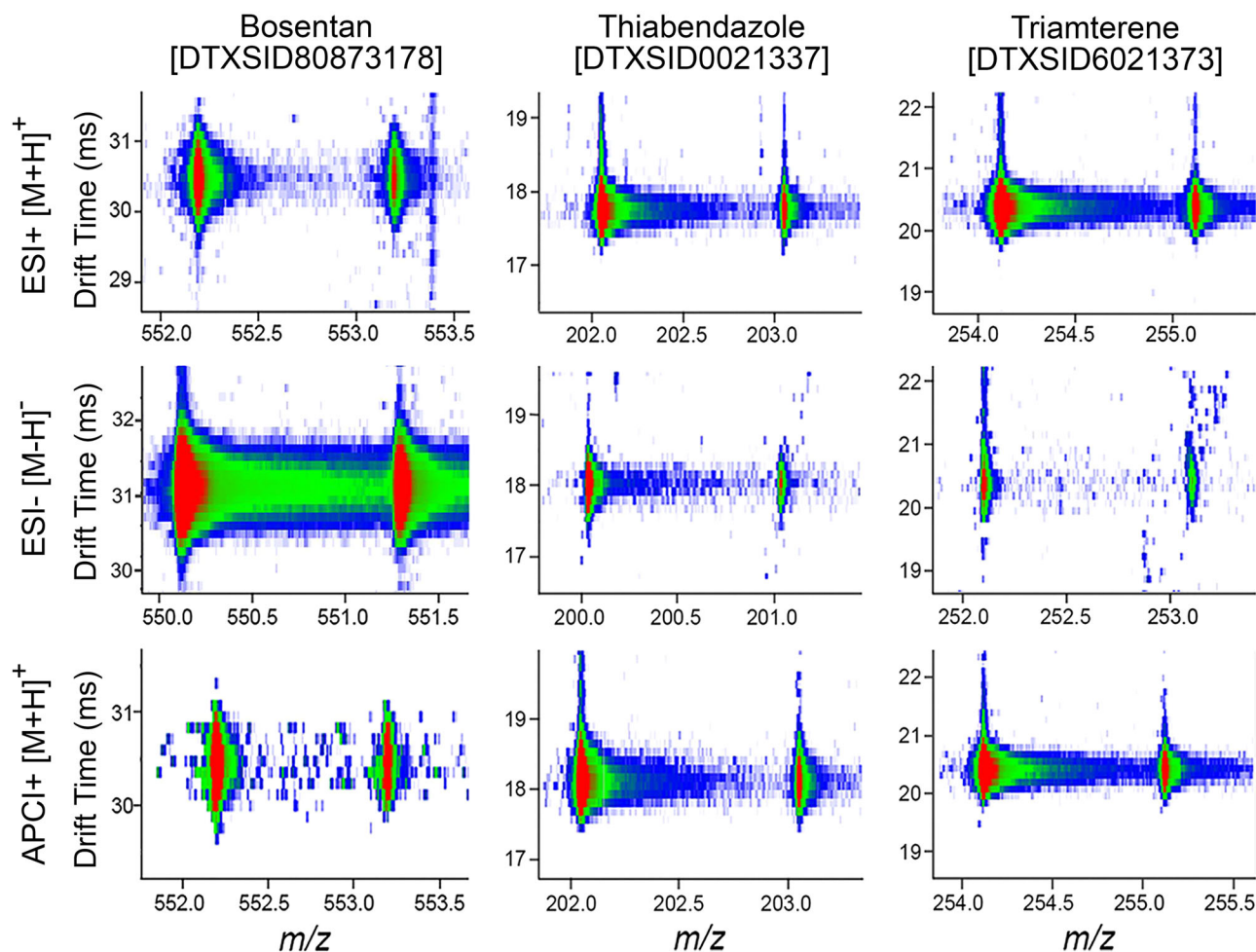
Fig. 2 | Nested IMS–MS spectra for three representative test chemicals. Images illustrate the abundance of the isotopic distribution of the precursor [M] and first ^{13}C isotope [M + 1] peak for each chemical across the three ionization modes (ESI[−], ESI[+] and APCI[+]). For each chemical shown (Bosentan, Thiabendazole, and Triamterene), *m/z* values are plotted on the x-axis, and drift times on the y-axis. Colours represent abundance from red (highest) to green (intermediate), to blue (lowest). The DTXSIDs are the substance identifiers associated with the US EPA's CompTox Chemicals Dashboard.

Instrumentation and analysis

DTIMS-MS analyses were performed using the 6560 IM-QTOF MS system (Agilent Technologies, Santa Clara), and the drift tube was filled with nitrogen gas. A liquid chromatography stack (Agilent Technologies: G7116B Column Compartment, G7167B Multisampler, and G7120A High Speed Pump) was used as an autosampler to inject samples into each ionization source at 0.400 mL/min with a constant solvent composition of 50% water and 50% acetonitrile. All chemical standards were ionized using the APCI source in positive ionization mode and the ESI source in both positive and negative ionization mode.

Prior to each experiment, the instrument was tuned and mass calibrated using the Agilent Tune Mix (G2421A/G2432A, Agilent Technologies, Santa Clara, CA). For the IMS analyses, the ions were passed through the inlet glass capillary, focused by a high-pressure ion funnel, and accumulated in an ion funnel trap[45]. Ions were then pulsed into the 78.24 cm long drift tube filled with ~3.95 torr of nitrogen gas, where they travelled under the influence of a weak electric field (17 V/cm). Ions exiting the drift tube were refocused by a rear ion funnel prior to quadrupole time-of-flight (QTOF) MS analysis and their drift time was recorded. The detailed instrumental settings are listed in **Supporting Information File**. For each detected ion, $^{DT}CCS_{N2}$ values were calculated using a single-field method (Tables S3–S5 **and uploaded to PubChem**), which has been detailed previously[12]. For quality control for the single-field method, Agilent Tune Mix ions were evaluated daily at the start of each experimental run and the calibration constants derived were stable for all studies performed in this

manuscript (see Results). The use of Agilent Tune Mix and the single-field method was the same process as carried out by Picache et al.[22], to obtain values that will be reproducible for those applying IMS CCS measurements. Methanol washes were performed after 24 sample injections to reduce the occurrence of carryover. Additionally, isomeric chemicals were also separated by several standards with greater mass differences to ensure no interferences occurred. Samples from the same plate were run in the same worklist over six days, alternating the ion mode each day. Each chemical was injected six times total, resulting in duplicate results for each ion mode. All ions were evaluated and only considered present if their mass error was less than 10 ppm from the expected chemical formula, they illustrated the correct ^{13}C isotopic distribution and the duplicates had <1% difference in CCS values. The average $^{DT}CCS_{N2}$ values for each ion are illustrated in Table S6.

Data analysis and curation

Agilent IM-MS Browser software (Agilent Technologies) was utilized for all single-field $^{DT}CCS_{N2}$ calculations, initially starting with the Tune Mix ions for use as calibrants. While software programmes can assist in calculating the CCS values, every instrumental run was manually examined for the presence of a peak and the software-assigned CCS value was verified. Quality control and confidence in CCS values were determined based on the $^{DT}CCS_{N2}$ duplicate values[46] having a percent difference of ≤1% for at least two replicates[47]. The $^{DT}CCS_{N2}$ and *m/z* values were then plotted to identify potential trendlines. Only $^{DT}CCS_{N2}$ values that fell within 15% difference of
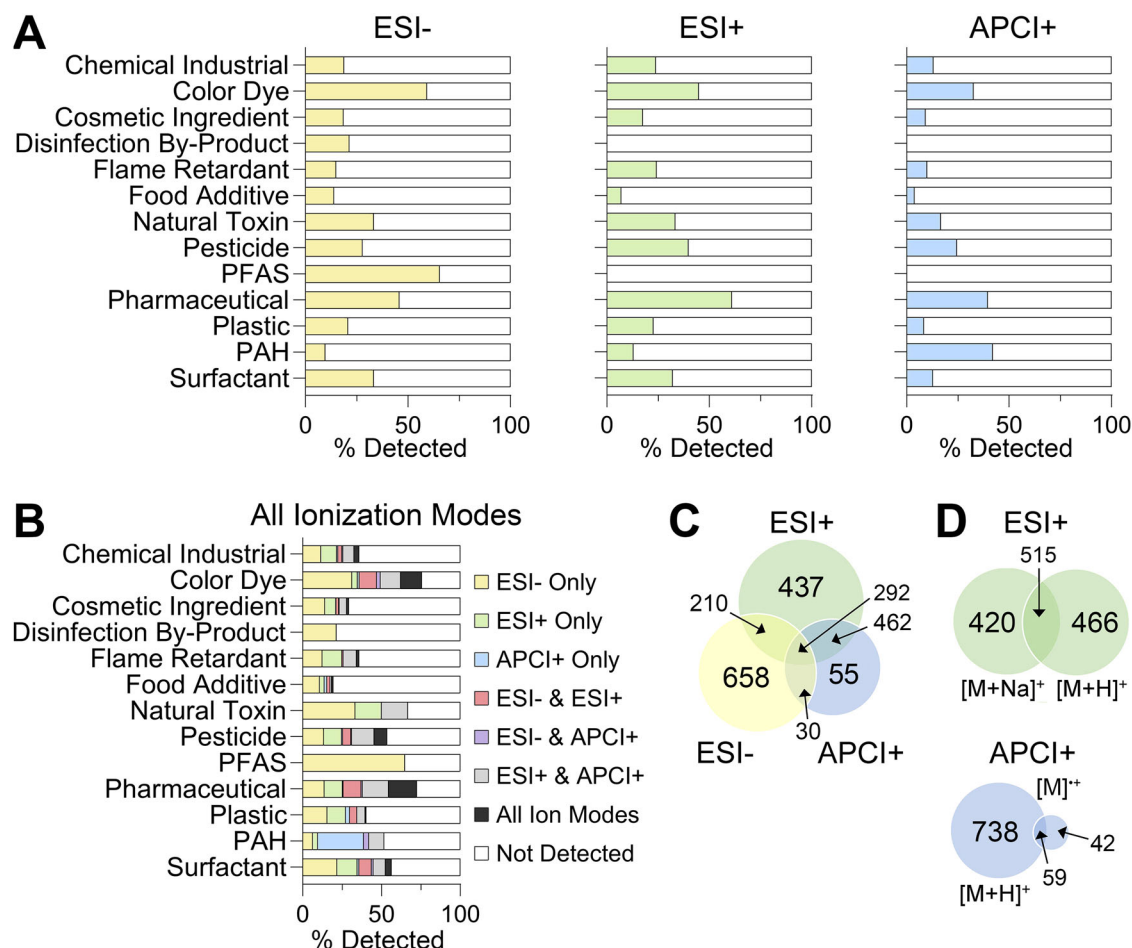
Fig. 3 | Detection frequency for tested chemicals across ionization modes and chemical classes. A Percentages of chemicals detected in each class (filled bars) using ESI[-], ESI[+], or APCI[+]. **B** Percentage of chemicals in each class detected by one or multiple ionization modes. **C** Venn diagram illustrating overlap of chemicals detected across the different ionization modes. **D** Venn diagrams showing number of ions detected in ESI[+] and APCI[+].

the main trendline were determined to be confident identities and listed in the final dataset. This filtering removes small molecules which multimerize in solution, drift through the ion mobility cell as multimers, and then break into monomers following drift analysis, thereby illustrating extremely large CCS values per m/z[48].

Of the 4685 standards evaluated herein, three chemicals were duplicated in the plates, resulting in 4682 unique chemicals. The chemicals were assigned into classes using expert judgement of the authors using PubChem "Use and manufacturing" classifications and other pertinent information in the US EPA's CompTox Chemicals Dashboard. Following ESI-IMS-MS and APCI-IMS-MS analyses, reproducible CCS values resulted for 2,144 unique chemicals. Moreover, due to the various ion types observed ([M + H]$^+$, [M+Na]$^+$, [M-H]$^-$ and [M]$^{\bullet+}$), the IMS-MS database contains 4,004 CCS values and the corresponding m/z ratios.

Results and Discussion
Many experimentally measured CCS datasets have been published for use in the fields of proteomics, metabolomics, steroids, and xenobiotics[8,18–22], but the number of CCS values is still limited, especially for xenobiotic studies. The purpose of developing a database of CCS values from the EPA's ToxCast library is therefore to assist in identifying toxicants and improve the assessment of human and environmental risks and exposures. To create the IMS-MS ToxCast database, 4685 chemical standards were analysed using the workflow depicted in Fig. 1. In this workflow, the chemicals were first categorized into classes, diluted, evaluated with IMS-MS, assessed for reproducible CCS values, and then added to a CCS database if they passed all

quality control steps stated in the Methods section. For the categorization step, the chemicals were placed into thirteen broad classes based on their structures and use by employing information from the EPA's ToxCast library[32]. Because most chemicals had several classifications listed, we used the source/use information in PubChem to determine the final assignments. Overall, the following assignments were made: Natural Toxin ($n = 6$), Disinfection By-Product ($n = 14$), PFAS ($n = 17$), Polycyclic Aromatic Hydrocarbon (PAH, $n = 31$), Surfactant ($n = 78$), Flame Retardant ($n = 120$), Plastic ($n = 154$), Colour Dye ($n = 172$), Cosmetic Ingredient ($n = 360$), Pharmaceutical ($n = 615$), Food Additive ($n = 634$), Pesticide ($n = 948$), and Chemical Industrial ($n = 1536$), where the three duplicates were from the Pharmaceutical, Food Additive, and Chemical Industrial classes. It should be noted that while these classifications are subject to interpretation, the assignment of substances into these classes is commonly used in regulatory science.

For the IMS-MS analyses, each chemical standard was diluted to 10 μM and injected twice into the instrument using ESI[+], ESI[-], and APCI[+] sources. While various ions were formed in the different sources, we specifically evaluated [M-H]$^-$ in the ESI[-] analyses, [M + H]$^+$ and [M+Na]$^+$ in the ESI[+] analyses, and [M + H]$^+$ and [M]$^{\bullet+}$ in the APCI[+] studies. Furthermore, an ion was considered present if the mass error was less than 10 ppm from the expected chemical formula, in addition to the presence of the correct ^{13}C isotopic distribution. Figure 2 shows three example chemicals following ionization using all three ion modes. These chemicals (bosentan, thiabendazole, and triamterene) represent two different chemical classes studied (two pharmaceuticals and a pesticide), each having ions observed in all analysis modes.
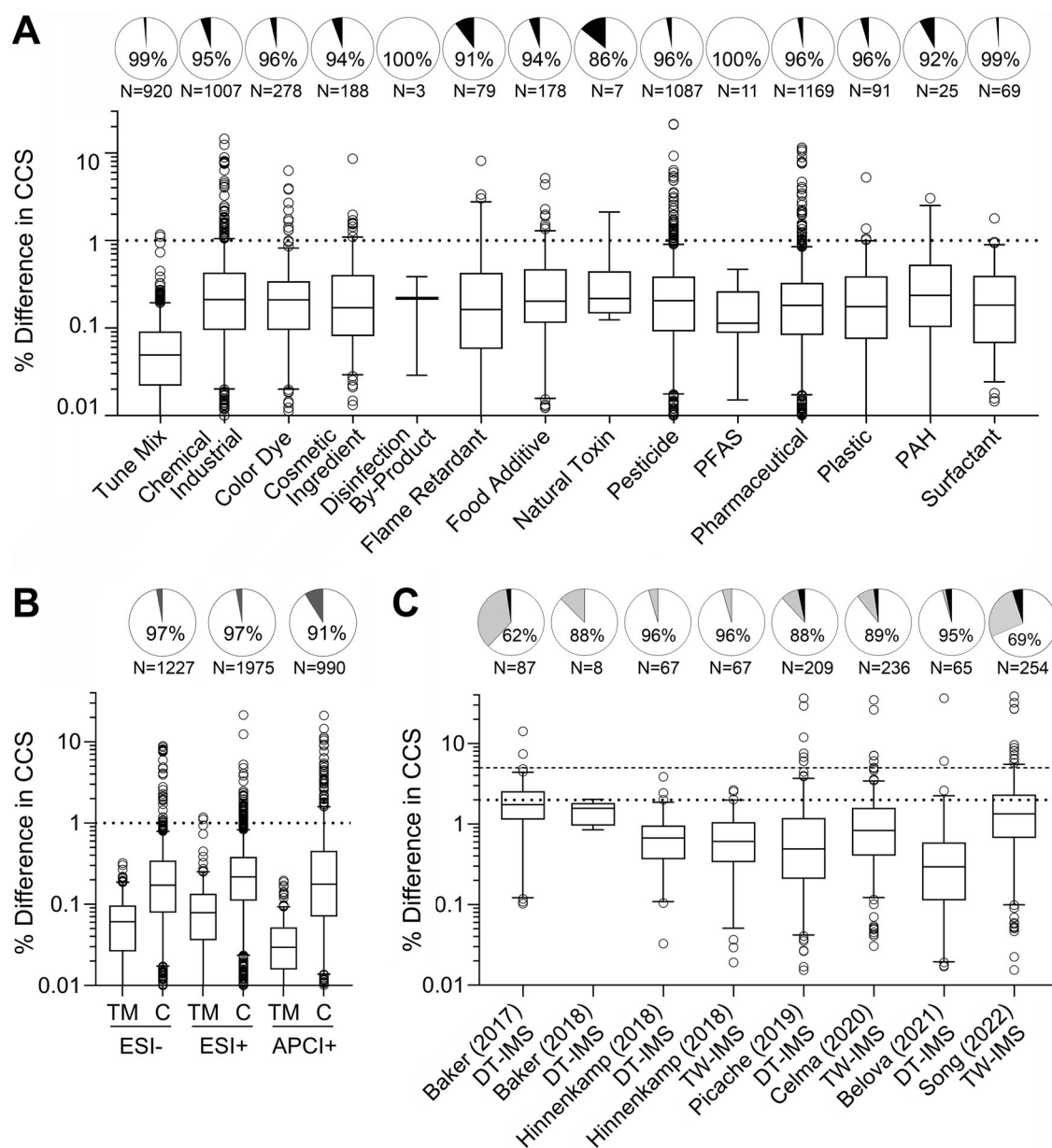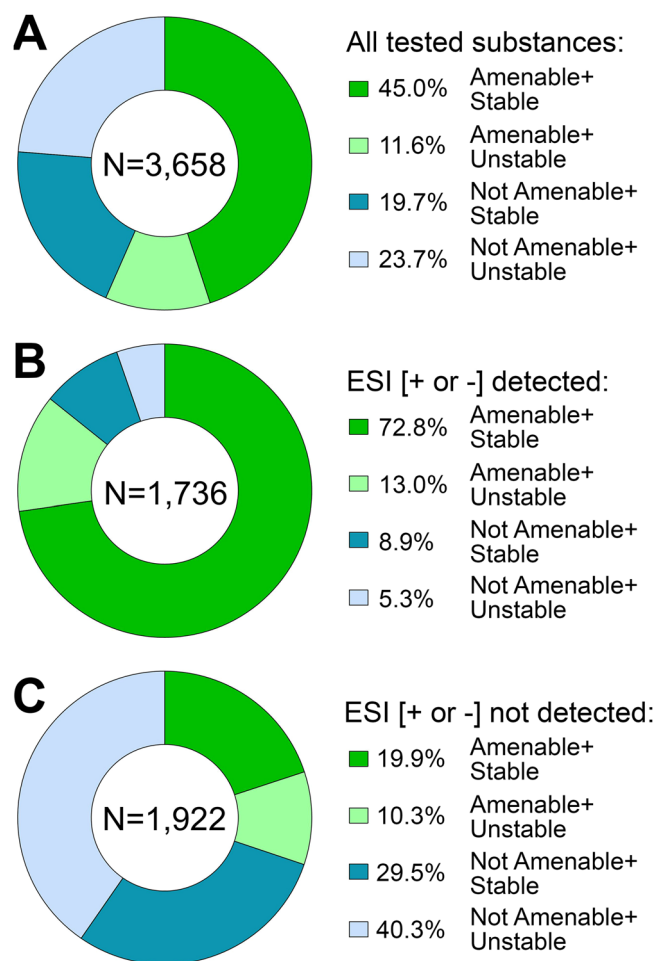
Fig. 4 | Reproducibility of Collision Cross Section (CCS) values. CCS reproducibility for the technical replicates examined in this study was assessed by **A** chemical class and **B** ionization mode (ESI[-], ESI[+], and APCI[+]). **C** CCS reproducibility between this study and other published large datasets that contained overlapping chemicals is shown in **C**. The pie charts in **A** and **B** show the fraction of the chemicals falling below a 1% threshold (white) and above 1% threshold (black), which were filtered out. The total number of chemicals is also noted below each pie chart.

B abbreviations are TM = tune mix and C = chemicals. In **C**, the pie charts show the fractions below 2% (white), between 2 and 5% (gray), and above 5% (black). The box and whiskers plots in each panel show the distribution of the values for each detected chemical (boxes are interquartile range, the line is median, the whiskers are 5–96 percentile, and the points are outliers). Horizontal lines indicate thresholds used in each analysis.

For the 4682 unique chemicals analysed, ~45.8% ($n = 2144$) could be detected in at least one analysis mode (Fig. 3). This number was not surprising as some chemicals may have degraded or were not amenable to the ionization sources used (such as some of the metals in the ToxCast library (Table S1) would need to be run with inductively coupled plasma (ICP)-MS). As expected, differences among chemical classes were observed with respect to what ionization mode was preferable for their detection. For example, of the ion species evaluated, PAHs were primarily detected with APCI[+], while PFAS were identified with ESI[-], due to the structural components of both molecular classes (Fig. 3A, B). Figure 3A breaks down the number of chemicals in each classification detected by each ion source. Figure 3B shows the total percentage of unique chemicals detected in every classification and their ion mode for detection. No class of chemicals was

completely detected in the study, and detectability ranged from less than 20% for "Food Additives" to 75% for both "Colour Dyes" and "Pharmaceuticals". Additionally, ESI[+] provided the majority of identifications ($n = 1401$), and of this total, 437 chemicals were detected only with ESI[+] (Fig. 3C). ESI[-] had the second most with $n = 1190$, and here 658 were detected only with ESI[-]. Finally, APCI[+] had the least amount of chemicals detected at $n = 839$, and only 55 were uniquely detected with this source and polarity. To understand the positive mode detections, we also analysed which ions were detected using ESI[+] and APCI[+] (Fig. 3D). For ESI[+], the [M + H]⁺ and [M+Na]⁺ were found to have a similar rate of detection with ~37% of chemicals in ESI[+] producing both ion types. On the other hand, APCI[+] ion types were abundantly represented by [M + H]⁺, while [M]•⁺ detection only included 13.1% of the chemicals.

A All tested substances:
- 45.0% Amenable+ Stable
- 11.6% Amenable+ Unstable
- 19.7% Not Amenable+ Stable
- 23.7% Not Amenable+ Unstable

N=3,658

B ESI [+ or -] detected:
- 72.8% Amenable+ Stable
- 13.0% Amenable+ Unstable
- 8.9% Not Amenable+ Stable
- 5.3% Not Amenable+ Unstable

N=1,736

C ESI [+ or -] not detected:
- 19.9% Amenable+ Stable
- 10.3% Amenable+ Unstable
- 29.5% Not Amenable+ Stable
- 40.3% Not Amenable+ Unstable

N=1,922

Fig. 5 | Analysis of chemical stability and amenability as factors influencing detection. Donut plots show fractions of chemicals based on whether they had predictions for amenability (for ESI [+ or −] ionization mode) and stability among all substances tested in this study. **A** Breakdown for the 3658 substances in our set that had amenability and stability predictions. **B** Breakdown for substances that were detected in our study. **C** Same for the substances that were not detected. The total number of chemicals included in each analysis is shown at the centre of each plot and % of that number is shown in the side legend for each category.

Because the library of CCS values for chemicals analysed herein may be used for suspect screening of complex substances and mixtures by other laboratories, we also assessed reproducibility of CCS values. For this examination, we evaluated reproducibility both between the technical replicates in our own analyses (intra-laboratory) and the differences in our values and those published by others (inter-laboratory) using DTIMS and travelling wave IMS (TWIMS) (Fig. 4). For most chemical classes, over 90% of the substances were within 1% difference. The percent difference in CCS values among the duplicate injections of each chemical were evaluated by chemical class (Fig. 4A) and ionization mode (Fig. 4B). Box and whisker plots show the distribution of the deviation among technical replicates, and the pie charts show the fraction of the substances with values under the generally accepted cutoff of 1%. Interestingly, only "Natural Toxins" had the lowest number of retained chemicals using these cutoffs with 14% removed, however, it was among the classes with the fewest number of chemicals represented in the library. With respect to ionization mode analysis, the overwhelming majority (97%) of the chemicals detected with either ESI[-] or ESI[+] had excellent reproducibility with differences in CCS ≤ 1%. Even though the reproducibility of CCS values was less robust with APCI[+], 91% were below the threshold of 1%. Of the 4192 CCS values for the different ion types evaluated, only 188 were excluded from the final database because they

exceeded 1% difference, resulting in 4004 ions in the final database (Table S6).

Interlaboratory reproducibility was also determined in the study using data from seven other suspect screening CCS libraries[8,17,18,20–22,49,50] (Fig. 4C and Tables S7–S13). Our database containing 4004 CCS values represents a great expansion on important toxicants. Furthermore, in overall CCS reproducibility assessments across laboratories (i.e., % difference), it is often considered acceptable if the difference in reported values is ≤2% to account for instrument variability. A 5% difference was also used for this comparison to consider any larger experimental differences as has been done in several previous studies[12,51,52]. Of the chemicals that could be cross-compared (from 8 to 254 depending on publication), over 82% of the CCS values had ≤2% difference, and only ~3% of the compared chemicals fell above the 5% threshold. Previous work has noted CCS differences for the DTIMS and TWIMS methods for different chemical classes[53], so if readers are interested in understanding this further for their chemicals of interest, this dataset will provide DTIMS reference values for the comparisons. However, having over 87% of the CCS values within 2% difference of the other labs provides us with confidence in our database, especially because all values included in the intra-laboratory comparison fell below the 1% difference threshold and we had very few values outside of the 5% difference when compared to others.

To finalize this study, we attempted to determine reasons for lack of detection of some of the tested compounds with IMS-MS and the ESI and APCI sources. The two main factors considered were stability of the chemicals in solution[42], and amenability of a chemical for detection using a particular ionization source and mode (e.g., ESI[−] and ESI[+])[54,55]. For the chemical stability analysis, we reasoned that chemicals that lack long-term stability in solution may not be detected due to potential degradation prior to analysis. To make this comparison, we were able to obtain information on long-term stability of the ToxCast library from the U.S. EPA and categorize them using this information as "stable" (i.e., no appreciable degradation) or "unstable" (chemical not detected in the analysis of a stock solution after 4 months of being stored at room temperature). The overall number of chemicals included in this analysis was less than the total based on the availability of stability data – out of the 4682 unique chemicals tested, data was only available for 3658. Of these, 37.6% were both detected and stable, while 26.6% were neither detected nor stable. This result was highly significant ($p < 0.0001$ using Fisher's exact test), indicating that stability of the chemicals was a considerable contributor to our ability to determine their CCS value. For the chemical amenability analysis, we reasoned that chemicals that could not be ionized will not be detected. Using previously published predictive chemical amenability analyses with LC-MS[54], we determined which chemicals were "amenable" to ESI[−] or ESI[+] ionization. Figure 5A divides the chemicals that had both stability and amenability information (3658 out of 4685 total analysed) into 4 categories. Figure 5B, C show the breakdown into these categories for chemicals either detected in ESI [+ or −] or not detected. It is clear that for amenable compounds, the majority was detected and vice versa; indeed, amenability was also a major contributing factor to our ability to derive CCS values ($p < 0.0001$ using Fisher's exact test for both ionization modes) as many chemicals were predicted as not amenable. To improve detection of the chemicals we were unable to derive a CCS for, additional analytical methods, such as GC-IMS-MS, will be needed to aid in the multidimensional separations of chemicals that are non-amenable to ESI, such as polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs) and polycyclic aromatic hydrocarbons (PAHs). Of additional note, all chemicals in our study were tested at 10 μM and there was no concentration response performed. Thus, it could not be determined if the concentration of the standard was below the limit of detection or if the chemical had degraded over time to a lower concentration. As mentioned previously, each data file was analysed manually for its specified chemical. This was done to ensure the most accurate analysis and incorporate expert opinions in cases of discrepancies regarding features like saturation, isomers, low abundances,

and n-mer streaking (e.g., where unstable multimers break into monomers during instrumental analysis). However, this does introduce bias into the experiment.

Conclusions

In this study, we utilized IMS-MS to develop a multidimensional database containing experimentally measured CCS and *m/z* values for 2144 unique chemicals from the U.S. Environmental Protection Agency ToxCast programme. The 4685 chemicals evaluated in this study covered thirteen classes including Natural Toxin (*n* = 6), Disinfection By-Product (*n* = 14), PFAS (*n* = 17), Polycyclic Aromatic Hydrocarbon (PAH, *n* = 31), Surfactant (*n* = 78), Flame Retardant (*n* = 120), Plastic (*n* = 154), Colour Dye (*n* = 172), Cosmetic Ingredient (*n* = 360), Pharmaceutical (*n* = 615), Food Additive (*n* = 634), Pesticide (*n* = 948), and Chemical Industrial (*n* = 1536). To create the database, each chemical standard was assessed twice with ESI[+], ESI[−] and APCI[+] followed by IMS-MS analyses. Chemicals were then identified based on expected *m/z* values for $[M]^{\bullet+}$, $[M + H]^+$, $[M+Na]^+$ and $[M\text{-}H]^-$ ion types. From these measurements, only chemicals having CCS values ≤ 1% error in the duplicate assessments were retained as high confidence identifications. This resulted in 2144 unique chemical entries and 4004 ion types. It was also noted that ESI[−] and ESI[+] worked for the largest majority of the chemicals, however APCI[+] covered a unique subset and provided value for certain chemical classes, including the PAHs. Furthermore, comparison of the CCS database values with 7 other laboratories showed ~82% of the overlapping values were within 2% error, indicating interlaboratory reproducibility. Thus, new toxicant knowledge in this chemical library expands on our analytical capabilities and provides a tool for the scientific community to enable rapid IMS-MS screening of a wide range of xenobiotic chemicals, ultimately decreasing response time of exposure assessments and the ability to probe novel disease linkages.

Data availability

Supplementary Tables (Tables S1–S13) have been supplied within the Supplementary Data file. The $^{DT}CCS_{N2}$ values for the 4,004 $[M + H]^+$, $[M+Na]^+$, $[M\text{-}H]^-$ and $[M]^{\bullet+}$ ion types from the 2144 unique chemicals detected are included in Tables S3–S5 and uploaded to PubChem.

References

1. May, J. C. & McLean, J. A. Ion mobility-mass spectrometry: time-dispersive instrumentation. *Anal. Chem.* **87**, 1422–1436 (2015).
2. Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. & Hill, H. H. Jr Ion mobility-mass spectrometry. *J. Mass Spectrom.* **43**, 1–22 (2008).
3. Ibrahim, Y. M. et al. Development of an Ion Mobility Spectrometry-Orbitrap Mass Spectrometer Platform. *Anal. Chem.* **88**, 12152–12160 (2016).
4. Metz, T. O. et al. Integrating ion mobility spectrometry into mass spectrometry-based exposome measurements: what can it add and how far can it go? *Bioanalysis* **9**, 81–98 (2017).
5. Luo, Y. S. et al. Rapid characterization of emerging per- and polyfluoroalkyl substances in aqueous film-forming foams using ion mobility spectrometry-mass spectrometry. *Environ. Sci. Technol.* **54**, 15024–15034 (2020).
6. Foster, M. et al. Uncovering PFAS and other Xenobiotics in the dark metabolome using ion mobility spectrometry, mass defect analysis, and machine learning. *Environ. Sci. Technol.* **56**, 9133–9143 (2022).
7. Kirkwood-Donelson, K. I., Dodds, J. N., Schnetzer, A., Hall, N. & Baker, E. S. Uncovering per- and polyfluoroalkyl substances (PFAS) with nontargeted ion mobility spectrometry-mass spectrometry analyses. *Sci. Adv.* **9**, eadj7048 (2023).
8. Hinnenkamp, V. et al. Comparison of CCS values determined by traveling wave ion mobility mass spectrometry and drift tube ion mobility mass spectrometry. *Anal. Chem.* **90**, 12042–12050 (2018).
9. Jurneczko, E., Kalapothakis, J., Campuzano, I. D., Morris, M. & Barran, P. E. Effects of drift gas on collision cross sections of a protein standard in linear drift tube and traveling wave ion mobility mass spectrometry. *Anal. Chem.* **84**, 8524–8531 (2012).
10. Ewing, M. A., Glover, M. S. & Clemmer, D. E. Hybrid ion mobility and mass spectrometry as a separation tool. *J. Chromatogr. A* **1439**, 3–25 (2016).
11. Gabelica, V. et al. Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **38**, 291–320 (2019).
12. Stow, S. M. et al. An interlaboratory evaluation of drift tube ion mobility-mass spectrometry collision cross section measurements. *Anal. Chem.* **89**, 9048–9055 (2017).
13. Dodds, J. N., Hopkins, Z. R., Knappe, D. R. U. & Baker, E. S. Rapid characterization of Per- and Polyfluoroalkyl Substances (PFAS) by Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS). *Anal. Chem.* **92**, 4427–4435 (2020).
14. Burnum-Johnson, K. E. et al. Ion mobility spectrometry and the omics: distinguishing isomers, molecular classes and contaminant ions in complex samples. *Trends Anal. Chem.* **116**, 292–299 (2019).
15. Ibrahim, Y. M. et al. Development of a new ion mobility (Quadrupole) time-of-flight mass spectrometer. *Int J. Mass Spectrom.* **377**, 655–662 (2015).
16. Dodds, J. N. & Baker, E. S. Ion mobility spectrometry: fundamental concepts, instrumentation, applications, and the road ahead. *J. Am. Soc. Mass Spectrom.* **30**, 2185–2195 (2019).
17. May, J. C., Morris, C. B. & McLean, J. A. Ion mobility collision cross section Compendium. *Anal. Chem.* **89**, 1032–1044 (2017).
18. Celma, A. et al. Improving target and suspect screening high-resolution mass spectrometry workflows in environmental analysis by ion mobility separation. *Environ. Sci. Technol.* **54**, 15120–15131 (2020).
19. Aly, N. A. et al. Utilizing ion mobility spectrometry-mass spectrometry for the characterization and detection of persistent organic pollutants and their metabolites. *Anal. Bioanal. Chem.* **414**, 1245–1258 (2022).
20. Zheng, X. et al. Utilizing ion mobility spectrometry and mass spectrometry for the analysis of polycyclic aromatic hydrocarbons, polychlorinated biphenyls, polybrominated diphenyl ethers and their metabolites. *Anal. Chim. Acta* **1037**, 265–273 (2018).
21. Belova, L., Caballero-Casero, N., van Nuijs, A. L. N. & Covaci, A. Ion Mobility-High-Resolution Mass Spectrometry (IM-HRMS) for the analysis of Contaminants of Emerging Concern (CECs): Database compilation and application to urine samples. *Anal. Chem.* **93**, 6428–6436 (2021).
22. Picache, J. A. et al. Collision cross section compendium to annotate and predict multi-omic compound identities. *Chem. Sci.* **10**, 983–993 (2019).
23. Luo, Y. S. et al. Relationships between constituents of energy drinks and beating parameters in human induced pluripotent stem cell (iPSC)-Derived cardiomyocytes. *Food Chem. Toxicol.* **149**, 111979 (2021).
24. Cordova, A. C. et al. Application of ion mobility spectrometry-mass spectrometry for compositional characterization and fingerprinting of a library of diverse crude oil samples. *Environ. Toxicol. Chem.* **42**, 2336–2349 (2023).
25. Roman-Hubers, A. T. et al. Temporal chemical composition changes in water below a crude oil slick irradiated with natural sunlight. *Mar. Pollut. Bull.* **185**, 114360 (2022).
26. Roman-Hubers, A. T. et al. Characterization of compositional variability in petroleum substances. *Fuel* **317**, 123547 (2022).
27. Valdiviezo, A. et al. Analysis of per- and polyfluoroalkyl substances in Houston Ship Channel and Galveston Bay following a large-scale industrial fire using ion-mobility-spectrometry-mass spectrometry. *J. Environ. Sci.* **115**, 350–362 (2022).

28. Escher, B. I. et al. From the exposome to mechanistic understanding of chemical-induced adverse effects. *Environ. Int.* **99**, 97–106 (2017).

29. Vermeulen, R., Schymanski, E. L., Barabasi, A. L. & Miller, G. W. The exposome and health: Where chemistry meets biology. *Science* **367**, 392–396 (2020).

30. Escher, B. I., Stapleton, H. M. & Schymanski, E. L. Tracking complex mixtures of chemicals in our changing environment. *Science* **367**, 388–392 (2020).

31. Scholz, S. et al. The eco-exposome concept: supporting an integrated assessment of mixtures of environmental chemicals. *Environ. Toxicol. Chem.* **41**, 30–45 (2022).

32. Williams, A. J. et al. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J. Cheminform* **9**, 61 (2017).

33. Barouki, R. et al. The exposome and toxicology: a win-win collaboration. *Toxicol. Sci.* **186**, 1–11 (2022).

34. Lai, Y. et al. High-resolution mass spectrometry for human exposomics: expanding chemical space coverage. *Environ. Sci. Technol.* **58**, 12784–12822 (2024).

35. Brack, W. et al. Effect-directed analysis supporting monitoring of aquatic environments-An in-depth overview. *Sci. Total Environ.* **544**, 1073–1118 (2016).

36. Wambaugh, J. F. et al. Assessing toxicokinetic uncertainty and variability in risk prioritization. *Toxicol. Sci.* **172**, 235–251 (2019).

37. Wambaugh, J. F. et al. Evaluating in vitro-in vivo extrapolation of toxicokinetics. *Toxicol. Sci.* **163**, 152–169 (2018).

38. Valdiviezo, A., Luo, Y. S., Chen, Z., Chiu, W. A. & Rusyn, I. Quantitative in vitro-to-in vivo extrapolation for mixtures: a case study of superfund priority list pesticides. *Toxicol. Sci.* **183**, 60–69 (2021).

39. Ulrich, E. M. et al. EPA's non-targeted analysis collaborative trial (ENTACT): genesis, design, and initial findings. *Anal. Bioanal. Chem.* **411**, 853–866 (2019).

40. Richard, A. M. et al. The Tox21 10K compound library: collaborative chemistry advancing toxicology. *Chem. Res Toxicol.* **34**, 189–216 (2021).

41. Phillips, A. L. et al. A framework for utilizing high-resolution mass spectrometry and nontargeted analysis in rapid response and emergency situations. *Environ. Toxicol. Chem.* **41**, 1117–1130 (2022).

42. Richard, A. M. et al. ToxCast chemical landscape: paving the road to 21st century toxicology. *Chem. Res Toxicol.* **29**, 1225–1251 (2016).

43. Richard, A. M., Truong, H., Wolf, M. & Thillainadarajah, I. (ed National Center for Computational Toxicology (NCCT)) (U.S. Environmental Protection Agency, Office of Research & Development, Research Triangle Park, NC, 2014).

44. Richard, A. M. et al. Analytical quality evaluation of the Tox21 compound library. *Chem. Res Toxicol.* **38**, 15–41 (2025).

45. Baker, E. S. et al. Ion mobility spectrometry-mass spectrometry performance using electrodynamic ion funnels and elevated drift gas pressures. *J. Am. Soc. Mass Spectrom.* **18**, 1176–1187 (2007).

46. GraphPad. KNOWLEDGEBASE - ARTICLE #591: Statistics with n=2, https://www.graphpad.com/support/faqid/591/ (2024).

47. Zhang, X. et al. SPE-IMS-MS: An automated platform for sub-sixty second surveillance of endogenous metabolites and xenobiotics in biofluids. *Clin. Mass Spectrom.* **2**, 1–10 (2016).

48. Baker, E. S. et al. METLIN-CCS: an ion mobility spectrometry collision cross section database. *Nat. Methods* **20**, 1836–1837 (2023).

49. Zheng, X. et al. A structural examination and collision cross section database for over 500 metabolites and xenobiotics using drift tube ion mobility spectrometry. *Chem. Sci.* **8**, 7724–7736 (2017).

50. Song, X. C., Canellas, E., Dreolin, N., Goshawk, J. & Nerin, C. Identification of nonvolatile migrates from food contact materials using ion mobility-high-resolution mass spectrometry and in silico prediction tools. *J. Agric Food Chem.* **70**, 9499–9508 (2022).

51. Song, X. C., Canellas, E., Dreolin, N., Goshawk, J. & Nerin, C. A collision cross section database for extractables and leachables from food contact materials. *J. Agric Food Chem.* **70**, 4457–4466 (2022).

52. Song, X. C. et al. Application of ion mobility spectrometry and the derived collision cross section in the analysis of environmental organic micropollutants. *Environ. Sci. Technol.* **57**, 21485–21502 (2023).

53. Dodds, J. N. et al. Evaluating ion mobility data acquisition, calibration, and processing for small molecules: a cross-platform assessment of drift tube and traveling wave methodologies. *J. Am. Soc. Mass Spectrom.*, https://doi.org/10.1021/jasms.5c00056 (2025).

54. Lowe, C. N. et al. Predicting compound amenability with liquid chromatography-mass spectrometry to improve non-targeted analysis. *Anal. Bioanal. Chem.* **413**, 7495–7508 (2021).

55. Charest, N. et al. Improving predictions of compound amenability for liquid chromatography-mass spectrometry to enhance non-targeted analysis. *Anal. Bioanal. Chem.* **416**, 2565–2579 (2024).

Acknowledgements

Author contributions

The manuscript was written through contributions of all authors, and all authors have given approval to the final version of the manuscript. Contributions for individual authors are noted below: Conceptualization: D.T., N.A.A., J.N.D., I.R., E.S.B. Methodology: D.T., N.A.A., J.N.D. Data Availability: J.Z., P.A.T., E.E.B., K.M.J., EIS. Data Interpretation: D.T., A.J.W., I.R., E.S.B. Resources: J.Z., P.A.T., E.E.B., A.J.W., E.I.S., I.R., E.S.B. Supervision: I.R., E.S.B. Writing—original draft: D.T., I.R., E.S.B. Writing—review & editing: D.T., N.A.A., J.N.D., J.Z., P.A.T., E.E.B., K.M.J., A.J.W., E.I.S., I.R., E.S.B.

Competing interests

The authors declare no competing interests.

Additional information