ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

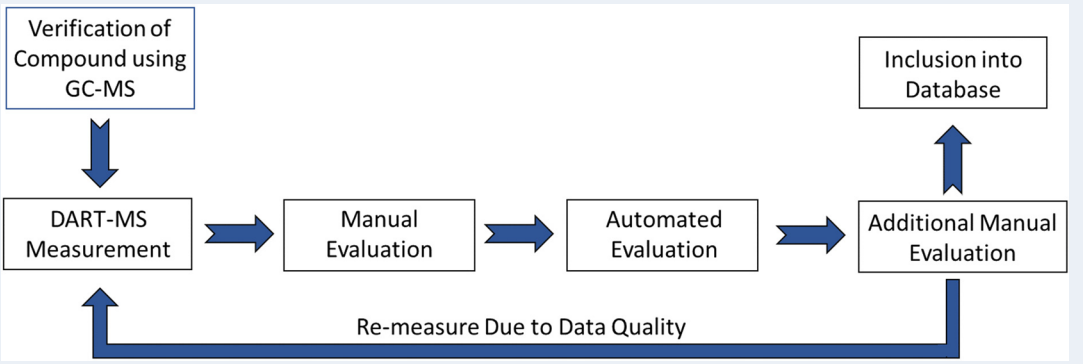
Creation and Release of an Updated NIST DART-MS Forensics Database

Edward Sisco,* Arun S. Moorthy, and Laura M. Watt

Cite This: https://dx.doi.org/10.1021/jasms.0c00416

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Facing increasing caseloads and an everchanging drug landscape, forensic laboratories have been implementing new analytical tools. Direct analysis in real time mass spectrometry (DART-MS) is often one of these tools because it provides a wealth of information from a rapid, simple analysis. The data produced by these systems, while extremely useful, can be difficult to interpret, especially in the case of complex mixtures, and therefore, mass spectral databases are often used to assist in interpretation of data. Development of these databases can be expensive and time-consuming and often relies on manual evaluation of the underlying data. The National Institute of Standards and Technology (NIST) released an initial DART-MS in-source collisional-induced dissociation mass spectral database for seized drugs in the early 2010s but it has not been updated to reflect the increasing prevalence of novel psychoactive substances. Recently, efforts to update the database have been undertaken. To assist in development of the database, an automated data evaluation process was also created. This manuscript describes the new NIST DART-MS Forensics Database and the steps taken to automate the data evaluation process.

KEYWORDS: *DART-MS, database, library, seized drugs, data evaluation*

Seized drug analysis is the most frequently requested forensic examination in the United States, accounting for approximately 33%[1] (or over one million[2]) of all submissions per year. Given the large number of submissions, laboratories are often challenged with trying to reduce turnaround times and backlogs. According to the 2019 National Forensic Laboratory Information System (NFLIS)-Drug Survey of Crime Laboratory Drug Chemistry Sections, nearly 60% of laboratories noted an increase in caseload over the last year and over 40% reported an increase in turnaround time.[3] The average turnaround time for laboratories in the survey ranged from 49 days to 151 days.[3]

One of the major drivers for increased case submissions and turnaround times is the emergence of new psychoactive substances (NPSs) and synthetic opioids. These substances, especially synthetic cathinones and cannabinoids, present analytical challenges due to the continued creation of analogues and other structurally similar chemicals that often make identification and confirmation difficult. Additionally, many of the traditional presumptive tests (e.g., color tests and microcrystalline tests) are not well suited for these compounds.[4] For synthetic cathinones with similar molecular structures, gas chromatography mass spectrometry (GC−MS) fragmentation patterns can often appear indistinguishable and the presence of a molecular ion may not occur.[5] Other compounds, like synthetic opioids, are often present at low concentrations creating detection challenges not only in current presumptive analyses but also in confirmatory analysis by GC−MS, where detection limits, coeluting peaks, or tailing from large amounts of cutting agent may hinder accurate identification. The multicomponent nature of many of these samples can present additional analytical difficulties if a

Received: November 17, 2020
Revised: December 28, 2020
Accepted: February 1, 2021

complete chemical profile is desired. Due to these challenges, over half of laboratories reported emerging drugs as a major contributor to their backlogs and over 80% reported limited analytical tools as a driver for emerging drug analysis issues.[3]

Direct analysis in real time mass spectrometry (DART-MS) is one of the analytical tools forensic chemists are utilizing to help tackle some of the issues posed by NPSs, synthetic opioids, and other emerging drugs.[6] DART-MS is an appealing tool for drug screening applications because it can provide a sensitive, near-complete mass spectral profile of a sample within a matter of seconds. The application of DART-MS for drug analysis has been rigorously demonstrated over the past decade[7] and has excelled with analysis of traditional drugs,[8] emerging drugs,[9] synthetic opioids,[10] pharmaceuticals,[11] plant materials,[12] and other compounds of interest.[13] While DART-MS is traditionally used to obtain a simple presumptive mass spectrum, novel applications of the technique are beginning to emerge. These applications include using solid-phase microextraction for sample cleanup,[14] probing drug residues for investigative purposes,[15] and psychoactive plant species identification.[16] Advances in areas like thermal desorption (TD)-DART-MS have unlocked the ability to use nitrogen as a source gas with little to no impact on detection capabilities,[17,18] relieving potential concerns about high helium consumption.

Like other mass spectrometry-based techniques, interpretation of DART-MS mass spectra is often completed by comparing unidentified mass spectra to those from a reference mass spectral database/library using search algorithms. These algorithms can be simple—such as determining whether peaks of interest are present in a spectrum—to more complex algorithms that take into account multiple mass spectra from different in-source collision-induced dissociation (is-CID) energies to provide probabilistic match values. All of these algorithms depend on a mass spectral database of known compounds to compare against, and for algorithms that incorporate is-CID it is crucial that spectra are obtained under similar instrumental conditions. Development of these databases can be time-consuming and costly, as individual chemical standards need to be procured and analyzed. The resulting data must also be evaluated to ensure it is representative of the compound of interest and adequate for inclusion into the database.

Because of these difficulties in database development and evaluation, laboratories sometimes rely on existing databases from vendors or other agencies. Perhaps the most widely known mass spectral database is the NIST Mass Spectral Library which contains over 350000 electron ionization mass spectrometry (EI-MS) spectra.[19] An initial DART-MS is-CID mass spectral library that included 3217 spectra of 828 compounds (up to four measurements at different is-CID voltages per compound) was created and released by NIST in 2012. Recently, efforts have been made to create an updated database that better represents the emergence, and prevalence, of NPSs and ensures that all compounds have the same number of is-CID spectra. In addition to expanding the database to incorporate these compounds, automated data evaluation tools were developed to streamline and better objectify the determination of spectral suitability for inclusion into the database. This manuscript provides a brief description of the new is-CID mass spectral database and automated evaluation steps.

COMPOUND SELECTION, MEASUREMENT, AND PRELIMINARY MS EVALUATION

The new NIST DART-MS Forensics Database contains 1989 spectra for 663 compounds (exactly three is-CID measurements per compound), though additional entries will be added as new spectra are acquired and more of the old spectra are re-evaluated. At present, only 393 spectra from the original DART-MS library from 2012 are included in the update. New compounds for inclusion were chosen based upon the availability of standards with an emphasis on new psychoactive substances and cutting agents that may be useful in forensic intelligence efforts. All compounds were purchased as solids or 1 mg/mL solutions from commercial vendors (AccuStandard, Alltech, Cayman Chemical, Cerilliant, Fluka, Grace, Millipore-Sigma, Spectrum, Steraloids, and Supelco). For those received as a solid material, solutions, with an approximate concentration of 1 mg/mL, were created in Chromasolv-grade methanol or acetonitrile. Compounds received as 1 mg/mL solutions were used as is. Prior to inclusion into the database, all compounds were analyzed by gas chromatography electron ionization mass spectrometry (GC−EI-MS) to confirm their identities and to ensure the absence of major degradation products.

Collection of mass spectra for inclusion in the database was completed using one of three JEOL AccuTOF mass spectrometers (two AccuTOF-DART 4G systems and one AccuTOF JMS-T100LP system) (Peabody, MA), all of which were coupled with DART-SVP ionization sources (IonSense, Saugus, MA). The operating settings for all instruments were identical, and all mass spectra were collected in positive ionization mode. Relevant DART parameters include the use of helium (99.999% purity) as the DART source gas, a DART gas temperature of 400 °C, and an exit grid voltage of 150 V. Relevant instrument parameters include an orifice temperature of 120 °C, an ion guide voltage of 500 V, an orifice 2 and ring lens voltage of 5 V each, and a detector voltage of 2300 V. The parameter switching function was employed to collect mass spectra at +30, +60, and +90 V orifice 1 voltages with function switching occurring every 0.2 s. Mass spectra were collected from m/z 50 to 800. A dilute, methanolic solution of polyethylene glycol (PEG-600) was used as the mass calibration compound for DART-MS data collection. Spectra at +20 V were not collected, though present in the previous database, as they were found to not typically be used by practicing forensic laboratories and often differ only minorly from the +30 V spectra.

Samples were analyzed by dipping the closed end of a clean glass microcapillary (Corning, Corning, NY) into the solution and introducing it into the DART gas stream. Three to five replicate analyses were completed for each compound. PEG-600 was run at the beginning and end of the run as well as intermittently throughout the run to aid in ensuring mass calibration. Glass microcapillaries not containing a standard were also run with each datafile to provide a signature for background subtraction.

Extraction of mass spectra was completed by integrating the signal of all replicates for a compound and background subtracting the resulting spectrum of a blank microcapillary. Mass calibration of the data was completed differently depending on the software package available for the particular instrument (MassCenter versus msAxel). For systems using MassCenter, a full internal mass calibration was completed for

https://dx.doi.org/10.1021/jasms.0c00416
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

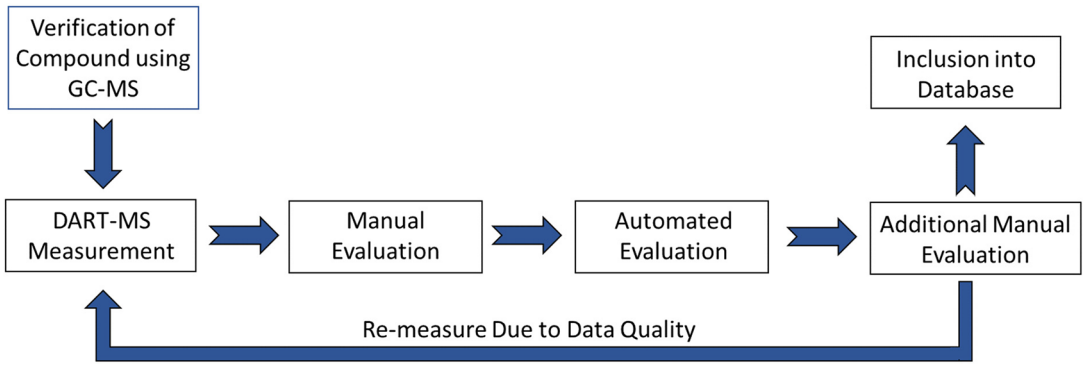


Figure 1. Schematic of overall workflow for inclusion of compound spectra into the database.

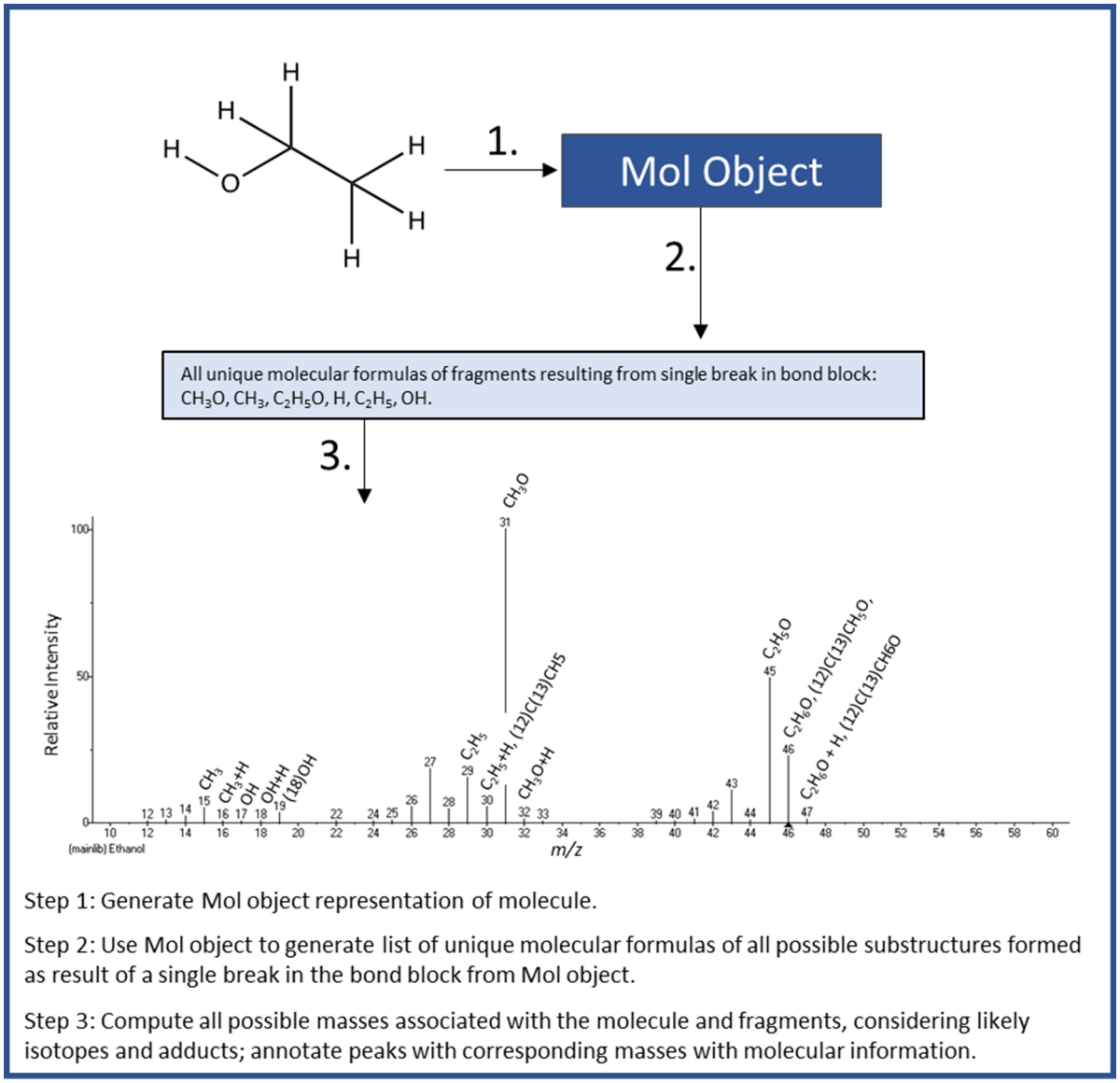


Figure 2. Example of mass spectral substructure annotation procedure for a unit-mass resolution electron ionization mass spectrum (for illustrative purposes).

each datafile while for the systems using msAxel a multipoint drift mass compensation function using an m/z value from the PEG-600 spectrum was employed.

After data extraction, an initial manual evaluation of each mass spectrum was completed, as shown in the workflow depicted in Figure 1. During this initial evaluation, the accurate mass of the protonated molecule (within ±0.005 Da) was verified and a visual inspection of the noise was completed. Compounds exhibiting noisy mass spectra, defined as a visual determination of high background or unexplainable peaks,

https://dx.doi.org/10.1021/jasms.0c00416
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

were remeasured. The presence of a dimer was also examined, and if the dimer was present over approximately 50% relative abundance, the solution was diluted and reanalyzed. Mass spectra that passed the initial manual evaluation were saved in folders based on the orifice 1 voltage used and named according to a unique five-character code assigned to each compound. Along with the extracted mass spectra, a master list of the selected compounds, including essential metadata, was also produced for automated database evaluation and collation. The metadata compiled for the master list included the unique code, conventional name, CAS number (when available), synonyms, IUPAC or formula name, molecular formula, Canonical SMILES, InChi Code, and InChiKey.

AUTOMATED EVALUATION AND DATABASE CONSTRUCTION

With the data produced from the laboratory, a computer program to automate a second evaluation of the mass spectra and collation of the database was written. The tool, referred to as the NIST DART-MS Database Builder, generated a structure data file (.SDF) that encompasses both the structure and spectral information for every compound in the library. A detailed description of the software is included in the Supporting Information. The following set of tests were employed to assess the suitability of the database entries and inform remeasurement or deletion from the final database.

Mass Calibration. For a given analyte, a protonated molecule $[M + H]^+$ peak is expected to be a major peak (e.g., at least 60% of the base peak) in the low orifice energy spectrum (e.g., + 30 V). Based on the specifications of the instrument used, the m/z value of the protonated molecule was evaluated using a mass accuracy of ±0.005 Da. Spectra containing protonated molecule peaks that fell outside of this tolerance were marked for remeasurement.

Probability of Dimer. Ideally, the highest mass peak in a low energy spectrum should be the protonated molecule $[M + H]^+$. However, if the analyte is insufficiently dilute during measurement, it is possible that a protonated dimer $[2\ M + H]^+$ peak is present. If there was a peak within ±0.2 Da of the projected dimer peak, and the peaks intensity was greater than 30% of the intensity of the protonated molecule peak, the spectrum was marked for remeasurement. A ± 0.2 Da window was chosen to ensure detection of dimers if the m/z calibration was off in the higher range.

Fragmentation Consistency. For a given analyte, the level of observable fragmentation between mass spectra recorded at multiple orifice energies is expected to increase with orifice energy values. As there is no obvious measure of the extent of fragmentation, a proxy measure was considered. In particular, the weighted average m/z in the spectrum, using every observed peak and weighted by relative intensity, was determined as a good representative of fragmentation—as the weighted average m/z in a spectrum should decrease with increased fragmentation. Accordingly, compounds where the computed weighted average m/z value did not decrease with increased orifice energy values were marked for further manual evaluation and possible remeasurement.

Noise Evaluation. As an ambient method, some level of background noise is expected in a DART-MS mass spectrum. However, an excess of noise decreases a spectrum's value to a reference library. To identify potentially problematic mass spectra, mass spectral peaks were annotated with potential substructures given the known structure of the analyte and assuming every bond in the molecule can be broken (see Figure 2). Structural rearrangements and more sophisticated fragmentation mechanisms were not considered. Potential substructures included hydrogen and ammonium adducts and considered isotopic masses. Peaks that could not be attributed to any potential substructures were deemed potential noise. The fraction of total intensity attributed to potential noise peaks in each spectrum was computed and mass spectra for which this fraction exceeded a threshold, which was initially set at 0.45 based on visual observations of a subset of compounds, were marked for further manual evaluation and possible remeasurement.

FINAL DATABASE EVALUATION

Mass spectra that were flagged during one or more of the automated evaluation steps were manually reviewed to verify the autogenerated issue and determine whether or not the spectrum was of sufficient quality for inclusion in the library. For compounds exhibiting mass calibration issues, the base peak was further examined. If the base peak was not the protonated molecule, such as compounds that form an $[M - OH]^+$ ion, then the accurate mass of that peak was verified. If the base peak was the protonated molecule and out of calibration, the compound was remeasured. For compounds exhibiting fragmentation inconsistency, the mass spectra for all three orifice 1 voltages were re-extracted and manually inspected to ensure increasing fragmentation was observed and to see if spectra were saved to the incorrect orifice 1 voltage. Compounds that were still in question after this step were remeasured. For compounds with high dimer peaks, the compound solution was diluted by a factor of 10 and remeasured.

Since the noise filter only accounted for individual bond breakages, fragments created through multiple bond breakages and/or rearrangements were flagged as noise. For mass spectra that were flagged as having high noise factors, all assigned peaks above 10% relative abundance were investigated to identify if they were reasonable fragment ions caused by multiple bond breakages and/or rearrangements. If peaks greater than 10% relative intensity could not be explained, the compound was remeasured. A number of other mass spectra were flagged for high noise due to numerous, low-level peaks which, cumulatively, drove up the overall noise factor. These compounds were remeasured to obtain mass spectra with a cleaner baseline and were further concentrated if the noise was due to low analyte signal.

DATABASE AVAILABILITY

The new NIST DART-MS Forensics Database can be downloaded as both a general purpose SDF which can be read through any text editor on any platform and in the NIST format for use with NIST MS Search software[19] on Windows platforms. Both database formats are freely available to download at https://chemdata.nist.gov. The database will be updated on a regular basis to include additional compounds. The NIST DART-MS Forensics Database Builder software tool is also freely available. Instructions to download the builder can be found in the Supporting Information. A list of compounds in the initial version of the database can also be found in the Supporting Information. To suggest compounds for inclusion in the database, email DARTdata@nist.gov.

https://dx.doi.org/10.1021/jasms.0c00416
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.0c00416.

NIST DART-MS Database Builder details (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Edward Sisco − *National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States;* orcid.org/0000-0003-0252-1910; Phone: 301-975-2093; Email: DARTdata@nist.gov

Authors

Arun S. Moorthy − *National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States;* orcid.org/0000-0002-5988-1389

Laura M. Watt − *National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.0c00416

Notes

Certain commercial products are identified in order to adequately specify the procedure; this does not imply endorsement or recommendation by NIST, nor does it imply that such products are necessarily the best available for the purpose.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Edward Erisman of the National Institute of Standards and Technology for evaluating preliminary versions of the library. The authors also recognize Amber Burns of the Maryland State Police Forensic Science Division and Luther Schaeffer of Homeland Security Investigations for their assistance and guidance in collection of spectra. Finally, the authors thank Jacob Easter and Linda Jackson from the Virginia Department of Forensic Science for providing several spectra that were included in the database update.

■ REFERENCES

(1) M. J.Durose, M. R., Burch, A. M., Walsh, K, Tiry, E. *Publicly Funded Forensic Crime Laboratories Resources and Services; 2014 Summary*; U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, 2016.

(2) U.S. Drug Enforcement Administration, Diversion Control Division. *National Forensic Laboratory Information System: NFLIS-Drug 2018 Annual Report*; U.S. Drug Enforcement Administration: Springfield, VA, 2019.

(3) U.S. Drug Enforcement Administration, Diversion Control Division. *NFLIS-Drug 2019 Survey of Crime Laboratory Drug Chemistry Sections Report*; U.S. Drug Enforcement Administration: Springfield, VA, 2019.

(4) Majchrzak, M.; Celiński, R.; Kuś, P.; Kowalska, T.; Sajewicz, M. The newest cathinone derivatives as designer drugs: an analytical and toxicological review. *Forensic Toxicol.* **2018**, *36* (1), 33−50.

(5) Levitas, M. P.; Andrews, E.; Lurie, I.; Marginean, I. Discrimination of synthetic cathinones by GC-MS and GC-MS/MS using cold electron ionization. *Forensic Sci. Int.* **2018**, *288*, 107−14.

(6) Pavlovich, M. J.; Musselman, B.; Hall, A. B. Direct analysis in real time—Mass spectrometry (DART-MS) in forensic and security applications. *Mass Spectrom. Rev.* **2018**, *37*, 171−187.

(7) Steiner, R. R.. Use of DART-TOF-MS for Screening Drugs of Abuse. *Analysis of Drugs of Abuse*; Humana Press: New York, 2018; pp 59−68.

(8) Steiner, R. R.; Larson, R. L. Validation of the direct analysis in real time source for use in forensic drug screening. *J. Forensic Sci.* **2009**, *54* (3), 617−22.

(9) Lesiak, A. D.; Musah, R. A.; Cody, R. B.; Domin, M. A.; Dane, A. J.; Shepard, J. R. E. Direct analysis in real time mass spectrometry (DART-MS) of "bath salt" cathinone drug mixtures. *Analyst* **2013**, *138* (12), 3424−32.

(10) Sisco, E.; Verkouteren, J.; Staymates, J.; Lawrence, J. Rapid detection of fentanyl, fentanyl analogues, and opioids for on-site or laboratory based drug seizure screening using thermal desorption DART-MS and ion mobility spectrometry. *Forensic Chem.* **2017**, *4*, 108−15.

(11) Easter, J. L.; Steiner, R. R. Pharmaceutical identifier confirmation via DART-TOF. *Forensic Sci. Int.* **2014**, *240*, 9−20.

(12) Lesiak, A. D.; Musah, R. A.; Domin, M. A.; Shepard, J. R. E. DART-MS as a Preliminary Screening Method for "Herbal Incense": Chemical Analysis of Synthetic Cannabinoids. *J. Forensic Sci.* **2014**, *59* (2), 337−43.

(13) Robinson, E. L.; Sisco, E. Detection of Brodifacoum and other Rodenticides in Drug Mixtures using Thermal Desorption Direct Analysis in Real Time Mass Spectrometry (TD-DART-MS). *J. Forensic Sci.* **2019**, *64* (4), 1026−33.

(14) Watt, L.; Sisco, E. Detection of trace drugs of abuse in baby formula using solid-phase microextraction direct analysis in real-time mass spectrometry (SPME-DART-MS). *J. Forensic Sci.* **2021**, *66*, 172−178.

(15) Sisco, E.; Robinson, E. L.; Burns, A.; Mead, R. What's in the bag? Analysis of exterior drug packaging by TD-DART-MS to predict the contents. *Forensic Sci. Int.* **2019**, *304*, 109939.

(16) Lesiak, A. D.; Cody, R. B.; Dane, A. J.; Musah, R. A. Plant Seed Species Identification from Chemical Fingerprints: A High-Throughput Application of Direct Analysis in Real Time Mass Spectrometry. *Anal. Chem.* **2015**, *87* (17), 8748−57.

(17) Sisco, E.; Forbes, T. P.; Staymates, M. E.; Gillen, G. Rapid analysis of trace drugs and metabolites using a thermal desorption DART-MS configuration. *Anal. Methods* **2016**, *8* (35), 6494−9.

(18) Sisco, E.; Staymates, M. E.; Forbes, T. P.. Optimization of confined direct analysis in real time mass spectrometry (DART-MS). *Analyst* **2020**. *1452743*.

(19) Mass Spectrometry Data Center, NIST. https://chemdata.nist.gov/ (accessed August 26, 2020).