analytical chemistry

pubs.acs.org/ac Article

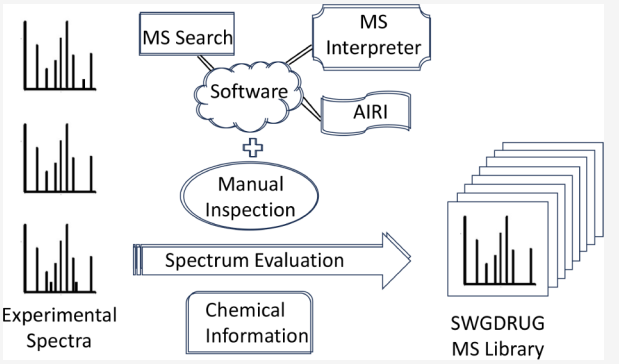
Comprehensive Data Evaluation Methods Used in Developing the SWGDRUG Mass Spectral Reference Library for Seized Drug Identification

Weihua Ji* and William E. Wallace

Cite This: https://doi.org/10.1021/acs.analchem.4c04425 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The mass spectral library of the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) is the most comprehensive free reference database of its kind in the world. It provides reliable mass spectra for identification of seized drugs, their metabolites, and related forensic compounds when using gas chromatography/mass spectrometry (GC/MS). The SWGDRUG library (version 3.13) contains spectra for 3598 compounds. All spectra are evaluated by the Mass Spectrometry Data Center (MSDC) at the National Institute of Standards and Technology (NIST). Over the past few years, new evaluation methods aided by improved software have been developed. First, all chemical information, such as chemical structure and name, is confirmed. Second, the product ions in each spectrum are verified to match the compound structure using the *NIST MS Interpreter* software tool. Subsequently, the mass spectra are compared to the same or similar compounds across six different mass spectral reference libraries using three distinct library search methods. Additionally, the *NIST Artificial Intelligence Retention Indices* (AIRI) software is used to help confirm the corresponding compounds of spectra, especially for those without molecular ions. Low-quality and incorrect spectra are rejected for inclusion in the library.



INTRODUCTION

Gas chromatography coupled with mass spectrometry (GC/MS) is an important analytical technique for identifying emerging seized drugs in forensic studies.[1−8] Such identification is dependent on the availability of comprehensive and high-quality reference mass spectral libraries. These libraries enable rapid and reliable data analysis by comparing an experimental spectrum with reference spectra in the libraries to find the best match.[9,10] This approach provides high specificity and sensitivity in identifying compounds without the need for certified samples during measurement. A critical aspect of library creation is spectral evaluation, which ensures the reliability and accuracy of these libraries. Notably, the most challenging part of building such a library is maintaining quality assurance through expert evaluation.

The Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) mass spectral library is a collaboration between SWGDRUG and the National Institute of Standards and Technology (NIST).[11,12] This library contains reference mass spectra of GC/MS with electron ionization (EI) for seized drugs and forensic-related compounds. It is distributed at www.swgdrug.org/ms.htm as the largest free resource of its kind. It is extensively used by federal, state, and local law enforcement forensic laboratories, and has broad international acceptance. Since it began evaluation of the library in 2015, NIST has been improving its spectral evaluation methods to ensure the library quality, including chemical information by utilizing various continually updated software programs and tools.

NIST previously reported on the spectral evaluation of the SWGDRUG mass spectral library, primarily focusing on the manual inspection of chemical information and spectral comparison among different libraries.[13] Since that time, advances in software development have led to the improvement and creation of new tools to aid spectral evaluation.[14−24] Here, the improved spectral evaluation methods facilitated by software programs and used during the SWGDRUG library evaluation are presented with examples. This work can also serve as a guide for others to build mass spectral reference libraries in any field of study.

Received: August 19, 2024
Revised: September 26, 2024
Accepted: September 30, 2024

Downloaded via NATL INST OF STANDARDS & TECHNOLOGY on October 8, 2024 at 19:26:37 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications
Not subject to U.S. Copyright. Published XXXX by American Chemical Society

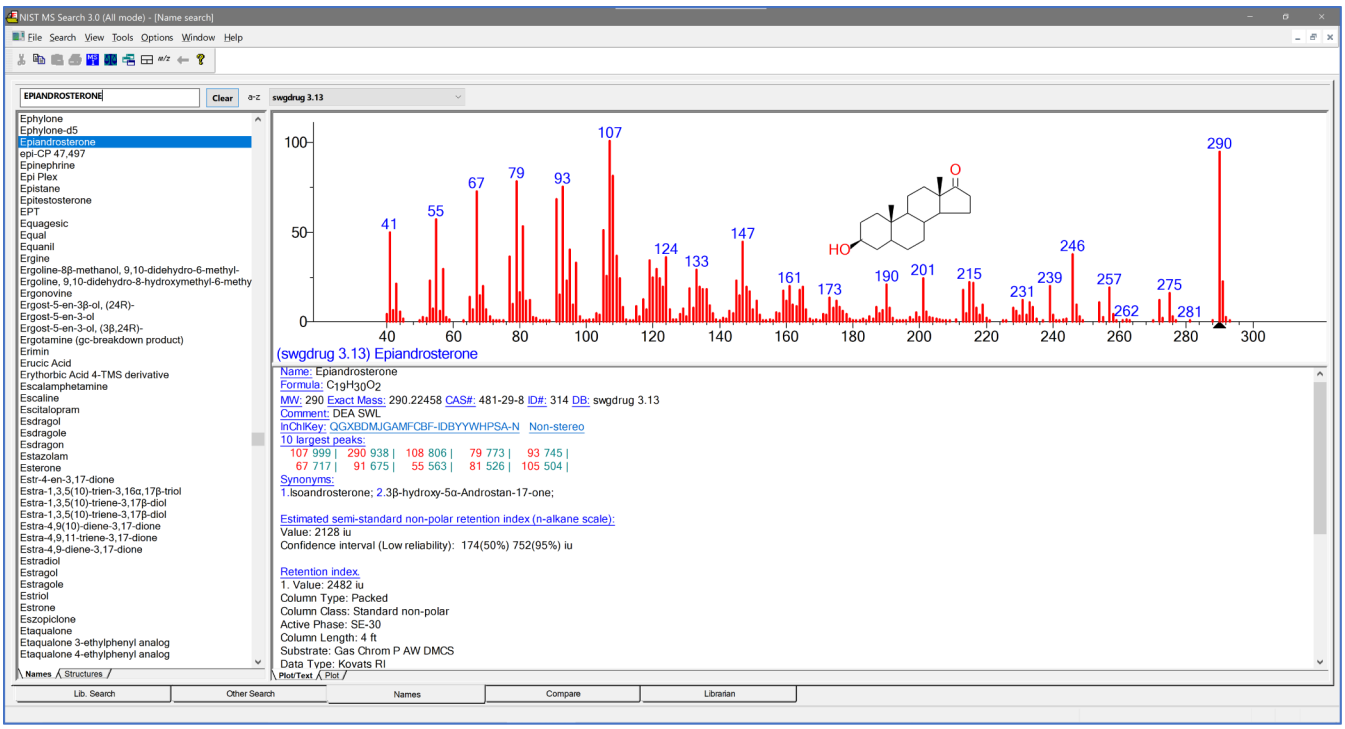


Figure 1. *NIST MS Search* interface for spectral evaluation. The left panel lists all of the compounds in the library. The top right panel displays the spectrum, while the bottom right panel shows detailed chemical, spectral, and retention index information. Pressing the F9 key on the keyboard allows for switching to *NIST MS Interpreter* (Figure 3), facilitating detailed examination of ion fragmentation.

EXPERIMENTAL SECTION

Source of Mass Spectra. Mass spectra of GC/MS with electron ionization were mainly obtained from Cayman Chemical, the US Drug Enforcement Administration (DEA), the American Academy of Forensic Sciences (AAFS), the European Network of Forensic Science Institutes-Drugs Working Group (ENFSI-DWG), and various other contributing institutions.[11] The library in the NIST format was used in this work. Other formats are available on the SWGDRUG website.[11]

Software Programs for Assisting Mass Spectral Evaluation. All software used is freely available at chemdata.nist.gov. *NIST MS Search* 3.0[19] was used to view each mass spectrum along with chemical information including structure, formula, molecular weight, CAS registry number, and synonyms (Figure 1). Spectrum information and the available retention index are also displayed. Additionally, an InChI key was linked to the Google website for users to quickly find more information about the compound. *MS Interpreter* (v3.4.5),[20] which predicts product ions based on ion thermochemistry, was used to annotate product ion peaks in the spectra. Spectral comparison of the same and similar compounds was performed by using six mass spectral libraries: SWGDRUG,[11] NIST23,[21] Cayman,[25] Adams Essential Oil Components,[26] Wiley Designer Drugs,[27] and Wiley Drugs, Poisons, Pesticides, and Pollutants.[28] Three library search methods in *MS Search* were employed for the comparison: identity search (matching each spectrum with spectra in the libraries), structure similarity search (comparing each spectrum with spectra of compounds having a similar chemical structure), and hybrid search (combining identity search and matching fragment ions via the "DeltaMass", that is, the mass difference between the query and library compounds).[22,23] The *NIST Artificial Intelligence Retention Indices* (AIRI)[24] software, utilizing a deep neural network, was also employed. It takes a molecular structure as input and predicts a retention index (RI),[29] which is commonly used in GC identification. The predicted RI is denoted as an AIRI.

Quality Control Procedure. Figure 2 shows the quality control procedure followed. The first step, checking chemical information, was described in reference[13] and has not changed. The product ions in each spectrum consistent with the compound structure were verified using *MS Interpreter*.

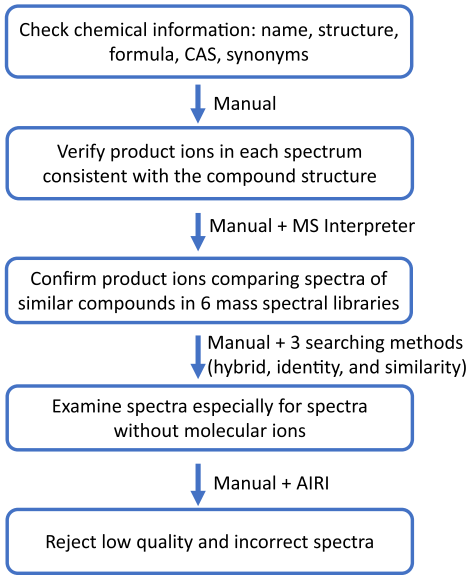


Figure 2. Quality control procedure for evaluating mass spectra in the SWGDRUG library.

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

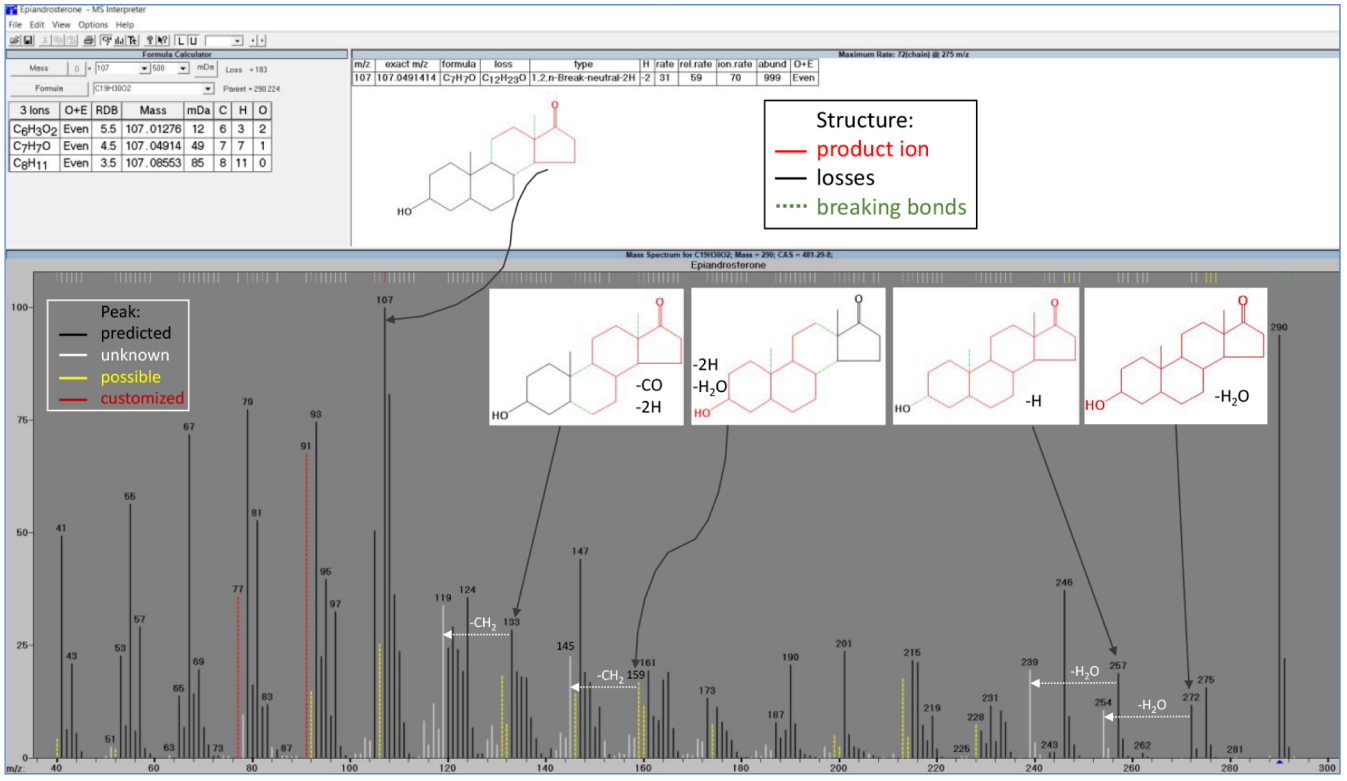


Figure 3. Evaluation of product ions in each mass spectrum in the SWGDRUG library using *MS Interpreter*. In the above view of the user interface, the spectrum is shown in the bottom panel, and proposed fragment structures are displayed in the top panel. When the pointer is moved to select the peak in black with m/z 107, the right top panel displays the putative product ion structure in red, losses in black, and bond breakages in green, along with the formula, loss from the molecular ion, and reaction type and rate. The top left panel lists all possible formulas for the peak. The peaks colored white, yellow, and red indicate unknowns, possible fragments, and *MS Interpreter*-customized fragments, respectively. Different colors make it easy to identify peaks that are not annotated. Product ions for the peaks in white with m/z 254, 239, 145, and 119 were elucidated manually, e.g., the losses ($-H_2O$ and $-CH_2$, respectively) from the predicted product ions are shown in white color.

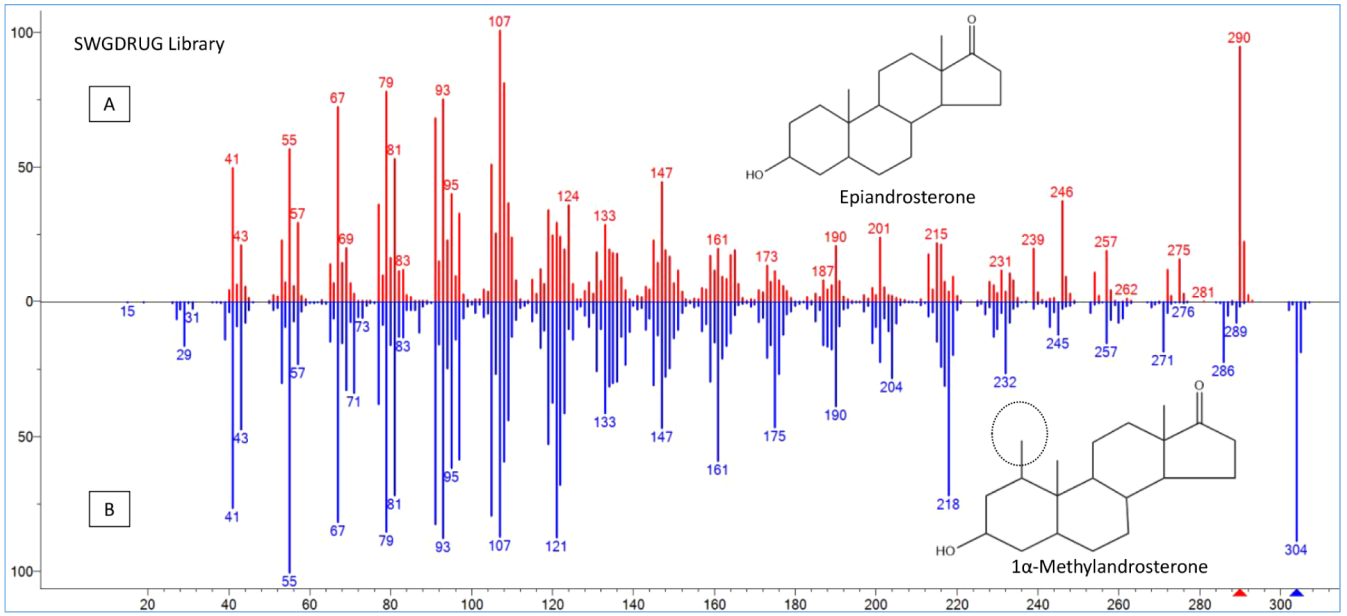


Figure 4. Evaluation of product ions in each mass spectrum for the SWGDRUG library using the similarity search method in *MS Search*. Each spectrum was compared with the reference spectra of similar structure compounds in six mass spectral libraries. A: the spectrum evaluated for the SWGDRUG library; B: a reference spectrum in a library. The chemical structure for the evaluated spectrum in A and that in B are similar except for an additional methyl group in B (shown in the dashed circle).

Subsequently, the spectrum is compared against spectra across the six mass spectral reference libraries utilizing each of the three library search methods (hybrid, identity, and similarity) to confirm common product ions or signature peaks and

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

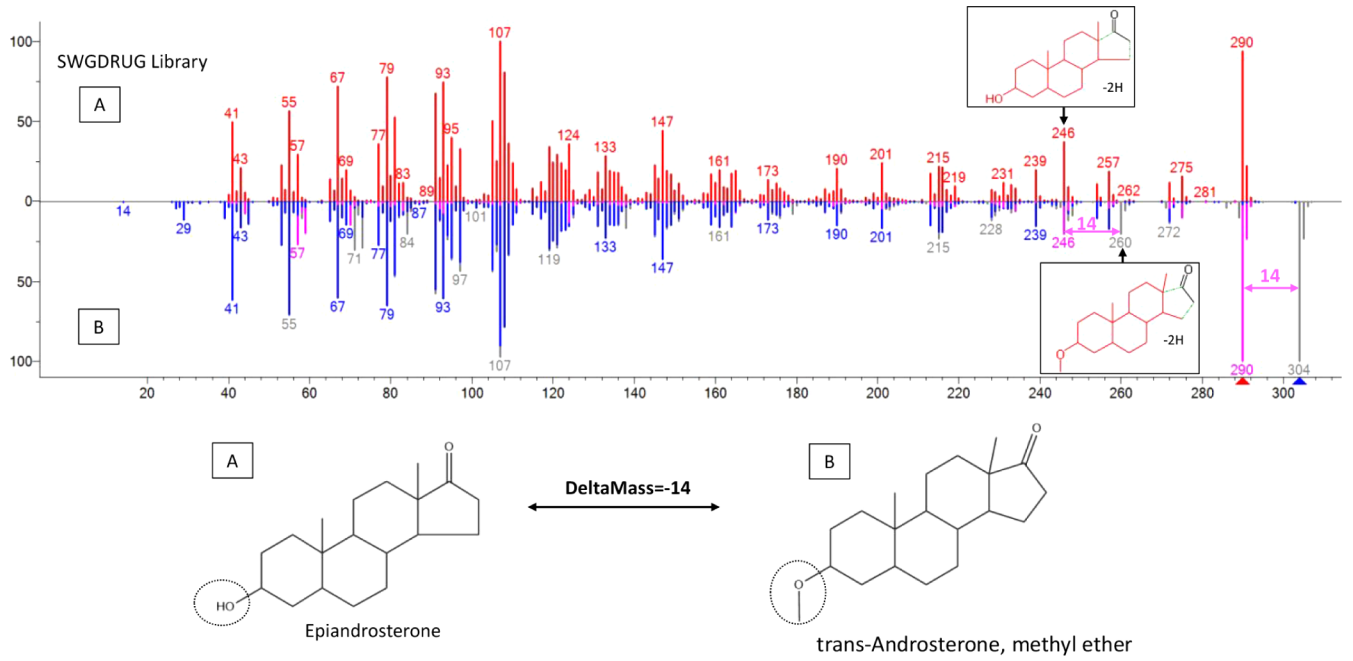


Figure 5. Evaluation of product ions in each mass spectrum for the SWGDRUG library using the hybrid search method in *MS Search*. Each spectrum was compared with similar spectra in six mass libraries. A: the spectrum evaluated for the SWGDRUG library; B: a reference spectrum in a library. Original peaks in gray are shifted by −14 Da (magenta peaks) in the reference spectrum. The chemical structures for the evaluated spectrum in A and the reference spectrum in B are similar except for the group in the dashed circle. Their mass difference (DeltaMass) is −14.

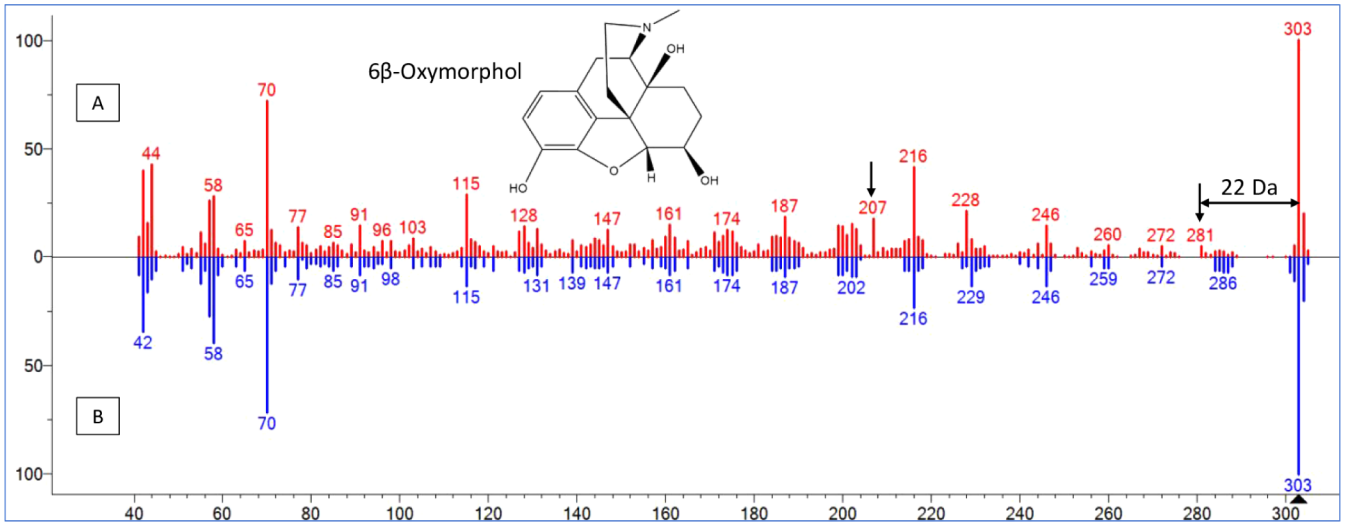


Figure 6. Comparison of a spectrum evaluated for the SWGDRUG library (A) with a reference spectrum of the same compound in the NIST23 Mass Spectral Library (B) using the identity search method in *MS Search* software. Peaks with m/z 281 and 207 in A were not found in B.

ensure spectrum consistency with known-good spectra. AIRI were also used to help confirm the corresponding compounds of spectra, especially for those without molecular ions. All the spectral evaluations involved new software to record evaluator comments and enable further review of spectra by a second analyst. Low-quality and incorrect spectra were rejected by manual inspection.

RESULTS AND DISCUSSION

Consistency of Product Ions with Molecular Structure. To maintain quality reference mass spectra in any library, it is essential to ensure that product ions in each spectrum correspond to the molecular structure. Using *MS Interpreter* greatly accelerated the evaluation process and helped identify errors that are difficult to detect manually. Nevertheless, due to limitations placed on it by the vast array of possible fragmentation mechanisms, the software may not be able to predict all product ions needed to annotate every peak in a spectrum. Therefore, manual elucidation of product ions was necessary. Features in *MS Interpreter* made this a convenient task. For example, in the spectrum of epiandrosterone shown in Figure 3, *MS Interpreter* predicted most product ions, but missed peaks with m/z 254, 239, 145, and 119. By manual elucidation, the product ions of m/z 254 and 239 were attributed to water loss from m/z 272 and 257, respectively, and the product ions of m/z 145 and 119 were attributed to CH_2 loss from m/z 159 and 133, respectively. Through this procedure, all of the major product ion peaks predicted and

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

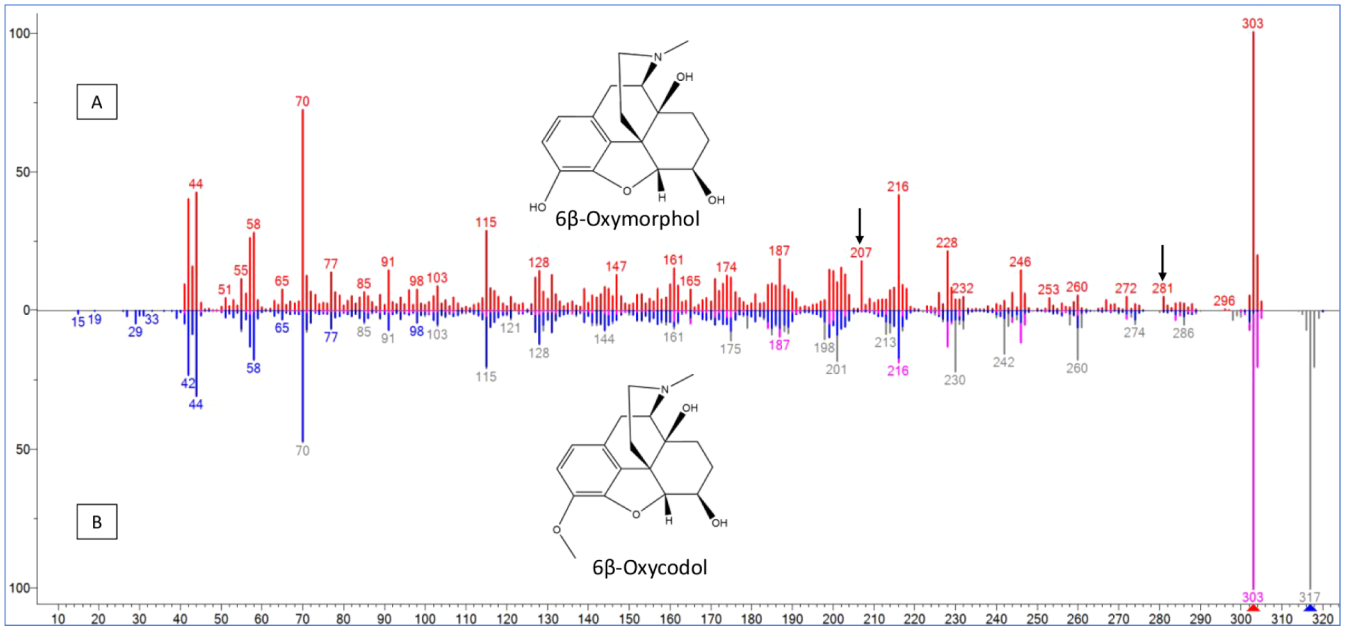


Figure 7. Comparison of a spectrum evaluated for the SWGDRUG library (A) with a reference spectrum of a similar compound in the NIST23 Mass Spectral Library (B) using the hybrid search method in *MS Search* software. Peaks with m/z 281 and 207 were not found in B.

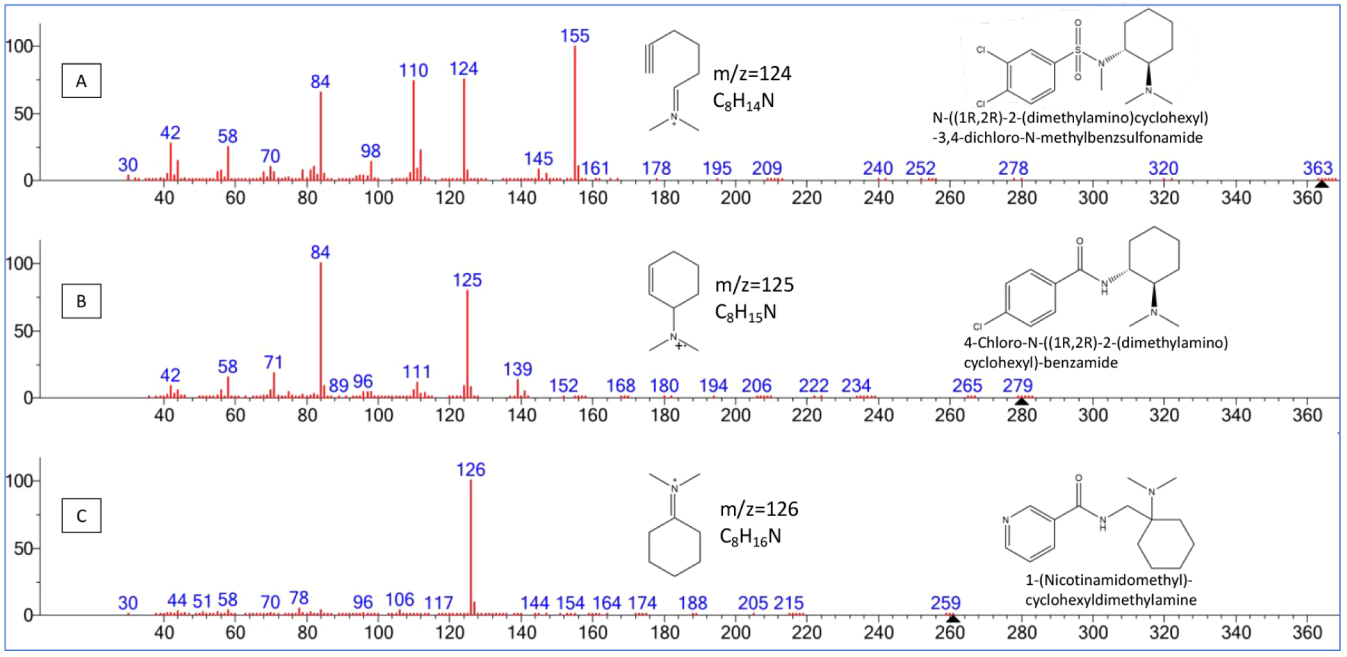


Figure 8. Verification of product ions with three fragmentation mechanisms for *N,N*-dimethylcyclohexanamine in different opioid compounds. Product ion peaks with m/z 124 in A, 125 in B, and 126 in C were elucidated and confirmed.

unpredicted by *MS Interpreter* to match the molecule's structure in each spectrum are verified.

Spectral Comparison of Similar Compounds. To confirm product ions in each spectrum, a comparison was made to similar spectra with the same and similar compounds in the SWGDRUG, NIST23, and four other libraries by using identity, structure similarity, and hybrid searching methods. For example, using the structure similarity search and after consistency checking as shown in Figure 3, the spectrum of epiandrosterone was compared with the one for its analogue 1α-methylandrosterone (Figure 4). Head-to-tail comparison shows that most product ions are common in both spectra between m/z values 40 and 235. Peaks at m/z 145 and 119, which were manually verified (Figure 3), were also observed in the analogue spectrum, confirming these product ions.

Moreover, when the hybrid search was used to compare the spectrum of epiandrosterone with that of *trans*-androsterone, methyl ether, which replaces the hydrogen in the hydroxy group with a methyl group (Figure 5), a match factor of 968 was achieved. Head-to-tail comparison clearly confirmed the product ions at high-end m/z values >235. For instance, the peak with m/z 246 in spectrum A and the peak with m/z 260 in spectrum B have a mass difference (DeltaMass, the mass difference between the query and library compounds) of −14.

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

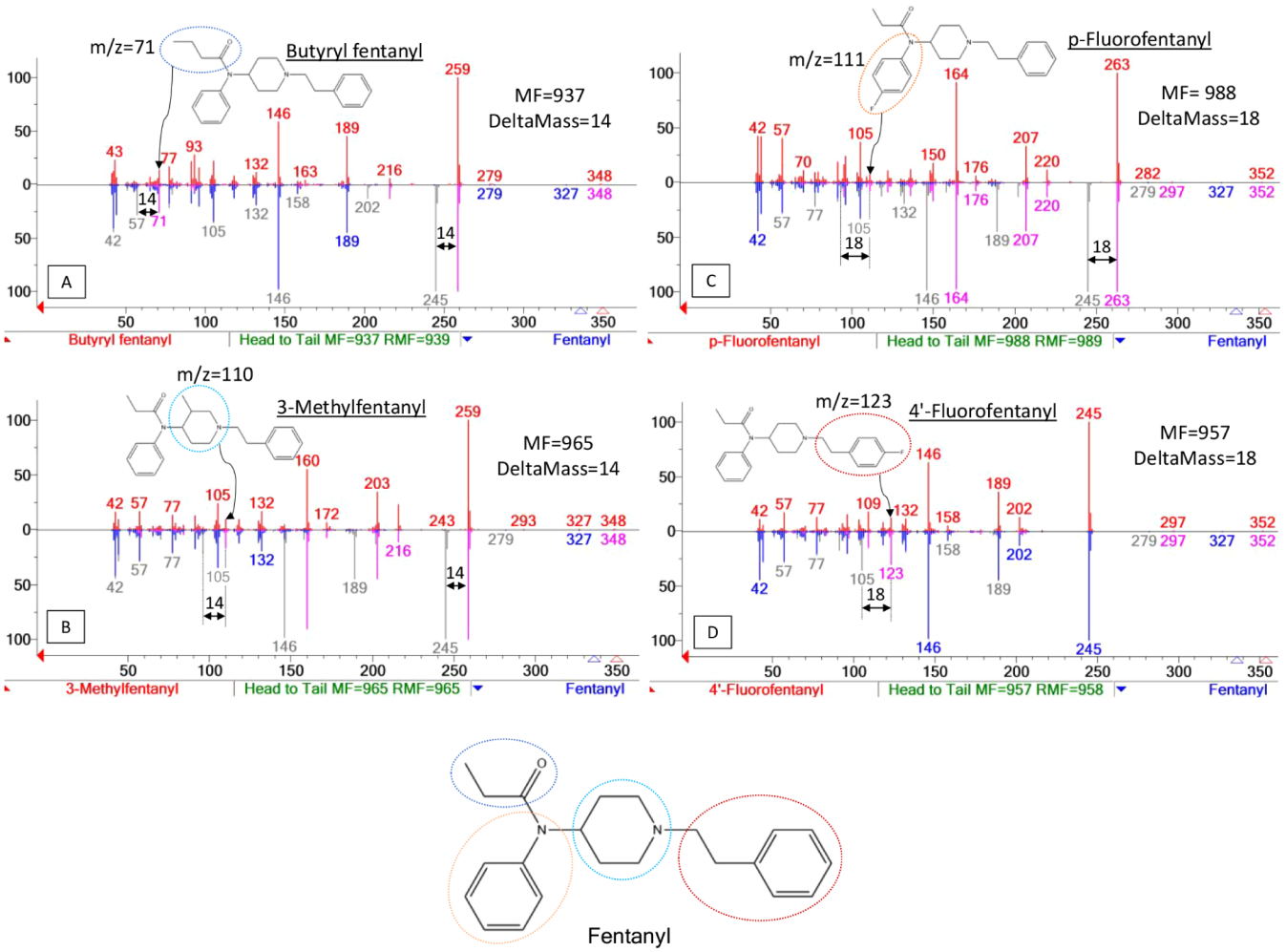


Figure 9. Examination of mass spectra of fentanyl analogues and localization of modification sites using the hybrid search. Modification sites in the fentanyl molecule and signature peaks are shown in A, B, C, and D, respectively. Each structure within the circle indicates the signature peak for the modification. The bottom spectrum depicts fentanyl. MF denotes match factor, i.e., library search score. RMF denotes reverse match factor, excluding peaks in the query spectrum that do not appear in the library reference spectrum.

Their elucidated product ion structures are the same except for the DeltaMass. Additionally, peaks with m/z 254 and 239 that were manually checked in Figure 3 were common in both spectra, as confirmed by the hybrid search. This spectrum was accepted for the library after major ions were all confirmed by consistency checking and spectral comparison with similar compounds.

Low-quality spectra with contaminant, missing, or spurious peaks inconsistent with the compound structure or incorrect product ion assignments were rejected by manual inspection. As an illustration, spectrum A in Figure 6, intended for inclusion in the SWGDRUG library, was questionable. Neither *MS Interpreter* nor manual inspection could identify product ions for the peaks at m/z 281 and 207. Specifically, the m/z 281 peak cannot originate from the molecular ion at m/z 303 due to the 22 Da mass difference. Through the identity search, this spectrum was compared to the reference spectrum for the same compound in the NIST23 Mass Spectral Library (Figure 6B). The result indicated that the reference spectrum did not have peaks at m/z 281 and 207. Moreover, using the hybrid search (Figure 7), the spectra were also compared. The reference spectrum of a similar compound also did not have these two peaks. So, by this process, it was confirmed that these two peaks were likely to be sample contaminants. This low-quality spectrum was rejected from inclusion in the SWGDRUG library.

Different Fragmentation Mechanisms of the Same Chemical Moiety. It was observed that the same chemical group can fragment into different product ions through various mechanisms, depending on the other groups in the compound. As an example, three fragmentation mechanisms for *N*,*N*-dimethylcyclohexanamine in more than 60 opioid compounds with this group in the current SWGDRUG library (Figure 8) were identified. Each mechanism resulted in a product ion at m/z 124, 125, and 126, corresponding to ring opening, double bond formation within the ring, and double bond formation outside the ring, A, B, and C, respectively. These fragments and their associated mechanisms are used to verify product ions and enhance the evaluation efficiency for similar compounds.

Examining Modifications of Fentanyl Analogues. Fentanyl and its analogues make up a crucial class of forensic substances due to their significant role in drug overdose deaths. The SWGDRUG library contains 283 fentanyl analogues. The hybrid search method is particularly useful for examining product ions and assessing the spectral quality of these compounds. With the hybrid search, a single modification on any of the four main structural elements of the fentanyl

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

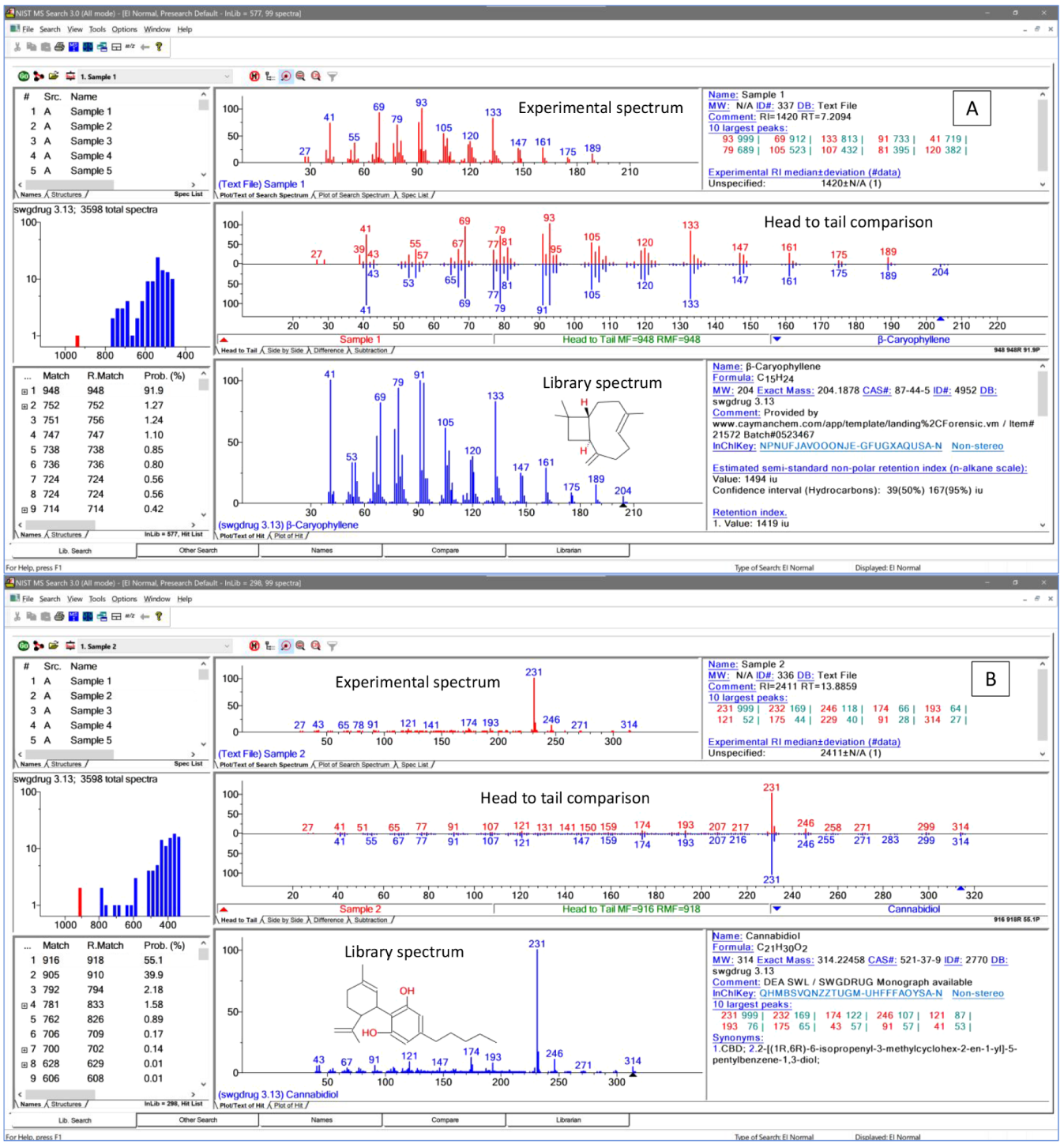


Figure 10. Search of the SWGDRUG library using the identity search method in *MS Search*. In A and B, the experimental spectrum is shown at the top, and the library reference spectrum is at the bottom. The middle window illustrates the head-to-tail comparison between the experimental and library spectra. A. The retention index in the library confirms the compound's identification. B. AIRI supports the confirmation of the library search identification.

molecule can be used to confirm novel fentanyl analogues. As shown in Figure 9, the spectra of four fentanyl analogues with modifications in each area were compared to that of fentanyl using the hybrid search. Each spectrum showed high similarity to fentanyl, with a match factor >930 (a score greater than 900 indicates a good match in general). Notably, the signature peak in each spectrum confirmed the modification site in each analogue, although isomers at different locations (e.g., a methyl group attached to different locations on a phenyl ring) cannot be distinguished. These signature peaks are useful for examining spectra of fentanyl analogues in the SWGDRUG library and for identifying these compounds when searching this library.

New Library Feature—GC Retention Index. In the latest SWGDRUG library, version 3.13, available retention index (RI) values have been appended to spectra where

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX

available, enhancing the reliability of library search results. It should be noted that independent of mass spectrometry, the latest SWGDRUG recommendations, version 8.1 (swgdrug.org/approved.htm), cover the use of gas chromatography, via retention time compared to certified samples, as a Category B method of identification.

Figure 10A shows an example where an experimental spectrum is searched against the SWGDRUG library. Identification is confirmed using both the spectrum searching score and the retention index. The library search score is 948. The retention index values of 1420 for the experimental spectrum and 1419 for the library spectrum confirm the identification, even though the experimental spectrum lacks the molecular ion. As another example shown in Figure 10B, an experimental spectrum was searched against the SWGDRUG library, and a search score of 916 was obtained. The reference spectrum in the library does not have a retention index. The experimental spectrum has a retention index of 2411, while the library spectrum's AIRI was calculated as 2458. These two values helped confirm the identification from the library search. The result indicates that AIRI is a valuable parameter for identification confirmation in library searches, especially for spectra without molecular ions.

CONCLUSIONS

Over the past few years, enhanced spectral evaluation methods applied to the SWGDRUG mass spectral library have been developed. These improvements were centered on enhanced software tools to augment expert manual inspection. The SWGDRUG library has been updated and now contains 3598 compounds, including 283 fentanyl analogues. The latest version 3.13 can be freely downloaded at https://www.swgdrug.org/ms.htm.

AUTHOR INFORMATION

Corresponding Author

Weihua Ji − *Mass Spectrometry Data Center, Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States;* orcid.org/0009-0005-2737-8059; Email: weihua.ji@nist.gov

Author

William E. Wallace − *Mass Spectrometry Data Center, Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.4c04425

Notes

The authors declare the following competing financial interest(s): Certain commercial reference libraries and software are identified in this paper to specify the evaluation procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the reference libraries and software identified are necessarily the best available for the purpose.

ACKNOWLEDGMENTS

The authors thank Drs. Stephen Stein, Yuri Mirokhin, Dmitrii Tchekhovskoi, Lewis Geer, Gary Mallard, Oleg Toropov, and other members of the NIST Mass Spectrometry Data Center for software development, helpful discussions, and technical support. The authors also thank SWGDRUG for their collaboration and support, especially Mr. Jason A. Bordelon (DEA Office of Forensic Sciences) for collecting mass spectra and for building the SWGDRUG library and Dr. Sandra E. Rodriguez-Cruz (DEA Special Testing and Research Laboratory) for detailed manuscript review.

REFERENCES

(1) Jackson, G. P.; Barkett, M. A. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (7), 1210−1224.
(2) Lehrer, M. *Clin. Lab. Med.* **1990**, *10* (2), 271−288.
(3) Hoffmann, W. D.; Jackson, G. P. *Annu. Rev. Anal. Chem.* **2015**, *8*, 419−940.
(4) Ferreira, A. B.; Lobo Castro, A.; Tarelho, S.; Domingues, P.; Franco, J. M. *Aust. J. Forensic Sci.* **2023**, *55* (1), 107−128.
(5) Ferguson, K.; Perr, J.; Tupik, S.; Gilbert, M.; Newman, R.; Winokur, A.; Vallejo, I.; Hokanson, S.; Pothier, M.; Knapp, B.; Icard, M.; Kramer, K.; Almirall, J. *J. Forensic Sci.* **2023**, *68* (5), 1504−1519.
(6) Wei, Q.; Su, F. H. *ACS Omega* **2022**, *7* (23), 19176−19182.
(7) Capistran, B. A.; Sisco, E. *Forensic Chem.* **2022**, *30*, 100435.
(8) Becue, I.; Van Poucke, C.; Van Peteghem, C. *Anal. Bioanal. Chem.* **2011**, *400* (9), 2837−2846.
(9) Stein, S. E. *Anal. Chem.* **2012**, *84* (17), 7274−7282.
(10) Stein, S. E. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (4), 316−323.
(11) *SWGDRUG Mass Spectral Library website.* https://www.swgdrug.org/ms.htm (accessed September 2024).
(12) *NIST and the SWGDRUG Mass Spectral Reference Library of Seized Drugs website.* https://www.nist.gov/programs-projects/nist-and-swgdrug-mass-spectral-reference-library-seized-drugs (accessed September 2024).
(13) Wallace, W. E.; Ji, W.; Tchekhovskoi, D. V.; Phinney, K. W.; Stein, S. E. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 733−738.
(14) Wallace, W. E.; Moorthy, A. S. *J. Forensic Sci.* **2023**, *68*, 1484−1493.
(15) Ji, W.; Wallace, W. E.; Tchekhovskoi, D. V.; Stein, S. E. *New Evaluation Methods for Building a Seized-drug Electron Ionization Mass Spectral Library*, Abstracts of Papers, ACS Fall 2019 ANYL−0299, San Diego, CA, United States, 2019.
(16) Ji, W.; Wallace, W. E. Improved Evaluation Methods for the SWGDRUG Mass Spectral Library, In *Oral presentation (WOG am 09:30), Proceedings of the 70th ASMS Conference on Mass Spectrometry and Allied Topics*; ASMS: Minneapolis, MN, 2022, pp. 59.
(17) Moorthy, A. S.; Kearsley, A. J.; Mallard, W. G.; Wallace, W. E. *Forensic Chem.* **2020**, *19*, 100237.
(18) Moorthy, A.; Kearsley, A.; Mallard, W. G.; Wallace, W. E.; Stein, S. E. *Anal. Chem.* **2023**, *95* (35), 13132−13139.
(19) *NIST MS Search.* https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:nistlibs#nist_search_software (accessed September 2024).
(20) *NIST MS Interpreter.* https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:interpreter (accessed September 2024).
(21) *NIST EI Mass Spectral Library website.* https://www.nist.gov/srd/nist-standard-reference-database-1a (accessed September 2024).
(22) Moorthy, A. S.; Wallace, W. E.; Kearsley, A. J.; Tchekhovskoi, D. V.; Stein, S. E. *Anal. Chem.* **2017**, *89* (24), 13261−13268.
(23) Burke, M. C.; Mirokhin, Y. A.; Tchekhovskoi, D. V.; Markey, S. P.; Heidbrink Thompson, J.; Larkin, C.; Stein, S. E. *J. Proteome Res.* **2017**, *16* (5), 1924−1935.
(24) Geer, L. Y.; Stein, S. E.; Mallard, W. G.; Slotta, D. J. *J. Chem. Inf. Model.* **2024**, *64* (3), 690−696.
(25) Cayman Spectral Library website: https://www.caymanchem.com/forensics/publications/csl (accessed September 2024).

(26) Adams, R. P. *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry*, 4th ed.; Allured Publishing, 2007.
(27) Rösner, P. *Mass Spectra of Designer Drugs*; Wiley, 2022.
(28) Maurer, H. H.; Pfleger, K.; Weber, A. A. Mass Spectral Library of Drugs. In *Poisons, Pesticides, Pollutants and Their Metabolites (DPPP2011)*; Wiley, 2011.
(29) Kováts, E. *Helv. Chim. Acta* **1958**, *41*, 1915−1932.

https://doi.org/10.1021/acs.analchem.4c04425
Anal. Chem. XXXX, XXX, XXX−XXX