Invited contribution to the *Journal of Forensic Sciences* special issue on fentanyl detection

NIST Mass Spectrometry Data Center Standard Reference Libraries and Software Tools: Application to Seized Drug Analysis

William E. Wallace PhD and Arun S. Moorthy PhD

Mass Spectrometry Data Center

Biomolecular Measurement Division

National Institute of Standards and Technology

100 Bureau Drive, Gaithersburg, MD 20899-8362 USA

*To whom correspondence should be addressed: william.wallace@nist.gov

ABSTRACT

The standard reference libraries and associated custom software provided by the National Institute of Standards and Technology's Mass Spectrometry Data Center (NIST MSDC) are described with a focus on assisting the seized drug analyst with the identification of fentanyl-related substances (FRS). These tools are particularly useful when encountering novel substances when no certified sample is available. The MSDC provides three standard reference mass spectral libraries, as well as six software packages for mass spectral analysis, reference library searching, data interpretation, and measurement uncertainty estimation. Each of these libraries and software packages are described with references to the original publications provided. Examples of fentanyl identification by gas chromatography – mass spectrometry (GC-MS) and by direct analysis in real time (DART) mass spectrometry are given. A link to online tutorials is provided.

KEYWORDS

DART-MS, GC-MS, LC-MS/MS, novel psychoactive substances, fentanyl, opioids, seized drug analysis

HIGHLIGHTS

- After reading this paper, the seized drug analyst will be better equipped to identify fentanyls by mass spectrometry.
- This includes never-before-seen fentanyl analogs.
- The methods presented are especially useful when a certified sample is not available.
- An online link to video tutorials is provided.

1. INTRODUCTION

The 2009 US National Research Council report *Strengthening Forensic Science in the United States: A Path Forward* identified the National Institute of Standards and Technology (NIST) as the Federal agency best positioned to provide the measurement methods and standards necessary to advance forensic science. Shortly after the report was published NIST began a long-term, wide-ranging program of research and standards development in the forensic sciences [1] working with other government agencies to create an integrated national response, for example to fight the national opioid epidemic [2]. As part of this broad effort the NIST Mass Spectrometry Data Center (MSDC) began a program aimed at aiding the seized drug analyst.

NIST has a long history of providing the Nation with standard reference data [3] including the first reference library of mass spectra in 1945 [4]. The MSDC currently provides an ever-expanding assortment of fully-evaluated mass spectral reference libraries [5]: an electron ionization (EI) library for gas chromatography-mass spectrometry (GC-MS), a tandem MS library for liquid chromatography-mass spectrometry (LC-MS/MS), and numerous material-specific libraries for metabolomics, glycomics, and proteomics. These material-specific libraries are not reviewed here as they are of less interest to the seized drug analyst. The MSDC also provides the NIST GC Retention Index (RI) Library with experimental and predicted retention indices for each of entries in the EI mass spectral library as well as other compounds. The general-purpose EI and tandem MS libraries, GC RI library, along with a

newly updated DART (Direct Analysis in Real Time) mass spectral library (aimed specifically at controlled substances), provide the seized drug analyst with a wide suite of standard reference data that can be used for compound identification.

To improve the practical utility of the reference data, the MSDC produces software for data processing, library building, library searching, uncertainty quantification, and quality assurance. General purpose tools include ***AMDIS*** for processing GC-MS raw data, ***MS Search*** for searching mass spectral libraries, and ***MS Interpreter*** for determining the origin of fragment ions using a measured mass spectrum and putative structure. The new DART-MS seized drug library can be searched with ***MS Search*** but is optimized for use with the new ***NIST/NIJ DART-MS Data Interpretation Tool*** (*DIT*) which employs the new ***Inverted Library Search Algorithm*** (*ILSA*) to analyze seized drug samples containing multiple components. In addition to the aforementioned software tools, the MSDC research portfolio includes novel algorithms for mass spectral comparisons and interpreting library search results. These algorithms are often published as open-source prototype software tools for early adopters and other research groups. For example, the ***MSDC min-max test*** provides a simple and objective method for discriminating samples from replicate mass spectra, and the ***NIST Fentanyl Classifier*** leverages multidimensional scaling and numerical cluster analysis in addition to hybrid similarity library searching to determine if an unknown mass spectrum is an unrecorded fentanyl analog.

This article will review the data and software products of the NIST MSDC with an eye towards the needs of the seized drug analyst, specifically with the identification of

fentanyl-related substances (FRS) by mass spectrometry. FRS are defined as compounds which have the fentanyl core molecular structure, typically modified in one or more positions on the molecule, or have a core molecular structure that is closely related to fentanyl. There is ambiguity in this definition; however, a strict definition is not necessary for the purpose of library searching. Additionally, fentanyl classifications which are based on legal or physiological considerations are not necessary for library searching and identification.

A summary of libraries and software to be discussed are given in Table 1. Hyperlinks are provided for easy access. These tools are particularly useful in cases where a certified standard is not available or when an unexpected adulterant has been used. Examples of using MSDC tools in concert to identify unknown compounds are provided, as is discussion of on-going work and future plans in the NIST MSDC.

2. THE NIST STANDARD REFERENCE MASS SPECTRAL LIBRARIES

The NIST23 EI library contains spectra for 346,964 compounds of which 259 are FRS. All entries are measured at 70 eV electron energy. Most of the 346,964 library entries include an experimentally-measured GC retention index (RI) [6], and all entries include GC-RI values predicted with a group additivity model [7] and a separate artificial intelligence model [8]. Each RI model was developed and tested extensively by the MSDC. Library entry metadata include a common name and any synonyms (often very many), molecular formula, molecular

mass (exact and nominal), chemical structure (both as a two-dimensional drawing and as a MOL file including stereochemistry as appropriate), InChI key, Chemical Abstracts Service (CAS) registration number (when available and unambiguous), and a clickable weblink to NIH/NCBI PubChem for additional information about the chemical compound. Prior to inclusion in the library, the spectrum, measured retention index, and metadata are carefully evaluated [9].

The NIST23 tandem MS library contains spectra for 51,502 compounds of which 224 are FRS. Samples were measured by electrospray ionization. MS/MS spectra are taken at multiple collision energies depending on how readily the precursor ion fragments. Spectra are measured for multiple precursor ion types and when there are in-source fragments of intensity comparable to that of the intact precursor ion. Metadata fields are the same as for the EI library. Entries to the tandem MS library are given an evaluation much like that for the EI library but tailored to the specifics of tandem mass spectrometry [10, 11]. Although LC-MS/MS is not typically used to identify seized drugs it has been demonstrated to have the required specificity to do so even when only trace amounts are available and has been used for that purpose [12, 13]. LC-MS/MS is used extensively in forensic toxicology which is beyond the scope of this review.

Recent additions have grown the NIST DART-MS library to better represent the emergence of novel psychoactive substances (NPS). The new library contains sets of in-source collision-induced dissociation (is-CID) mass spectra for 1232 controlled substances of which 226 are FRS. All compounds have three is-CID mass spectra taken at the same three fragmentation energies,

+30 V, +60 V, and +90 V [14]. Using all three is-CID mass spectra improves identifiability of unknowns during a library search as will be described in the next section. In addition to expanding the library to incorporate these compounds, automated data evaluation tools were developed to streamline and better determine spectral suitability for inclusion into the library.

3. ANALYSIS SOFTWARE

In preface to this section, it must be stated that each piece of software described is multifunctional and the result of years of development by the staff of the MSDC. Listing all functionalities for every program would make any review prohibitively long. The best approach is to download the software and use it while supplementing the experience by reading the embedded help file and the user manual. For additional support, online tutorials and archived workshops can be found with a simple web search.

AMDIS (Automated Mass Spectral Deconvolution and Identification System) is the software embodiment of an integrated set of procedures for extracting pure chemical component mass spectra from complex gas chromatograms that may exhibit overlapping elution peaks [15]. The overall data analysis process involves several steps. The first step extracts statistical signal characteristics from the data file for later use in noise processing and threshold setting. The second step identifies the individual chromatographic components and determines a numerical model peak shape for each. The third step extracts spectra from the individual ion

chromatograms using the model shape, explicitly subtracting nearby background ions when necessary. The extracted spectra can then be automatically searched with two elementary library search algorithms implemented within ***AMDIS***. More importantly, the extracted spectra can be forwarded to the ***MS Search*** program for more comprehensive library searching.

The NIST ***MS Search*** program [16] treats each mass spectrum as an n-dimensional vector and uses a modified vector dot product [17] to determine the closest matching spectra in the searched libraries. The closest matches are then sorted by similarity to form a 'hit list' showing the highest match factor at the top of the list. ***MS Search*** has the ability to search 126 libraries simultaneously including user-created libraries. Tools and instructions within the program aid the analyst in creating their own libraries. In 2017, NIST added the 'hybrid similarity search' (HSS) [18] to the ***MS Search*** program. The HSS calculates a similarity score for spectra of compounds with substantially similar fragmentation mechanisms that differ by the insertion, deletion, or replacement of a discrete, inert chemical moiety that is not active in the fragmentation process. This has the effect of increasing the coverage of a library to include neighboring compounds. These neighbors exist in a notional 'mass spectral chemical' space. The algorithm works by allowing each peak in the unknown query spectrum match library peaks in two ways: by the typical direct match or by matching a mass-shifted library peak. This is an algorithmic expression of the Biemann 'shift technique' [19, 20] something that most mass spectrometrists use intuitively when evaluating unknown spectra without a library. The HSS automates this intuitive process. To work effectively, the HSS requires the precursor mass

of the query compound. The molecular ion is not often present in EI spectra. Recent work on the ***MSDC Molecular Mass Estimator for Electron Ionization Mass Spectrometry*** has shown the utility of several methods for inferring the unknown compounds molecular mass from either characteristics of the spectrum itself, or from an analysis of the simple search hit list [21]. Computer code for this is available, see Table 1 for a link. Lastly, the HSS has been shown to work well not only for EI mass spectra but for tandem mass spectra as well [22, 23].

The NIST ***MS Interpreter*** program [24, 25, 26] takes as input a chemical structure and a mass spectrum obtained from any time of mass spectrometer. It then assigns peaks to plausible molecular substructures using an extensive set of thermochemical bond strength estimates and mechanistic fragmentation rules based on the proposed chemical structure. A calculated rate for every bond dissociation gives its relative reactivity. From these rates relative peak heights are estimated. The fragmentation rules were developed initially from well-known fragmentation patterns in organic mass spectrometry [27, 28] and then optimized by analysis of the NIST mass spectral libraries. ***MS Interpreter*** makes several assumptions about ion fragmentation during the generation of mass spectra: 1) reaction rates follow the general principles of transition state theory; 2) only relative rates are significant and peak origins are arranged in order of increasing energy; 3) energy distributions can be ignored and reactions occur solely through electronic ground states. It estimates relative dissociation rates using bond dissociation energetics for ions published in NIST data compilations, for example, see [29]. Individual rules account for more complex ring opening, H-transfer, and dissociation

reactions that occur during ion fragmentation. Recent modifications are also made to explicitly account for sequential dissociation commonly found at elevated energies in collision cell fragmentation [26].

The ***NIST Fentanyl Classifier*** [30] takes results of a targeted hybrid similarity library search and constructs a mass spectral similarity map. This symmetric, square, *pseudo* distance matrix can then be further interrogated numerically and visually; for example, the elements of this map each representing a single library entry plus the unknown can be clustered. This gives the analyst a visual relationship between the unknown and the compounds most similar to it in the library. Statistical analysis of the clusters can be used to determine the uncertainty in the fentanyl classification.

The strength of DART mass spectrometry for the seized drug analyst has three aspects: 1) ease of sample preparation, 2) short measurement time, and 3) the presence of a protonated/deprotonated intact analyte molecule in the mass spectrum in most cases. What DART-MS lacks is chromatographic separation before introduction into the mass spectrometer. This makes mixture analysis difficult. The ***inverted library search algorithm*** (ILSA) [31, 32] searches input query mass spectra for partial patterns that are similar to reference pure compound mass spectra contained in the reference library, an approach that is particularly useful when the query spectrum contains signatures of multiple compounds. The ILSA also simultaneously compares multiple is-CID mass spectra, collected at varying fragmentation energies, thereby providing a more comprehensive assessment of similarity between a set of

query and library spectra. The ILSA has been implemented in the interactive ***NIST/NIJ DART-MS Data Interpretation Tool*** (DIT) [33] developed with assistance from the National Institute of Justice (NIJ).

In library searching, a common question is "what is a good score?" for confidently deciding that the unknown compound has been correctly identified. Using numerical scores is not recommended without visual inspection of the mass spectra—libraries and software are tools that should be used carefully, not blindly [34]. However, mathematical approaches can be used to shed light on the task of mass spectral discrimination. The ***MSDC min-max test software*** [35] uses two or more measurements for both reference compounds and for the unknown to approximate uncertainty in the measurements. The measurements can be taken at the same time or different times. They can be taken on the same instrument or different instrument. All that is required is that they are the same type of ionization mechanism. The decision if two samples are different compounds is made by comparing the most conservative estimate of intrasample spectral similarity to the most generous estimate of intersample similarity. Thus, the min-max test provides exclusionary evidence and supports negative confirmation tasks, that is, it aids in deciding if two compounds are *not* the same. In spirit, min-max test is similar to the work of Bodnar Willard and coworkers [36, 37, 38] using an unequal variance t-test to compare two spectra peak-by-peak across to find statistically significant differences in measurement intensity. A more comprehensive approach for discriminating mass spectra using replicate mass spectra involves constructing and comparing

high-dimensional consensus (HDC) mass spectra [39]. In theory, the performance of the HDC method improves with the number of measurements. These exclusionary approaches are under development in the MSDC. Lastly, positive confirmation—deciding whether two samples are the same compound—is appreciably more difficult. This is also an active area of research in the MSDC [40] but is still in its early development stages.

4. EXAMPLES

The first of two examples provided demonstrates NIST software tools for GC-MS analysis, while the second demonstrates the ***DART-MS Data Interpretation Tool***.

Figure 1 shows an example GC-MS data file loaded and analyzed in ***AMDIS***. The analysis was done with low sensitivity settings in ***AMDIS***. Increasing sensitivity settings in ***AMDIS*** will increase the number of components identified by the program, which may come in handy if the user suspects the sample had trace components. With low sensitivity settings, ***AMDIS*** determined that there are mass spectra representing 12 components in the sample.

If the analysis were performed using only the default mass spectral library and basic identification system implemented in ***AMDIS***, only one of the components is provided with proposed structure because of the limited size of the default AMDIS mass spectral library. However, the user can also analyze the data by searching the newest NIST EI-MS database by selecting the *'Search NIST Library'* option in the *'Analyze'* dropdown menu. When searching the

NIST library with ***MS Search*** using its more advanced methods, all 12 components return possible matches.

In many cases, using the basic search function in ***AMDIS*** can be sufficient for data interpretation. However, it can be fruitful to send extracted mass spectra of interest to ***MS Search*** for deeper interrogation. Figure 2 shows the results of searching the extracted mass spectrum of component 8 (RT 4.6458 min) using ***MS Search***. ***AMDIS*** had proposed the structure of this component as *para*-fluorobutyryl fentanyl (FIBF), and the identity search in ***MS Search*** makes clear that the spectrum also matches the library entries for FIBF and *ortho*-fluorobutyryl fentanyl with high similarity (shown in the hit list in lower left of the figure). In addition to spectral searching, ***MS Search*** has several other ways to interact with NIST libraries and probe query mass spectra, for example, searching by molecular mass, by specific fragment ions, by neutral losses, among many others.

Both ***AMDIS*** and ***MS Search*** propose identifications for query mass spectra using spectral similarity matching to library spectra. While this identification procedure is generally effective, using ***MS Interpreter*** to further assess the consistency between a query mass spectrum and a proposed chemical structure can improve confidence in the identification. Figure 3 shows how well a query mass spectrum can be explained by the structure of *para*-fluorobutyryl fentanyl and the fragmentation rules implemented in ***MS Interpreter***. The black peaks in Figure 3 are peaks that are identified by the fragmentation logic of ***MS Interpreter***. The white colored peaks indicate peaks that are not explained by the implemented fragment rules—in the example, the

highest intensity peaks in the mass spectrum are all well-explained. For example, the base peak at *m/z* 277 is the well-known loss of the benzyl moiety [41].

Between ***AMDIS***, ***MS Search*** and ***MS Interpreter***, most compounds found in case work (including analysis beyond fentanyl-related substances) should be identifiable. The ***NIST Fentanyl Classifier*** is a newly developed experimental tool that expands the interpretation of library search results and can be helpful for the specific challenge of fentanyl analog identification. Figure 4 shows the results of searching a query mass spectrum in the ***NIST Fentanyl Classifier***. The results are presented as both a hit list (similar to what one might expect from a traditional library search) and a spectral similarity map. Using the hit list, spectral similarity map, and a set of expert rules, the classifier proposes that the unknown spectrum is a fentanyl analog with a single modification occurring at the *N*-propionyl group on the fentanyl scaffold [42]. The fentanyl classifier has shown to work well in identifying fentanyl analogs that differ by up to two moieties from unmodified fentanyl; the tool is continuously being evaluated and the underlying algorithms are being extended to other classes of drugs. The ovals around each of the three clusters in Figure 4 represent 95% confidence intervals that a compound is of that cluster type.

In addition to GC-MS analysis, the NIST MSDC in collaboration with other groups supports the analysis of DART-MS. Figure 5 shows the results of analyzing two is-CID mass spectra of a sample, collected at +30 V and +60 V voltages, using the ***NIST/NIJ DART-MS Data Interpretation Tool*** (DIT). These mass spectra can be searched against the NIST DART-MS

Forensics Database using ***MS Search***, but the user would need to then reconcile the results manually. Using the DIT, target peaks (that is, potential molecules) are identified in the +30 V is-CID mass spectrum and cumulative scores using all available is-CID spectra are provided for each possible match to each target peak. Note that a user is able to include a +90 V is-CID mass spectrum in the search as well; including additional spectra can improve discernability of the results. Additionally, it should be noted that work by Couch, *et al.* [43] has demonstrated the applicability of the NIST/NIJ DIT across different mass spectrometer platforms and even different ionization sources (i.e., electrospray ionization), which helps extend the application of these resources beyond that of specific model mass spectrometer used to create the NIST DART-MS library.

A useful source of free, online tutorials can be found at the Forensics@NIST 2022 website. The videos titled "Mass Spectral Interpretation" (parts 1 through 4) provide 5 hours of expert description on the application of the NIST MSDC libraries and tools.

5. CONCLUSIONS AND RECOMMENDATIONS

Mass spectral reference libraries, used in conjunction with custom software tools, can aid seized drug analysts in identifying compounds. This is especially useful when no certified standard is available. However, reference libraries and software are an aid to the analyst, not a replacement. The tools are also useful for quality audits of compound identification procedures, for proficiency testing, and for educational purposes.

The surest way to become adept at using the MSDC tools is to begin to integrate them into the analyst's daily work. Only with continual use will proficiency follow. Each library and software tool described here comes with user documentation and is described in the peer review publications cited. The online tutorials are the best place to start toward the goal of improving analyst accuracy and efficiency.

REFERENCES

1. Ballou SM. The NAS Report: Ten Years of Response. J Forensic Sci. 2019;64(1):6-9. doi: 10.1111/1556-4029.13961.

2. Morrow JB, Ropero-Miller JD, Catlin ML, Winokur AD, Cadwallader AB, Staymates JL, et al. The Opioid Epidemic: Moving Toward an Integrated, Holistic Analytical Response. J Anal Toxicol. 2019;43(1):1-9. doi: 10.1093/jat/bky049.

3. Semerjian HG, Jr DRB. Data Programs at NBS/NIST: 1901-2021. J Phys Chem Ref Data. 2022;51(1). doi: 10.1063/5.0084230.

4. Brewer AK, Dibeler VH. Mass Spectrometric Analyses of Hydrocarbon and Gas Mixtures. J Res Natl Bur Stand. 1945;35(2):125-139. doi: 10.6028/jres.035.003.

5. Stein S. Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical Identification. Anal Chem. 2012;84(17):7274-7282. doi: 10.1021/ac301205z.

6. Babushok VI, Linstrom PJ, Reed JJ, Zenkevich IG, Brown RL, Mallard WG, et al. Development of a database of gas chromatographic retention properties of organic compounds. J Chromatogr A. 2007;1157(1-2):414-421. doi: 10.1016/j.chroma.2007.05.044.

7. Stein SE, Babushok VI, Brown RL, Linstrom PJ. Estimation of Kovats retention indices using group contributions. J Chem Inf Model. 2007;47(3):975-980. doi: 10.1021/ci600548y.

8. Geer LY, Stein SE, Wallace WE, Using Generalized Chemical Artificial Intelligence to Calculate Molecular Properties, including GC Retention Indices. In: Proceedings of the 67th ASMS Conference on Mass Spectrometry and Allied Topics; 2019; Atlanta, GA. American Society for Mass Spectrometry. Proceedings Paper 300504

9. Wallace WE, Ji WH, Tchekhovskoi DV, Phinney KW, Stein SE. Mass Spectral Library Quality Assurance by Inter-Library Comparison. J Am Soc Mass Spectrom. 2017;28(4):733-738. doi: 10.1007/s13361-016-1589-4.

10. Yang XY, Neta P, Stein SE. Quality Control for Building Libraries from Electrospray Ionization Tandem Mass Spectra. Anal Chem. 2014;86(13):6393-6400. doi: 10.1021/ac500711m.

11. Yang XY, Neta P, Stein SE. Extending a Tandem Mass Spectral Library to Include MS2 Spectra of Fragment Ions Produced In-Source and MSn Spectra. J Am Soc Mass Spectrom. 2017;28(11):2280-2287. doi: 10.1007/s13361-017-1748-2.

12. Sisco E, Najarro M, Burns A. A snapshot of drug background levels on surfaces in a forensic laboratory. Forensic Chem. 2018;11:47-57. doi: 10.1016/j.forc.2018.09.001.

13. Sisco E, Najarro M, Burns A. Quantifying the effectiveness of cleaning agents at removing drugs from laboratory benches and floor tiles. Forensic Chem. 2019;12:1-7. doi: 10.1016/j.forc.2018.11.002.

14. Sisco E, Moorthy AS, Watt LM. Creation and Release of an Updated NIST DART-MS Forensics Database. J Am Soc Mass Spectrom. 2021;32(3):685-689. doi: 10.1021/jasms.0c00416.

15. Stein SE. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. J Am Soc Mass Spectrom. 1999;10(8):770-781. doi: 10.1016/s1044-0305(99)00047-1.

16. Stein SE. Estimating Probabilities of Correct Identification from Results of Mass Spectral Library Searches. J Am Soc Mass Spectrom. 1994;5(4):316-323. doi: 10.1016/1044-0305(94)85022-4.

17. Stein SE, Scott DR. Optimizing and Testing of Mass-Spectral Library Search Algorithms for Compound Identification. J Am Soc Mass Spectrom. 1994;5(9):859-866. doi: 10.1016/1044-0305(94)87009-8.

18. Moorthy AS, Wallace WE, Kearsley AJ, Tchekhovskoi DV, Stein SE. Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification. Anal Chem. 2017;89(24):13261-13268. doi: 10.1021/acs.analchem.7b03320.

19. Biemann K. The Determination of Carbon Skeleton of Sarpagine by Mass Spectrometry. Tetrahedron Lett. 1960;1(36):9-14. doi: 10.1016/S0040-4039(01)99327-X.

20. Biemann K, Bommer P, Burlingame AL, McMurray WJ. The High-Resolution Mass Spectra of Ajmalidine and Related Substances. Tetrahedron Lett. 1963;4(28):1969-1973. doi: 10.1016/S0040-4039(01)90951-7.

21. Moorthy AS, Kearsley AJ, Mallard WG, Wallace WE, Stein SE. Inferring the nominal molecular mass of an analyte from its electron ionization mass spectrum. ChemRxiv. 2023. doi: 10.26434/chemrxiv-2023-wpff0.

22. Burke MC, Mirokhin YA, Tchekhovskoi DV, Markey SP, Thompson JH, Larkin C, et al. The Hybrid Search: A Mass Spectral Library Search Method for Discovery of Modifications in Proteomics. J Proteome Res. 2017;16(5):1924-1935. doi: 10.1021/acs.jproteome.6b00988.

23. Cooper BT, Yan XJ, Simon-Manso Y, Tchekhovskoi DV, Mirokhin YA, Stein SE. Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries. Anal Chem. 2019;91(21):13924-13932. doi: 10.1021/acs.analchem.9b03415.

24. Mayorov A, Mirokhin Y, Tchekhovskoi DV, Stein SE, A New Release of NIST Mass Interpreter for High Mass Accuracy Spectra. In: Proceedings of the 65th ASMS Conference on Mass Spectrometry and Allied Topics; 2017; Indianapolis, IN. American Society for Mass Spectrometry. Proceedings Paper 288991

25. Mayorov A, Mirokhin Y, Tchekhovskoi DV, Stein SE, Enhancements to NIST MS Interpreter for Modeling High Mass Accuracy Tandem Mass Spectra. In: Proceedings of the 66th

ASMS Conference on Mass Spectrometry and Allied Topics; 2018; San Antonio, TX. American Society for Mass Spectrometry. Proceedings Paper 293146

26. Mayorov A, Mirokhin Y, Tchekhovskoi DV, Stein SE, New Developments in the Modeling of Ion Fragmentation by MS Interpreter Software. In: Proceedings of the 67th ASMS Conference on Mass Spectrometry and Allied Topics; 2019; Atlanta, GA. American Society for Mass Spectrometry. Proceedings Paper 296051

27. Mikaia A. Protocol for Structure Determination of Unknowns by EI Mass Spectrometry. I. Diagnostic Ions for Acyclic Compounds with up to One Functional Group. J Phys Chem Ref Data. 2022;51(3):031501. doi: 10.1063/5.0091956.

28. Mikaia A. Protocol for Structure Determination of Unknowns by EI mass Spectrometry. Part 2. Diagnostic Ions in One Ring Alicyclic, Heterocyclic, and Aromatic Compounds. J Phys Chem Ref Data. 2023. doi:

29. Lias SG, Bartmess JE, Liebman JF, Holmes JL, Levin RD, Mallard WG. Gas-Phase Ion and Neutral Thermochemistry. J Phys Chem Ref Data. 1988;17:1-861. doi: srd.nist.gov/JPCRD/jpcrdS1Vol17.pdf.

30. Moorthy AS, Kearsley AJ, Mallard WG, Wallace WE. Mass spectral similarity mapping applied to fentanyl analogs. Forensic Chem. 2020;19(Jun):100237. doi: 10.1016/j.forc.2020.100237.

31. Moorthy AS, Sisco E. A New Library-Search Algorithm for Mixture Analysis Using DART-MS. J Am Soc Mass Spectrom. 2021;32(7):1725-1734. doi: 10.1021/jasms.1c00097.

32. Moorthy AS, Tennyson SS, Sisco E. Updates to the Inverted Library Search Algorithm for Mixture Analysis. J Am Soc Mass Spectrom. 2022;33:1260-1266. doi: 10.1021/jasms.2c00090.

33. Sisco E, Appley MG, Tennyson SS, Moorthy AS. Qualitative Analysis of Real Drug Evidence Using DART-MS and the Inverted Library Search Algorithm. J Am Soc Mass Spectrom. 2022;33(9):1784-1793. doi: 10.1021/jasms.2c00166.

34. Sparkman OD. Evaluating electron ionization mass spectral library search results. J Am Soc Mass Spectrom. 1996;7(4):313-318. doi: 10.1016/1044-0305(95)00705-9.

35. Moorthy AS, Sisco E. The Min-Max Test: An Objective Method for Discriminating Mass Spectra. Anal Chem. 2021;93(39):13319-13325. doi: 10.1021/acs.analchem.1c03053.

36. Willard MAB, Hurd JE, Smith RW, McGuffin VL. Statistical comparison of mass spectra of salvinorins in Salvia divinorum and related Salvia species. Forensic Chem. 2020;17:9. doi: 10.1016/j.forc.2019.100192.

37. Willard MAB, McGuffin VL, Smith RW. Statistical comparison of mass spectra for identification of amphetamine-type stimulants. 2017;270:111-120. doi: 10.1016/j.forsciint.2016.11.013.

38. Willard MAB, Smith RW, McGuffin VL. Statistical approach to establish equivalence of unabbreviated mass spectra. Rapid Commun Mass Spectrom. 2014;28(1):83-95. doi: 10.1002/rcm.6759.

39. Roberts MJ, Moorthy AS, Sisco E, Kearsley AJ. Incorporating measurement variability when comparing sets of high-resolution mass spectra. Anal Chim Acta. 2022;1230:340247. doi: 10.1016/j.aca.2022.340247.

40. Kearsley AJ, Moorthy AS. Identifying fentanyl with mass spectral libraries. ChemRxiv. 2021. doi: 10.26434/chemrxiv.14176952.v1.

41. Kanamori T, Iwata YT, Segawa H, Yamamuro T, Kuwayama K, Tsujikawa K, et al. Characterization and Differentiation of Geometric Isomers of 3-methylfentanyl Analogs by Gas Chromatography/Mass Spectrometry, Liquid Chromatography/Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy. J Forensic Sci. 2017;62(6):1472-1478. doi: 10.1111/1556-4029.13395.

42. Drug Enforcement Adminstration— Schedules of Controlled Substances: Temporary Placement of Fentanyl-Related Substances in Schedule I. Fed Reg. 2018;83(25):5188-5192. doi: www.govinfo.gov/app/details/FR-2018-02-06/2018-02319.

43. Couch AN, Sharp J, Davidson JT. Assessing the effectiveness of the NIST DART-MS Forensics Database and Data Interpretation Tool for designer drug screening with alternative instrumentation. Int J Mass Spectrom. 2023;483:10. doi: 10.1016/j.ijms.2022.116964.

Table 1: List of reference mass spectral libraries and software discussed in this manuscript. Included are web links for direct access to the resources.

NIST/EPA/NIH EI Mass Spectral Library	List of Commercial Distributors (Chemdata.NIST.Gov) DOI: 10.18434/T4H594
NIST Tandem Mass Spectral Library	List of Commercial Distributors (Chemdata.NIST.Gov) DOI: 10.18434/T4H594
NIST DART Mass Spectral Library	NIST DART-MS Library (Chemdata.NIST.Gov) DOI: 10.18434/mds2-2313
AMDIS	AMDIS (Chemdata.NIST.Gov)
MS Search	MS Search (Chemdata.NIST.Gov)
MS Interpreter	MS Interpreter (Chemdata.NIST.Gov)
MSDC Molecular Mass Estimator for Electron Ionization Mass Spectrometry	MSDC Molecular Mass Estimator (GitHub)
NIST/NIJ DART-MS Data Interpretation Tool	NIST/NIJ DART-MS Data Interpretation Tool (Data.Gov)
MSDC Min-Max Test	MSDC min-max Tool (Data.Gov)
NIST Fentanyl Classifier	NIST Fentanyl Classifier (GitHub)

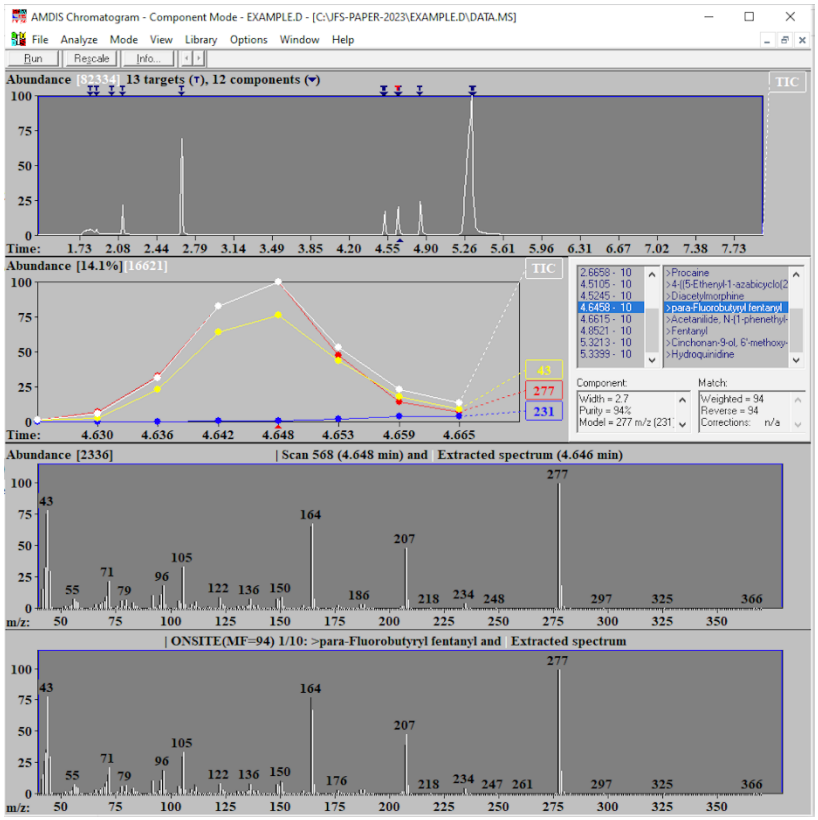

Figure 1: The **AMDIS** user interface showing the gas chromatogram in the top pane, an expanded view of the elution peak at 4.648 min below that, the mass spectrum extracted from the chromatographic peak, and on the bottom the NIST library match to the unknown. In this example the unknown, component 8, is identified as *para*-fluorobutyryl fentanyl. In the chromatographic pane the total ion current (TIC) is shown in white and three select ions at *m/z* 43, 231, and 277 are shown in different colors (yellow, blue, and red respectively). Any *m/z* value can be selected to display its partial ion chromatogram. Under "Component" AMDIS gives a measure of the 'Purity' of the spectrum, that is, what percentage of ion intensity in the elution peak is accounted for in the extracted spectrum. In this example it is 94%.

Figure 2: The full **MS Search** from component 8 in the GC-MS datafile presented in Figure 1. On the right the three panes show, from the top, the unknown spectrum, a butterfly plot comparing the unknown to the library spectrum, and the unknown spectrum shown on the bottom pane. The lower left pane shows the library search 'hit list'. Each entry in the hit list can be examined in comparison to the unknown. The histogram in the middle left pane shows the hit list values showing three high likelihood identifications with many other unlikely identifications having much lower match factors.

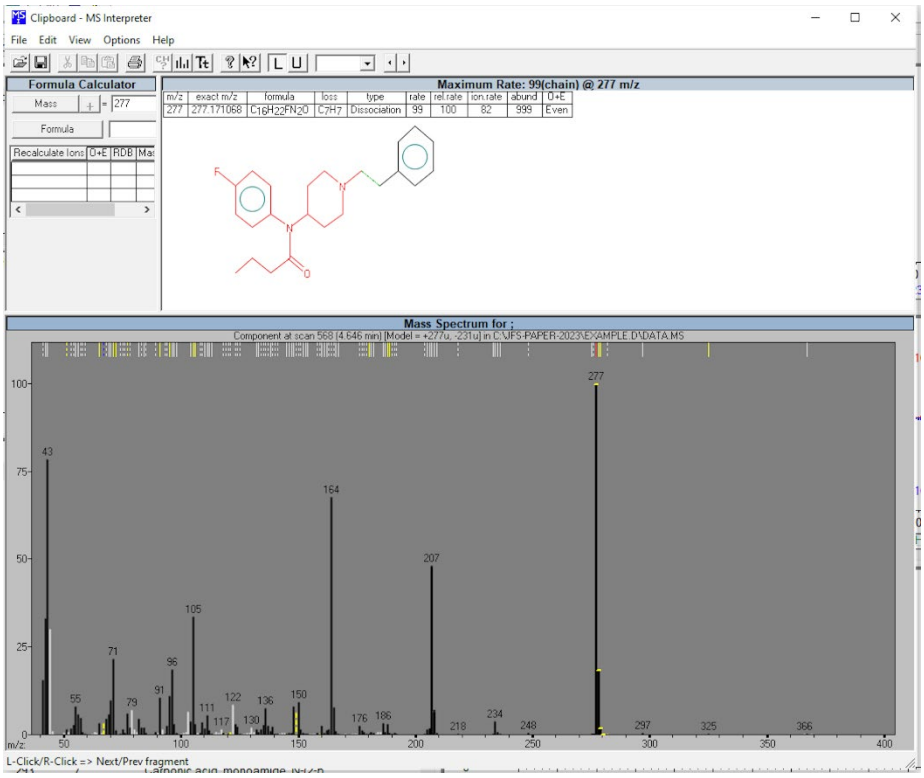

Figure 3: The **MS Interpreter** user interface. The fragment at *m/z* 277 is shown in red on the chemical structure with the bond necessary to create that ion shown as a dotted green line. Choosing different peaks in the spectrum highlights the part of the molecule from which they originated. Using the computer pointer to circle any part of the chemical structure will calculate the mass of the hypothetical ion. Black peaks are those correctly predicted by **MS Interpreter** while white peaks (e.g., m/z 122) are not. For m/z 122 the program can give possible formulas based on the formula of the analyte molecule which would be shown in the upper left under 'Formula Calculator.'

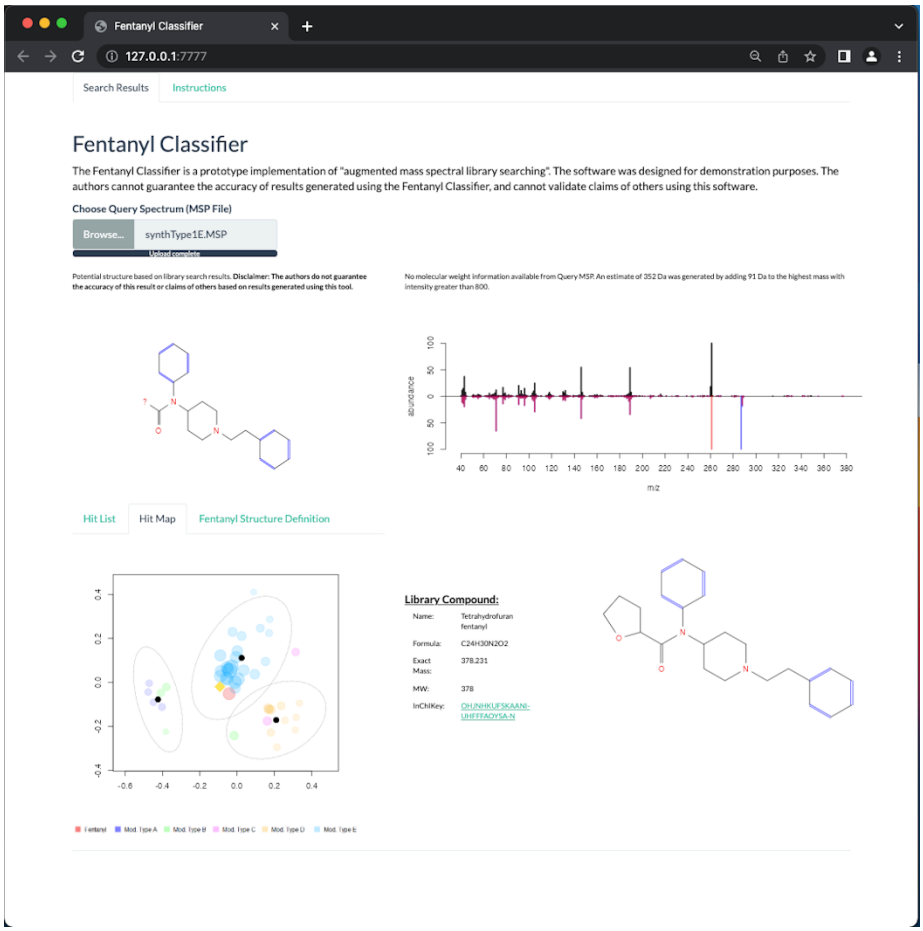

Figure 4: The **NIST Fentanyl Classifier** user interface showing the mass spectral similarity map in the lower left corner. In the top left corner, the classifier proposes a modification at the position shown at the red question mark in the left side chemical structure. The butterfly plot in the upper right shows the unknown spectrum compared to the (user) selected library spectrum (tetrahydrofuran fentanyl) with the structure and metadata of the library compound shown below. The light red and light blue peaks in the bottom spectrum show which peak was shifted by the hybrid similarity search to match the unknown with the library spectrum. The key to modification locations can be seen by choosing the 'Fentanyl Structure Definition' tab.

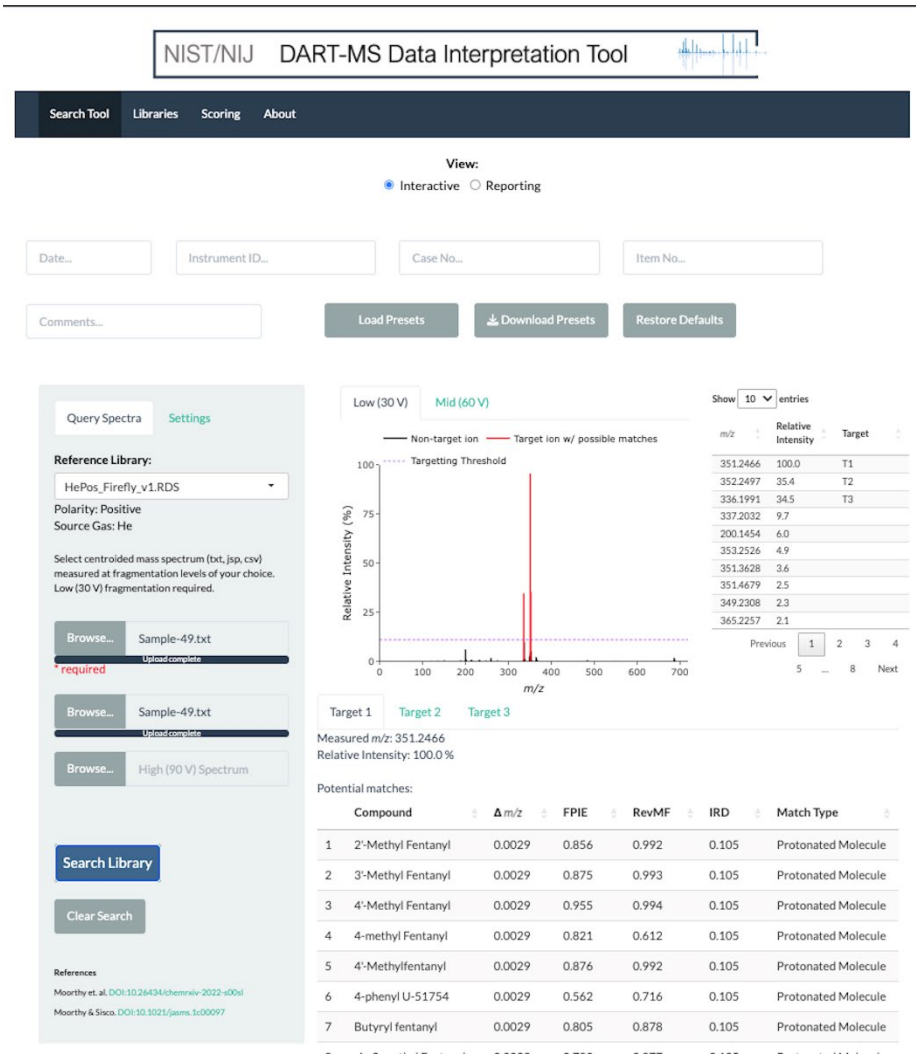

Figure 5: The **NIST/NIJ DART-MS Data Interpretation Tool** (DIT) showing the query spectrum in the middle and the search hit list at the bottom. The pink horizontal line in the spectrum adjusts the level below which peaks are ignored in the library search.