Advances in the Use of Mass Spectral Libraries for Forensic Toxicology

Beat Aebi* and **Werner Bernhard**

Institute of Legal Medicine, University of Berne, Buehlstrasse 20, CH-3012 Berne, Switzerland

Abstract

Gas chromatography in combination with mass spectrometry (GC–MS) plays an important role in the field of analytical toxicology. The identification of unknown compounds is very frequently undertaken with GC–MS and utilizing mass spectral libraries. Currently available libraries for analytical toxicology were compared for overlapping and uniqueness of their entries. Furthermore, the widely known Pfleger-Maurer-Weber-Drugs-and-Pesticides-Library for toxicology (PMW_tox2) was used to compare the search algorithms PBM (Probability Based Matching, Agilent Technologies), INCOS (Finnigan/Thermoquest), and MassLib (Max Planck Institute). To our knowledge, direct comparisons of mass spectral libraries and search programs for analytical toxicology have not been published previously. The capabilities and necessities of modern MS technology in the field of general unknown analysis are revealed, and some of the potential pitfalls are described.

Introduction

Screening techniques for unknown compounds in the forensic sciences often involve the combination of gas chromatography and mass spectrometry (GC–MS). Mass spectrometry is the method of choice for general unknown analyses (GUA) or systematical toxicological analysis (STA) (1,2). In the early days, mass spectra were compared manually. The introduction of powerful and affordable computers enabled the continuous development of mass spectra search programs. With this progress in computer technology, these search programs were and still are continuously improved. Helpful features such as similarity searches, structure editors, and new databases were added. Global networks (e.g., Internet) connect computers around the world and enable fast information gathering and data exchange. The forensic scientist strongly benefits from this fast progress, but he or she is also confronted with new problems including, for example, the large amounts of electronic data.

Even under the GC–MS standardized ionization technique of

electron ionization (EI) with an electron energy of 70eV, somewhat differing ion abundances may occur when different mass spectrometers are used. This is quite often the case for MS libraries because they are usually compiled from different sources. These differences may directly influence the library search. Different parameters may affect relative ion abundances in a spectrum.

The nature of inlet, source, analyzer, and detector are considered the most important factors affecting ion abundances. The use of different analyzers systems results in different ion lifetimes before detection. Rearrangements and metastable decompositions may occur in the ion source or during the transfer of the ions to the detector. The temperature of the ion source may influence the ionization and fragmentation, hence the ion abundances. Ion sources with high background pressure (especially ion-trap mass analyzers) produce significant amounts of ions by partial chemical ionization (CI). High concentrations of analytes in the ion source promote self-CI. Shifts and nonlinearities of the mass scale may disturb the peak recognition and the centralization of data. An inadequate mass resolution highly affects the ion abundances. Inadequate scan time, detector saturation, low dynamics of an aged detector, mass discrimination of the complete system, intense chemical background, electronical noise, detector offset misadjustment, and more factors affect relative ion abundances. Therefore, it is very important to regularly check the instrument not only for sensitivity, but also for general system performance and validity of data.

The type and features of the unknown spectrum are other important considerations. Many pharmacologically active molecules produce very few abundant ions when ionized under electron impact (EI) ionization conditions. If, for instance, a nitrogen atom is present in a side chain, then the fragmentation will most likely occur on that side chain. The nitrogen will often localize the positive charge, and as a consequence, only the small fragment of the side chain is detected (as a positively charged ion). The rest of the molecule with no charge (a radical) remains invisible. Because of the fragmentation of the aliphatic side chain and the stabilization of the fragment with the positive charge, the molecular ion is most frequently low in intensity. This adds to the difficulties of searching for reference spectra within a library and makes the distinction of spectra of homologues diffi-

* Author to whom correspondence should be addressed: Dr. Beat Aebi, Institute of Legal Medicine, University of Berne, Buehlstrasse 20, CH-3012 Berne, Switzerland.
E-mail beat.aebi@irm.unibe.ch.

cult. An example that we frequently encounter is the low level detection of methadone. Methadone is widely used in Switzerland as a substitute drug for opiate dependency. The ionization of methadone produces a base peak at m/z 72. All other ions are of weak intensity, that is, in the low percent region. The molecular ion at m/z 309 has no significant intensity. Depending on the intensity of matrix underground, the detection of methadone at low levels can be rather difficult. Because of its chemical structure as a tertiary amine, methadone is not easy to derivatize. As described previously (3), the use of a nitrogen-phosphorus specific detector in a parallel mode strongly enhances both the qualitative identification and the subsequent quantitation. An intense molecular ion with several typical intense fragments certainly makes the interpretation and comparison of mass spectra easier. However, reality is often different. One possible way to shift the fragmentation to higher masses is chemical derivatization (e.g., perfluoroacylation, silylation). Chemical derivatization also enables the analysis of more polar and thermically unstable molecules (e.g., hydroxy metabolites). An overview of derivatization procedures used in toxicological analyses was recently published by Pfleger et al. (4).

The quality of the library spectra is obviously very important. When creating or extending a mass spectral library, all new entries must be carefully checked. Wrong or incomplete reference data may lead to false-positive (wrong identification) or false-negative (no corresponding spectrum found) results. The analyst should be aware of this fact. Whenever possible, the pure standard compound should be analyzed to obtain a reference spectrum using identical conditions as the unknown. By this, artifacts and degradation products can also be recognized as such. The procedure to evaluate library reference spectra differs from evaluating an unknown spectrum. Several questions arise when a new spectrum should be integrated into an existing library or if an existing library entry needs to be evaluated.

Data consistency (without spectrum): Are the name, molecular formula and mass, and the structure consistent?

Spectrum consistency: Can the spectrum be verified against a reference of another previously verified library? Is the molecular ion present? Do no ions exist above the molecular ion cluster? Is the molecular ion cluster in accordance with the ion ratios calculated from the molecular formula (check for identity and detector signal offset)? Are the major fragment masses in accordance with the chemical structure? If no, are these fragment masses plausible when compared with spectra of similar compounds?

Is the mass range of the spectrum adequate to show the fragmentation of the compound? Mass spectra can differ strongly when an inadequate mass range is selected. Typical examples are found in amphetamines. Two different mass spectra of 4-bromo-2,5-dimethoxyamphetamine (DOB) are presented in Figure 1. The lower spectrum was acquired with a mass range starting too high (m/z 50). As a result, the base peak of m/z 44 was not recorded. The subsequent library search will in fact most likely end without a correct result, even if a (correct) reference spectrum exists in the used library. The upper spectrum was properly recorded from m/z 40. The spectroscopist must be aware of the fact that the mass ranges of the acquisition and the library must be compatible. In the new revision 3 of the Pfleger-Maurer-Weber-Drug-and-Pesticide-Library (PMW_tox3), the mass spectra of such compounds are also presented in the mass range from m/z 40 in addition to the standard mass range from m/z 50 (4,5).

Spectrum purity: Is the mass spectrum from a pure compound? Has the spectrum been checked for interfering signals, solvents, column bleed, degradation products, and artifacts? Is the detector signal offset correctly adjusted? Is the spectrum distinguishable from similar spectra originating from different compounds? If all questions can be answered by yes, the spectrum is of acceptable quality. If one or several answers are no, the evaluator should mark the spectrum (e.g., unconfirmed, proposed). When the spectrum quality is very low, it should not yet be entered into the library, but the fact still remains that a poor quality reference spectrum is of higher value than no spectrum at all.

Experimental

Description of the software on the 5972 MSD

The GC–MS instrument from Agilent Technologies (formerly Hewlett-Packard, Palo Alto, CA) used in this system was described previously (3). The system is working under Microsoft Windows NT 4.0. The Chemstation software (including data analysis) is G1701BA (version B.00.00). The MS search is using the probability-based matching algorithm PBM, developed by Fred McLafferty at Cornell University (6). The PBM algorithm is identifying the most significant peaks by a combination of their mass-to-charge ratio values and abundances and compares them with the condensed spectra of the library. This prefilter is usually eliminating
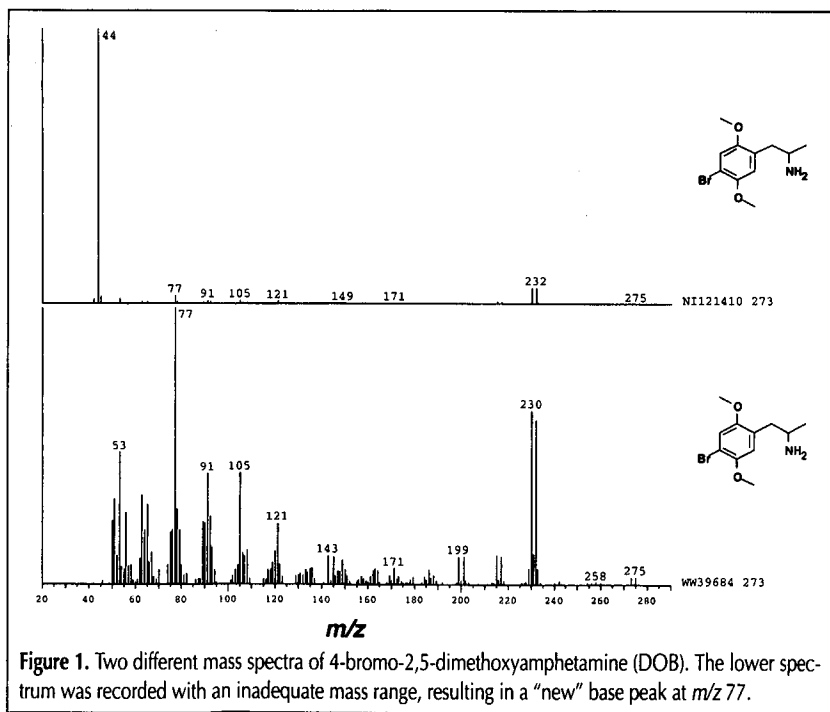

Figure 1. Two different mass spectra of 4-bromo-2,5-dimethoxyamphetamine (DOB). The lower spectrum was recorded with an inadequate mass range, resulting in a "new" base peak at m/z 77.

approximately 95% of all spectra, leaving 5% of the total number of the library entries for direct comparison and ranking for the hit list. If more than one library (maximum of three) is selected, the algorithm will check the following libraries accordingly and prepare a continued hit list from all selected libraries. A minimum match quality percentage can be used to weight the selected libraries. The libraries can also be searched manually for data using the parametric retrieval option. The search parameters for the PBM search were used as (strongly) recommended by the manufacturer: significance (uniqueness U and abundance A) U+A = 2, tilting = on, flag threshold = 3, cross-correlation sort = off, minimum estimated purity = 50%, low report MW = 0, high report MW = 9999, maximum hits = 20, remove duplicate CAS numbers = off. The values recommended have been carefully chosen by the manufacturer, and thus no additional experiments on further optimization have been performed.

Description of the software on the Finnigan SSQ7000 (3)

The SSQ7000 (Finnigan/Thermoquest Corp., San Jose, CA) is controlled by a DEC 3000 alpha-workstation (Compaq Corp., Houston, TX) with Tru64 UNIX 4.0E and the Finnigan software ICIS 8.3.0 SP. The library search is the INCOS algorithm-based program LIBR (7). The searches using the Pfleger-Maurer-Weber-Drug-and-Pesticide-Library (version 2) (8) were performed on the Finnigan SSQ7000 at the Institut für Gerichtliche Medizin (Vienna, Austria). Presearch and main search parameters have been used as recommended by the manufacturers and developers of the search program. No additional experiments on variation of the recommended values have been performed.

Description of MassLib

The mass spectra analysis program MassLib (version V8.6-B5) (MSP, Koeniz, Switzerland) is run on the SSQ7000 alpha-workstation and on a PC under Windows 95 (MassLib version 8.6-F). The search algorithm SISCOM (search for similar and identical compounds) originates from Henneberg, Weimann, and Ziegler (Max-Planck-Institute, Muehlheim a.d. Ruhr, Germany) and was developed in the 1970s (9). Originally optimized for the identification of isomers of aliphatic and aromatic hydrocarbons from petroleum products, the algorithm was extended for the search of identical mass spectra. The search is based on characteristic ions. The occurrence of matches, non-matched intensities of characteristic ions, and pattern correlations are considered. MassLib provides neutral loss and structure similarity searches in addition to SISCOM. In case of an unknown spectrum with very few abundant ions besides an intense base peak, MassLib offers an identity search, which is optimized for such spectra. Additionally, the MassLib system has been extended to perform INCOS searches according to the original literature (7).

Description of the libraries

The specific mass spectral libraries for toxicology used in this work are presented. Because of the transformation of the data into the MassLib-library format, the total number of entries of the libraries may slightly differ from its original number.

Commercial libraries. To obtain and evaluate mass spectra and compile all the data into libraries, is certainly a very resource consuming and expensive task. In order to honor the work per-

formed by others and enable the future development of such valuable databases, its users definitely must comply with legal regulations. As with any other software product, copyright laws protect commercial libraries. The unauthorized use, reproduction, and distribution are not only against the law, but also impair the interests of the scientific community as data collection and publishing come to a stop without adequate indemnification. The Pfleger-Maurer-Weber-Drug-and-Pesticide-Library, version 2 (PMW_tox2) (10) contained 4367 spectra, including 2000 metabolites and 800 pesticides. The new version (PMW_tox3) (4,5) contains 6350 entries, including the data on new drugs, poisons, and their metabolites, as well as of new types of their derivatives like methylated, trimethylsilylated, and perfluoroalkylated compounds. As already mentioned, the mass spectra of amphetamine derivatives were additionally depicted beginning from m/z 40. Both versions only contain reduced spectra. The Finnigan Toxicology Library with 2211 spectra, including 70 metabolites and 150 artifacts, was used on the SSQ7000 with the Finnigan operating system for the INCOS-searches and the PC system for the MassLib-search. The forensic library of MassLib (version 1) with 2275 spectra, including 300 metabolites and 400 artifacts, originates from the work of Dr. Rolf Kuehnle of the German Landeskriminalamt Niedersachsen. A library of designer drugs and other new drugs of abuse from DigiLab Software GmbH (Scharnhagen, Germany) with 440 entries was used. This library originates from the work of Dr. Giselher Fritschi and coworkers of the German Landeskriminalamt Hessen. Several other libraries such as the Wiley 6th edition and the NIST 98 library are also in use. However, in the last two, mostly the IUPAC names are used instead of the INN.

Noncommercial libraries. The library of the MS committee of the American Academy of Forensic Sciences AAFS (version 3) contains 1716 spectra (several spectra of the same compound) with additional natural compounds, herbicides, and pesticides (11). The toxicology library of The International Association of Forensic Toxicologists (TIAFT) includes 204 spectra (after conversion into the MassLib format) of silylated compounds of forensic interest (12).

Results

Combination of libraries

The daily work has shown that the combination of libraries is not only beneficial because of increased efficiency but also because of increased quality. The search results of several libraries are ranked and thus can be directly compared. An example of a hit list resulting from a search of several libraries is shown in Figure 2. The spectra of the tricyclic antidepressant trimipramine show significantly different intensities for the molecular ion at m/z 294 in relation to the base peak of m/z 58. The intensity of m/z 249 also varies considerably. The second spectrum shows additional minor signals at m/z 120 and 137. Different ion abundances are problematic for identity-based search algorithms and generally when performing general unknown analyses at low concentrations.

Overlapping of libraries

Four toxicological libraries have been tested for common entries by comparison of the name and molecular formula of each entry, using the lists of entries of the four libraries. The mass spectra themselves were not used for comparison. For the AAFS Library, only version 1 with 1550 entries was tested (Table I). Unsurprisingly, the four toxicology libraries had many entries in common. Widely used medications and drugs of abuse were present in all libraries, whereas rarely encountered medications, alkaloids, metabolites, artifacts, and derivatized compounds might be present in only one library.
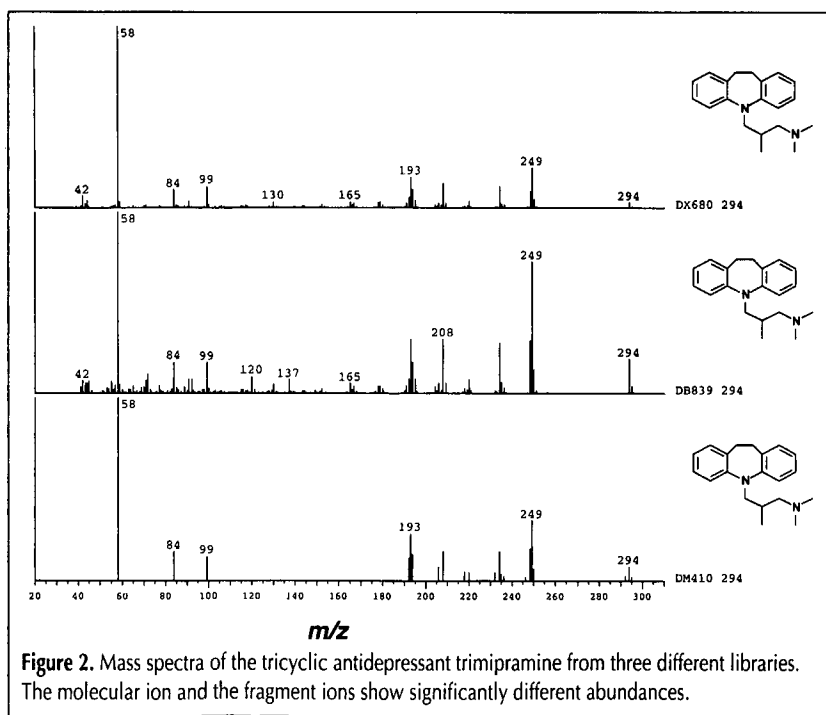
An estimation of the number of unique spectra for each library is also of interest (Table II). Using the MassLib system, every entry of each of the four libraries was searched in the other three libraries. The libraries were searched spectrum by spectrum, and the search results of each evaluated and validated by visual inspection. Counting of unique and nonunique spectra was done by hand. To perform this task efficiently, the option of interactive searching of multiple spectra (in blocks of 100 spectra) was used. With this option, the user can select multiple spectra from either a chromatogram or a mass spectral library and let them process one after the other. Three modes of display (result list, spectra list, structure list) can be switched at the end of each search. After clicking the continue button, the next spectrum is searched, and the results are again displayed. It is thus unnecessary to load each spectrum manually. The criteria for unique spectra were as follows: spectrum completely different (mass-to-charge ratios not matching at all); spectrum somewhat different (few not matching mass-to-charge ratios with > 10% absolute intensity); and spectrum slightly different (assigned name, molecular mass or formula different).

Spectra with different mass ranges were counted as unique when ions of absolute abundances greater than 10% were present in the differing mass range. Spectra with four or fewer significant ions were counted as unique when no appropriate molecular formula and chemical name were found.

It must be stated that the results of Tables I and II can only be an estimation of the real number of spectra overlapping and the real number of unique spectra. The widely known PMW_tox2 library (10) contained not only the highest number of spectra but also the highest percentage of unique spectra. The new version (PMW_tox3) is expected to contain even more unique spectra and also to overlap more with the other libraries. The addition of the unique spectra of all four libraries yields a total of about 6000 unique spectra. The AAFS MS-libraries do contain several entries of the same compound. The spectra were obviously acquired under different experimental conditions. These additional spectra increase the probability of searches for identity and similarity. As mentioned before, mass spectra of the same compound may have somewhat different ion intensities. As these differences are mostly due to complex mechanisms of the ionization process, they usually cannot be predicted. In our opinion, it is of great value to have multiple entries for the same compound.

Table I. Overlapping of Library Entries

	da*	df	dm	dx
da	1550	507	363	381
df		2211	503	592
dm			4367	778
dx				2275

Abbreviations: da, AAFS-Library, version 1, 1550 spectra; df, Finnigan Toxicology Library, 2211 spectra; dm, PMW_tox2, 4367 spectra; and dx: Toxicology Library of MassLib, version 1, 2275 spectra.

Table II. Comparison of Unique Spectra

Percent of Total Number		Library	Version	Total Number of Spectra
259	16.7%	AAFS-Library	version 1	1550
987	44.6%	Finnigan Toxicology Library		2211
3287	75.3%	PMW_tox2		4367
1243	54.6%	Toxicology Library of MassLib	version 1	2275

Figure 2. Mass spectra of the tricyclic antidepressant trimipramine from three different libraries. The molecular ion and the fragment ions show significantly different abundances.

Comparison of the three search programs

All the screenings (general unknown analyses) for forensic toxicology performed during the years 1998 and 1999 were re-checked for difficult or formerly unknown mass spectra. Of these spectra, 16 were selected and classified as follows: type A, spectrum present in the PMW_tox2 library but with considerably different relative ion abundances; type B, spectrum present in PMW_tox2 library with strong background signals; or type C, spectrum not present in the

PMW_tox2 library.

The spectra of the types A, B, and C were searched with the HP ChemStation, the Finnigan INCOS, and the MassLib software, all using the PMW_tox2. The HP ChemStation library was transformed into the MassLib library format only for testing purposes. The tests with the INCOS reversed-search algorithm were performed using the Finnigan PMW_tox2-library on the Finnigan SSQ 7000 at the Institut für Gerichtliche Medizin in Vienna, Austria.

Summary of the results (Table III)

Type A. All three systems performed well. Spectrum no. 3 gave very low hit rates on the MassLib and the ChemStation systems. It was only found at hit position 7 with the INCOS reversed search. Some of these examples are of caffeine. Caffeine is a frequently found compound and proved to be well suited to investigate on different ion abundances. The spectra were selected at different positions on the chromatographic peaks of several chromatographic runs. The spectra were used without background subtraction or spectra addition. A trained spectroscopist could easily obtain the pure spectra of caffeine. Our intent, however, was to use non-optimum spectra to test the search algorithms.

Type B. All three spectra were correctly assigned by the three search algorithms. For both types of searched spectra (A and B), the MassLib system was able to produce acceptable results even with the use of reduced spectra.

Type C. This type of test was used to show the capabilities of the search algorithm to find similar mass spectra. Because of the different search algorithms, different results were obtained.

Spectrum no. 11. For the compound proadifen, both the INCOS and the MassLib algorithms found the structurally related compound adiphenine. The ChemStation search listed a metabolite of etafenone as best hit (within the PMW_tox2-library). Reference spectra are present in the AAFS-libraries, versions 1 to 3, and in the Finnigan library. The structures are shown in Figure 3.

Spectrum no. 12. The novel antidepressant venlafaxine (Figure 4) contains an aliphatic ring with a hydroxy group at position 1. The INCOS search has found a compound with these chemical characteristics (Tramadol) as seventh hit, whereas MassLib found another compound with these characteristics as first hit (cyclopentolate, a parasympatholytic). The mass spectra of this new drug and their metabolites are included in the new
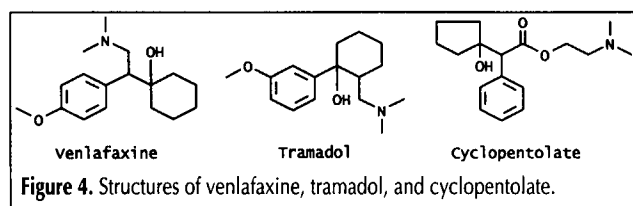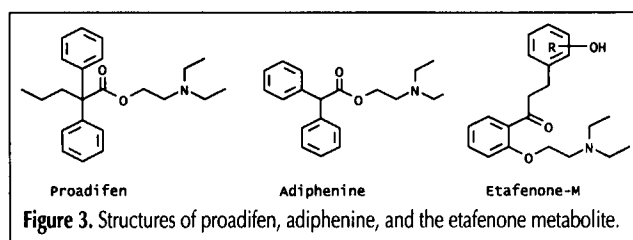

Figure 3. Structures of proadifen, adiphenine, and the etafenone metabolite.

Figure 4. Structures of venlafaxine, tramadol, and cyclopentolate.

Table III. Results of the Searches for Compounds 1–16

Spec. No.	Type	Comments	INCOS Reversed Search rank match (1...1000)		MassLib rank match (1...100)*		HP Chemstation rank match (1...100)	
1	A	Caffeine	1	329	1	(26)*	1	59
2	A	Caffeine, low ab.†	1	325	1	(25)*	1	76
3	A	Caffeine, low ab.	7	266	1	(17)*	1	14
4	A	Caffeine, low ab.	1	447	1	(28)*	1	52
5	A	Trimipramine	1	767	1	(45)*	1	91
6	A	Thioridazine	1	688	1	(29)*	1	50
7	A	Diazepam	1	658	1	(27)*	1	96
8	B	Nicotinamide	1	603	1	(80)*	1	94
9	B	Methaqualone	1	423	1	(34)*	1	96
10	B	Diphenhydramine	1	672	1	(79)*	1	72
			1st hit, match		**1st hit, match**		**1st hit, match**	
11	C	Proadifen (SKF-525)	Adiphenine, 720		Adiphenine, (56)*		Etafenone-M, 59	
12	C	Venlafaxine	Mexiletine AC, 570		Cyclopentolate, (34)*		Lidocaine-M, 38	
13	C	Tricresylphosphate	2,4,5-T octylester, 466		Tetradecane, (18)*		Brofaromin-M, 9	
14	C	Disulfiram-artifact (?)	Disulfiram, 341		Disulfiram, (19)*		Etiroxate-Art., 52	
15	C	Citalopram	Chlorprothixene-M, 547		Chlorprothixene-M, (43)*		Promocarb, 47	
16	C	Thioridazine-M (?)	Thioridazine-M, 428		Mepivacaine, (24)*		Thioridazine-M, 55	

* The results of the MassLib searches were obtained from the PMW_tox2 library with spectra reduced according to the PBM algorithm. These results are not based on a normalized parameter and hence no significance lies in the absolute magnitude. The generally low values shown reflect the fact that the reference library used contained non-complete (reduced) spectra, leaving many searched characteristic ions unmatched.
† Abbreviations: ab., abundance; AC, acetylated; -M, metabolite; -Art., artifact.

PMW_tox3, as well as in the AAFS-libraries (all three versions).

Spectrum no. 13. For this compound, tricresylphosphate, no adequate matches were found. The mass spectrum of this compound contains a strong ion cluster at *m/z* 368. All other ions are only of low intensity. The PMW_tox2-library used for comparison does contain the structurally related compound triphenylphospate with a strong ion cluster at *m/z* 326. But the molecular ion cluster and the other ions of minor intensity do not match those of tricresylphosphate at all. Reference spectra for this compound were found in the AAFS-libraries (all three versions).

Spectrum no. 14. The identity of this compound could not be confirmed. In this forensic case, the alcohol deterrent disulfiram and several related compounds (artifacts and metabolites) were located and confirmed. Because of the other findings and the correlation of spectrum no. 14 to the compound disulfiram itself, we strongly suspect it to be an artifact or metabolite. This spectrum is present neither in the PMW_tox2-library nor in the any of the other databases that we currently use. The library used in this comparison (pmw_tox2) does not contain other spectra of compounds structurally similar to disulfiram. The new PMW_tox3-library however does contain other metabolites and artifacts of Disulfiram. Both the INCOS and the MassLib systems found the original compound Disulfiram as first hit, although with low hit rates.

Spectrum no. 15. The mass spectrum of the novel antidepressant citalopram is shown in Figure 5. Citalopram contains an aromatic ring with an oxygen containing 5-membered ring attached. Additionally, the base peak at *m/z* 58 is very common. The molecular ion and several fragment ions are present, although with intensities below 20% relative to the base peak of *m/z* 58. The PMW_tox2-library does not contain any structurally related compounds. In such a case, the identification of the compound will not be possible without further information, either from the case itself or from other spectroscopic methods. The

spectrum is now included in the PMW_tox3 library as well as in the AAFS library version 3.

Spectrum no. 16. This spectrum could not be definitively confirmed. The antidepressant thioridazine was also detected and confirmed in this analysis of gastric content, so we suspect this spectrum to be from a hydrolysis product or artifact of thioridazine. For the INCOS and MassLib searches, several thioridazine-related hits were obtained.

Spectra nos. 11–16. The search of an unknown spectrum that is not present in the library is a difficult task. The PMW_tox2-library in combination with the PBM-algorithm of the Chemstation software clearly focuses on the search for identical spectra. For this purpose, reduced spectra were used. The PMW_tox2-library used with the MassLib systems also contained only reduced spectra. Because of the lack of ions with minor intensities, the search for similarity loses efficiency and reliability. Still, several similar mass spectra could be found. A further in-depth visual inspection for similarities is difficult, when no ions of minor intensity are available. The reduction of mass spectra according the search algorithm (PBM or INCOS) follows mathematical rules. The presence of signals with minor intensities in the spectrum is of course due to the fragmentation process. When comparing mass spectra of structurally related compounds or trying to establish a link between different mass spectra, ions of very low intensities may contain much more information than ions of high intensity.

Discussion and Conclusions

Currently available GC–MS libraries for toxicology were presented and compared. The four libraries compared showed different degrees of overlapping. The widely known PMW_tox2-library with the highest total number of entries also showed the highest percentage of unique spectra. With the use of the Finnigan and the MassLib toxicology library, as well as the non-commercial AAFS and TIAFT libraries, the number of unique spectra can be further increased. The newly published PMW_tox3-library was not used for the library statistics. It is to be expected that this new version overlaps even more with the other tested libraries. It must be noted, however, that the use of libraries of different manufacturers requires either a data transformation or the use of a format and system independent software. In either case, the copyright laws must be applied. We have experienced that multiple entries of the same substance increased the chance of a positive identification, especially when (apart from the base peak) only signals with low intensities were present. However, a risk is present with the use of multiple entries: a great number of similar spectra may push the correct entry down in the hit list. Whenever multiple hits of the same compound are found, the hit list must be checked accordingly.
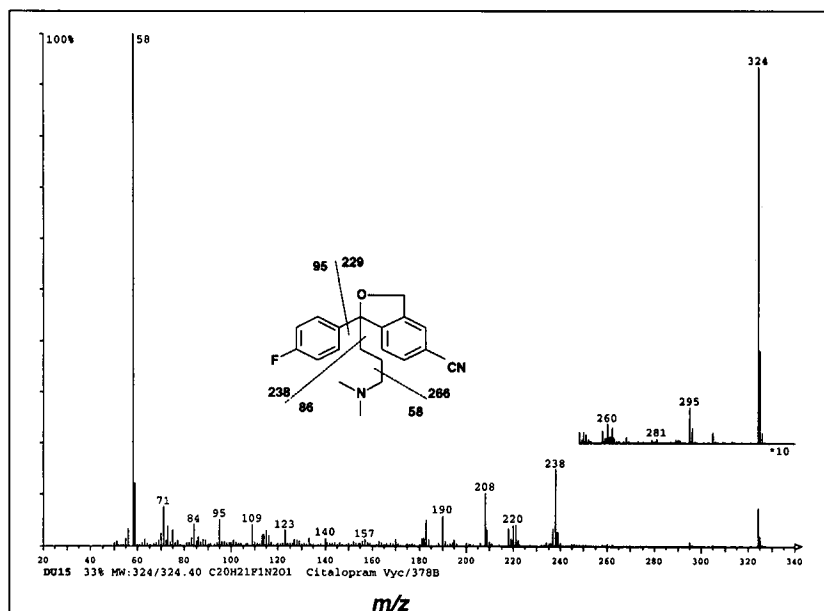
Figure 5. Mass spectrum and structure of the novel antidepressant citalopram (MW 324 Da). MassLib contains a structure editor with a built-in fragment editor. When a spectrum is loaded, the corresponding structure is automatically displayed.

The comparison of the three search algorithms presented in this work did not exhibit significant differences when the unknown spectrum is present in the library (identity search). The search for spectra that were not yet present in the reference library (PMW_tox2) gave different results. The INCOS system was originally built for the search of similar mass spectra. The MassLib system, however, was originally designed to search characteristic ions to find similar structures even with dissimilar spectra. Reduced mass spectra were also successfully used with the similarity-oriented system such as MassLib both for identity and similarity searches. Because the amount and quality of information obtained from library searches are directly dependent on the quality of the library spectra, complete mass spectra of high quality are recommended. Further work on this topic will be undertaken, including the newest versions of the PMW_tox and AAFS libraries. The intent of the proposed work was to show the capabilities of three different search algorithms to handle unusual and difficult as well as spectra that were not present in the reference library. The search parameters of the search programs were used as recommended by the manufacturers and developers of the search programs. It cannot be excluded that in some very special cases, the changing of these parameters could further improve the search quality.

When no spectrum of high or even medium similarity is present, the spectroscopist must perform a more thorough evaluation of the unknown spectrum. He or she must then rely on the possibilities offered by the search program. The spectroscopist will try to establish a link between the unknown spectrum and selected library spectra. Pharmacologically active substances usually have either "common" names or "chemical" names, mostly following INN, IUPAC, or CAS nomenclature. Additionally, not all libraries do contain the Chemical Abstract Registry Number (CAS-RN). Thus, the comparison of data may become very time consuming. The use of structures to support spectra interpretation is very helpful.

The ease of use (user-friendliness) of the program is also important. The author (B.A.) has experienced all three systems and found them to be very well suited for use. To define or "quantitate" user-friendliness is a difficult task because it strongly depends on what the user is expecting from the system.

Both the ChemStation software and MassLib support automatic library searches of an entire chromatogram. The user of the ChemStation software can select summarized or detailed outputs on either the screen, the printer, or as a text file. MassLib offers an interactive step-through the entire analysis or an automatic output on the computer screen or printer. The time required to perform a search for identity using the reference library (PMW_tox2) was less than or in the order of a second for all three systems, depending on the complexity of the spectrum and the number of similar spectra in the library. Because the speed for spectra searches is further highly depending on hardware (e.g., processor type and speed, memory size and speed, hard disk size and speed, graphics speed, and network and peripheral connections) and software factors (operating system, network system), this topic was not investigated in this work. All three systems compared used high-quality search algorithms that can give satisfactory results in a very short amounts of time. If the time spent for searching mass spectra in a library is very

important, then the system should of course have a high-quality data system that meets the requirements. This may include the use of high-end parallel (e.g., dual) processors with separate high-throughput graphics processors. On the other hand, the slowest but most important "part" is still the spectroscopist. The evaluation of complex spectra is a difficult task and may be quite time consuming. The knowledge and routine of a skilled spectroscopist can still not be outmatched by the most sophisticated technical equipment. But by bringing together skilled personnel and adequate equipment, time consumption can be reduced, increasing both productivity and quality.

Data analysis can sometimes be further accelerated by the means of additional programming, for example, macro programming with the Agilent/Hewlett-Packard ChemStation software (13,14) or the Thermoquest/Finnigan ICIS software. A macro-based method for enhanced data treatment was published by Polettini (15).

The MassLib system is independent of the type and format of the data to be analyzed. Most mass spectrometric data formats can be imported with the integrated filters. Additional programming is in this case not necessary. Identity, including automatic background correction, similarity, similarity, and identity (preferred use); neutral loss; and few peak searches are some of the options offered by MassLib. General data and structure searches can be performed. To efficiently use MassLib, advanced knowledge of mass spectrometry is of prime importance. We routinely use MassLib for analytical method development and for interpretation of difficult and unknown spectra. Additional tools, such as structure search, fragmentation visualization, and neutral loss searches, have definitely been helpful so far.

Every year, thousands of new compounds are discovered worldwide. Chemical Abstracts Services (CAS, www.cas.org) currently lists over 16 million organic and inorganic compounds and almost 7 million protein sequences. Many new compounds will be tested as drug candidates in the pharmaceutical industries. A total of 120,000 registered medications for Europe, the U.S., and Japan are estimated by the pharmaceutical industry of Switzerland (personal communication). According to the molecular weight table published by Pfleger et al. in 1992 (10), the total number of known and possibly toxic compounds can be (at that time) estimated to about 8000. To this, many new metabolites, artifacts, and degradation/hydrolysis products as well as illicit synthetic drugs, by-products, and precursors are continuously discovered and identified. Furthermore, the consumption of pharmacologically active substances is often nationally and geographically different. This will define the areas of prime interest of analytical laboratories and experts. For a very fast spreading of information of new compounds, the distribution of non-commercial libraries via the Internet seems to be very helpful. Two libraries of this type were used in this work. Both the AAFS (11) and TIAFT (12) have started to build mass spectral libraries for the forensic community. In addition to this very positive development, the analyst must still be aware of the development of new libraries and computer software. Many different compounds have very similar mass spectra. In combination with the technically imposed minor differences in ion abundances, the distinction of these spectra is sometimes difficult. With the current, very fast development of computer tech-

nology, computers are capable of handling large mass spectral libraries, even with multiple entries of the same compound. The financial investment in powerful computer equipment capable of enhanced mass spectra processing is low compared to the analytical system itself. The mass spectrum is the product of the entire analysis. Sufficient care should be taken in this matter.

The chances of finding a reference spectrum increase with the total number of spectra in the libraries. The decision as to which spectrum is the most convincing one cannot be made by the computer. A successful identification of an unknown compound should nevertheless be confirmed either by the analysis of reference material (if available, identification based on chromatographic and spectroscopic comparison) or by an independent second method. The fact that good quality libraries will produce good quality search results cannot overrule the need for careful spectra examination by an experienced analyst and the overall examination and evaluation of the entire case by the forensic expert.

Acknowledgments

References

1. H.H. Maurer. Systematic toxicological analysis procedures for acidic drugs and/or metabolites relevant to clinical and forensic toxicology or doping control [review]. *J. Chromatogr. B* **733**: 3–25 (1999).
2. A. Polettini. Systematic toxicological analysis of drugs and poisons in biosamples by hyphenated chromatographic and spectroscopic techniques [review]. *J. Chromatogr. B* **733**: 47–63 (1999).
3. B. Aebi and W. Bernhard. Gas chromatography with dual mass spectrometric and nitrogen-phosphorus specific detection: a new and powerful tool for forensic analyses. *Forensic Sci. Int.* **102**: 91–101 (1999).
4. K. Pfleger, H.H. Maurer, and A. Weber. *Mass Spectral and GC Data of Drugs, Poisons, Pesticides, Pollutants and their Metabolites,* 2nd ed., part 4, Wiley-VCH, Weinheim, Germany, 2000.
5. K. Pfleger, H.H. Maurer, and A. Weber. *Mass Spectral Library of Drugs, Poisons, Pesticides, Pollutants and their Metabolites,* 3rd rev. Agilent Technologies, Palo Alto, CA, 2000.
6. D.B. Stauffer, F.W. McLafferty, R.D. Ellis, and D.W. Peterson. Probability-based-matching algorithm with forward searching capabilities for matching unknown mass spectra of mixtures. *Anal. Chem.* **57**(6): 1056–1060 (1985).
7. S. Sokolow, J. Karnofsky, and P. Gustafson. The Finnigan library search program. *Finnigan Application Report* **2** (1978).
8. K. Pfleger, H.H. Maurer, and A. Weber. *Mass Spectral Library of Drugs, Poisons, Pesticides, Pollutants and their Metabolites,* 2nd rev. Finnigan, San Jose, CA, 1992.
9. H. Damen, D. Henneberg, and B. Weimann. SISCOM—a new library search system for mass spectra. *Anal. Chim. Acta* **103**(4): 289–302 (1978).
10. K. Pfleger, H.H. Maurer, and A. Weber. *Mass Spectral Library of Drugs, Poisons, Pesticides, Pollutants and their Metabolites,* 2nd ed., parts 1–3. VCH, Weinheim, Germany, 1992.
11. Mass Spectral Database Committee. mslib, http://www.ualberta.ca/~gjones/mslib.tm, 2001.
12. TIAFT. User contributed collection of EI mass spectra, http://www.tiaft.org/main/mslib.html, 2001.
13. H.H. Maurer. Toxicological analysis of drugs and poisons by GC–MS. *Spectroscopy Europe* **6**: 21–23 (1994).
14. H.H. Maurer and J. Bickeboeller-Friedrich. Screening procedure for detection of antidepressants of the selective serotonin reuptake inhibitor type and their metabolites in urine as part of a modified systematic toxicological analysis procedure using gas chromatography–mass spectrometry. *J. Anal. Toxicol.* **24**: 340–347 (2000).
15. A. Polettini. A simple automated procedure for the detection and identification of peaks in gas chromatography–mass spectrometry. Application to systematic toxicological analysis of drugs in whole human blood. *J. Anal. Toxicol.* **20**: 579–586 (1996).